Breast Cancer Classification as Malignant or Benign based on Texture Features using Multilayer Perceptron

Nassir H. Salman [1], Semaa Ibrahim M.Ali [2]

Computer Science Department, College of Science, Baghdad University, Baghdad, Iraq.
[1] nassir.s@yahoo.com, [2] Semaa.ibrahim@muc.edu.iq

Abstract - **Breast cancer is one of the most common malignant diseases among women. Mammography is at present one of the available methods for early detection of abnormalities, which is related to breast cancer different lesions to show characteristics such as masses, which can be detected through this technique. The images are divided according to the Mini-MIAS database. In order to classify the cluster as malignant or benign, the 2nd order co-occurrence statistical properties of the image such as contrast, correlation, homogeneity, and energy were adopted. The gray-scale convergent matrices (GLCM) are used with a suggested feature of gray level density matrices (GLDM) to identify abnormal tissues (malignant) and natural tissue (benign). The proposed method of analysis is tested on several images (taken from a UK organization interested in understanding breast X-ray images) from the database of digital mammograms and outstanding results were obtained for breast cancer detection. The proposed multilayered design performance significantly improved the diagnosis of breast cancer by more than 95% and 95.8% sensitivity, 95.5% specificity for all datasets, 91.1% with 100% sensitivity and 84% specificity for 70% training data and 30% testing data.**

Keyword - Breast cancer, digital mammogram, a texture feature, classification, multilayer perceptron.

I. INTRODUCTION

Currently Computer Aided Detection (CAD) is gaining popularity for the detection of breast cancer by analyzing mammograms. This practice is being done in many screening sites and hospitals around the world. This fact shows that CAD may be successfully used around the globe to detect and diagnose several problems present in medical images. These images are being collected from different examinations by exploiting several types of imaging modalities. Research shows that CAD studies have gain popularity among researchers in the field of medical imaging and diagnostic radiology. Research on CAD started in the 1960s; however, proper systematic investigation on CAD began in 1980s. In this era, the dominant conceptual change was about the output results, which transformed from automated computer diagnosis to computer-aided diagnosis [1]. Medical images proved to be helpful in treating and aiding diagnosing. These images are also useful in a medical student to understand their studies more appropriately. With the recent growth of digital imaging technologies, medical images helped medical students in their studies concerning analysis and interpretation [2]. These processes of imaging happened to be successful because of computer vision and pattern recognition technologies. The CAD system is becoming more and more popular among many medical researchers and clinical centers to help physicians in diagnosing and treating diseases like breast cancer [3]. In medical imaging, interior view of the body can be created and processed for clinical analysis. These digital images are used to expose structures hidden under the skin and bones of the human body. This technology dramatically helps in diagnosing and treating the ailment. Furthermore, these images are kept safe for further analysis of different health issues and may be used for future references.

Due to medical reasons, imaging done for detached organs and tissues. This imaging process is considered under pathology studies [4]. Several medical applications are now dependent on image analysis techniques and procedures. Relevant information, to diagnose and treat the ailment, is retrieved through automatic feature extraction from medical images. This information may be used for several classification tasks like differentiating normal healthy tissues from abnormal tissues. The extracted information may include morphological properties, specific textural properties and color properties of the image [5].

II. METHODOLOGY

In the current research work, radiation is passed through the tissues in these processes. The image of the anatomical structure is projected on the film screen or image sensor. X-ray imaging causes reduction of anatomical 3D organ information to 2D film/image. It is stated in the research [9] that if women, of age 50 or greater, have the regular mammogram, there are greater chances of reducing the number of breast cancer patients. Low-frequency X-rays are, which pass through the compressed breast, are recorded by mammograms. Captured images are categorized concerning the X-rays viewpoint. Two main viewpoints generally used in the process of extraction of mammograms: (i) Medio-lateral Oblique view (MLO) (ii) Cranio-Caudal (CC) view. The important point to be considered is that MLO views have one region conforming to a portion of pectoral muscle that we used in this research. Due to privacy

concerns, it is more challenging to acquire real clinical images for analysis. Mammographic Image Analysis Society (MIAS) has provided us the data for our analysis [6]

This studies painting is executed at the MATLAB R2014a picture processing toolbox in figure (1) firstly, mammogram image pre-pressing then extraction the Region Of Interest by appropriate equation then enhancement image by adaptive histogram equalization.

III. A PROPOSED NEW TECHNIQUE

Different types of artifacts in the background, noises as well as pectoral muscles, etc. are contained in the original mammogram images. These are all unwanted features concerning feature extraction as well as classification. For this reason, the mammogram image has been subjected to a cropping operation. This enables the extraction of Regions of Interest (ROIs) which are containing the abnormalities. This is beside the unwanted parts of the original image. All the possible details related to each of the mammogram images, e.g. background tissue related character, size in pixels, abnormality class, etc. are given by the MIAS database. These include 'r' radius of a circle enclosing the abnormality and the Xc and Yc coordinate value of the center of abnormality. Manual cropping operation is utilized for ROI extraction. This is done in consideration of the abnormal area center as the center of the Region of Interest ROI. The approximate radius of the circle taken in pixels is then considered which encloses the abnormal area as displayed in Figure (1). The rectangular ROIs, as well as the ROIs which are free from artifacts and background information, are extracted in this phase. In this particular phase, the rectangular Regions of Interest (ROIs) are extracted, and these ROIs are deemed free of the artifacts and background information. This is defined by the following equation (1):

$$IROI=I\ [xc-r,\ 1024-y,\ 2r,\ 2r] \qquad 1$$

Image resizing for uniformity can then be accomplished as128 x 128, 512 x512, and 256 x 256. ROI were resized to 256 x 256 as part of this work.
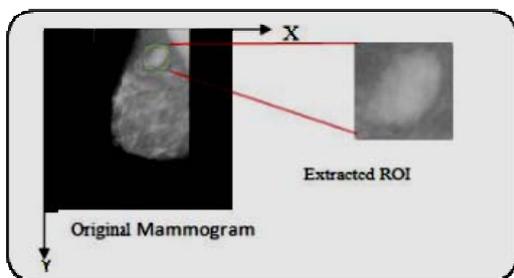

Figure 1. Cropping of ROI

The cropped part of the A mammogram image might still contain a lot of noise, and also the quality of the image might be quite poor. In order to force the image to become suitable for classification, a preprocessing step is required most of the time. At first, the image is converted from RGB to Grayscale, then the noise is removed by applying the median filter technique, and the utilization of the enhancement technique further improves the image quality. Preservation of information and concurrently removal of noise is ensured by using the median filter in this work. A process termed the Contrast Limited Adaptive Histogram Equalization (CALHE) is then applied to enhance the pixel contrast concerning its immediate localized neighborhood adaptively. This also results in improving the appearance of the image contrast. Figure (3) shows the block diagram of the proposed detection method based database classification, while Figure (2) displays the results of this process utilized for the enhancement of contrast.
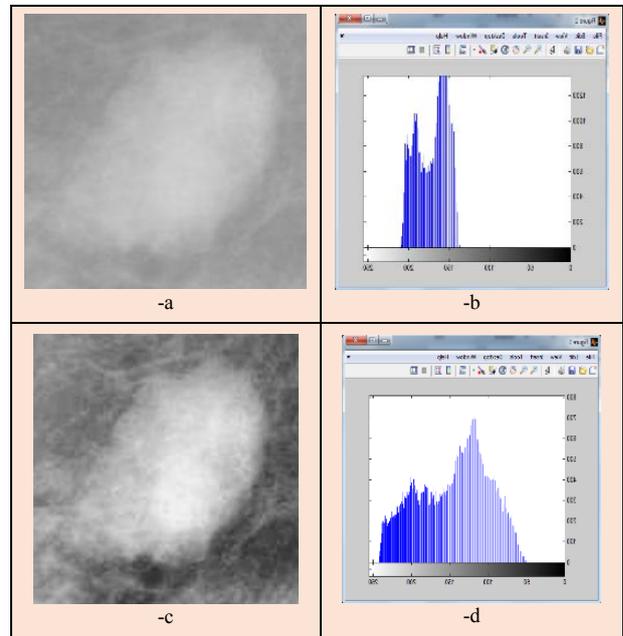

Figure 2. (a) De-noised image; (b) Histogram of image (a); (c) Contrast-enhanced image (d) Histogram of image (c)
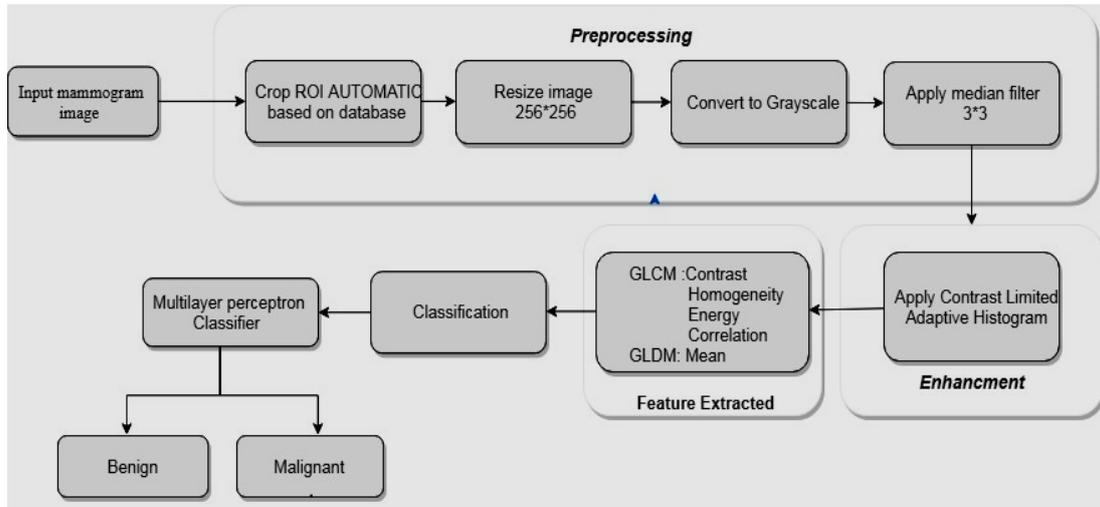
Figure 3. Block diagram of Proposed Detection Method Based on Database

IV. FEATURES EXTRACTION

When the amount of data entered for processing is too large and is subject to repetition (i.e. the amount of data is large versus the amount of less information), the entered data can be converted to a set of attributes that are less representative than the original data known as the feature vector. If these features are selected carefully, the relevant information from the input data is expected to be extracted by the attribute set so that the task requirements are met. For obtaining these features, first crop the masses from an entire image so that features can be extracted of only that region which is required, that has been done by using GLCM and GLDM.

A. Texture Feature

The limitation constrains the employment of histograms to compute measures of texture-based features that they do not possess any information on the position of pixels concerning one another. The texture analysis can make use of this information through the consideration of not just the distribution of intensities but also the pixel positions with either equal or almost equal intensity values. Feature extraction of this type is accomplished from gray-level co-occurrence matrix.

The probability of a pixel with a gray level of i occurring in terms of a particular spatial relationship to another pixel with a value of j can always be computed with GLCM, The size of the GLCM is determined by how many of the gray pixels are found in the image. In this thesis, the calculation of GLCM feature descriptor is done in four angles (0°, 45°, 90°, 135° degrees). Four texture features (Correlation, Contrast, Homogeneity, and Energy) and displacement distance d=1; thus, the implementation has 16 texture measurement features.

GLDM is used for extraction of features, which are used to figure out the likelihood density capabilities for a given or known picture. For a given, image power function I(x,y) as well as the displacement denoted as $\delta = (\Delta x, \Delta y)$, the value of probability density function f(i|) is estimated from the number of time modify in strength $I\delta(x,y)$, take place for a known or given displacement d, i.e., f(i| δ)=p(I$\delta(x,y) = i$).

In this case, the four possible forms of inter-sample spacing distance d were considered, and these were :(0, d), (d, 0), (-d,d), and (-d,-d). GLDM calculated in four possible forms. The four descriptors extracted from GLDM texture measurement include the mean of each descriptor for inter-sample spacing d =5.

It is often convenient to normalize the co-occurrence matrices so that they approximate discrete joint probability distribution of co-occurrence grey levels this is accomplished by dividing each entry in the matrix by the total number of paired occurrences:

$$P_N = \frac{1}{N_\theta} P_\theta \qquad 2$$

Haralick et al [65] have proposed a variety of measures that can be employed to extract useful information from matrices. Only four of these 2nd order measures have been adopted in this thesis and have been used to classify between malignant and benign breast tissue[9].

A1. Contrast (Moment of Inertia)

The measure of relative intensity between pixels and their neighbors over the complete image is termed as contrast, which is shown in Eq. (3). It is the local variation extent existing in a certain image.

$$f1 = \sum_{n=0}^{Ng-1} n^2 \left\{ \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} p(i,j) \right\} |i-j| = n \qquad 3$$

A2. Correlation

$$f2 = \sum_{i=1}^{Ng-1}\sum_{j=1}^{Ng-1}\frac{(i,j)p(i,j)-\mu_x\mu_y}{\sigma_x\sigma_y} \qquad 4$$

μ_x, μ_y And σ_x, σ_y are defined as following:

$$\mu_x = \sum_{i=1}^{Ng-1}\sum_{j=1}^{Ng-1} i.p(i,j) \qquad 4a$$

$$\mu_y = \sum_{i=1}^{Ng-1}\sum_{j=1}^{Ng-1} j.p(i,j) \qquad 4b$$

And

$$\sigma_x = \sqrt{\sum_{i=1}^{Ng-1}\sum_{j=1}^{Ng-1} i^2.p(i,j) - (\mu_x)^2} \qquad 4c$$

$$\sigma_y = \sqrt{\sum_{i=1}^{Ng-1}\sum_{j=1}^{Ng-1} i^2.p(i,j) - (\mu_y)^2} \qquad 4d$$

Where μ_x, μ_y and σ_x, σ_y are the rows, columns-sum means and standard deviations respectively computed from the marginal densities $p^x(i)$ and $p^y(j)$ respectively. The marginal densities are given as:

$$p^x(i) = \sum_{j=0}^{G-1} p(i,j) \qquad 4e$$

$$p^y(j) = \sum_{j=0}^{G-1} p(i,j) \qquad 4f$$

Correlation is the measure of connection of pixel with its neighbors in complete image, which is depicted in eq(4). The range of feature values is from -1 to 1. -1 represents perfect negative while perfect positive correlation is represented by 1[9].

A3. Energy (Angular Second Moment):

$$f3 = \sum_{i=1}^{Ng-1}\sum_{j=1}^{Ng-1}\{p(i,j)\}^2 \qquad 5$$

Energy is also termed as the angular second moment (ASM) or uniformity. The total of squared elements present in GLCM is measured by energy, shown in eq.(5). The uniformity of texture image is provided by energy. [0 1] is the energy range, where the value of energy is 1 in case of a constant image[9].

A4. Local Homogeneity

$$f4 = \sum_{i=0}^{Ng-1}\sum_{j=0}^{Ng-1}\frac{p(i,j)}{1+(i-j)^2} \qquad 6$$

The proximity of elements distribution from GLCM towards GLCM diagonal is measured by homogeneity, which is described in Eq. (6). [0 1] is the homogeneity range, where for diagonal GLCM, its value is 1. Its value is high in case of high inverse GLCM and uniform local gray level[9].

B. Gray Level Difference Matrix (GLDM)

The Gray Level Difference Matrix process carries out the calculation for probability density functions of gray level difference method, for the pre-processed gray images. The GLDM method is derived from the occurrence of two pixels having a provided absolute gray level difference, been separated by a precise displacement. For a given image, the probability density functions have been calculated by the GLDM method. The mathematical equation that describe mean shown below that is a new feature to classification breast cancer.

$$\text{Mean} = \mu = \frac{1}{N}\sum_{i=1}^{N} PDF(i) \qquad 7$$

Assuming that PDF (i) is a probability density form value for gray level, i and N denote the number of gray levels in the image [9].

Features in 0 direction					
Cancer type	Mean pdf1 E+04	Correlation	Energy	Homogeneity	Contrast
Malignant	6.51	0.9771	0.3239	0.9816	0.0383
Malignant	6.51	0.969	0.4709	0.9886	0.0247
Malignant	6.50	0.9958	0.3052	0.9794	0.0587
Malignant	6.46	0.9827	0.1888	0.9706	0.0596
Malignant	6.48	0.9911	0.1535	0.9742	0.0523
Benign	6.53	0.9812	0.186	0.9693	0.0621
Benign	6.53	0.9947	0.2268	0.9884	0.0241
Benign	6.51	0.987	0.4006	0.9885	0.0243
Benign	6.53	0.9779	0.2798	0.9826	0.0359
Benign	6.49	0.9884	0.2231	0.9821	0.0368

Features in 45 direction					
Cancer type	Mean pdf2 E+04	Correlation	Energy	Homogeneity	Contrast
Malignant	6.53	0.9706	0.318	0.9758	0.0492
Malignant	6.53	0.9584	0.4666	0.9841	0.0326
Malignant	6.53	0.9951	0.3024	0.9746	0.0683
Malignant	6.49	0.9772	0.1817	0.9612	0.0781
Malignant	6.49	0.9876	0.1471	0.9639	0.0729
Benign	6.53	0.0851	0.9578	0.1778	0.9741
Benign	6.52	0.0418	0.9793	0.2209	0.9907
Benign	6.49	0.0277	0.9865	0.3995	0.9851
Benign	6.52	0.0549	0.9728	0.2687	0.9662
Benign	6.46	0.043	0.9788	0.2207	0.9864

Features in 90 direction					
Cancer type	Mean pdf3 E+04	Correlation	Energy	Homogeneity	Contrast
Malignant	6.51	0.9802	0.3252	0.984	0.0334
Malignant	6.52	0.9646	0.47	0.987	0.0279
Malignant	6.50	0.9984	0.3087	0.9891	0.0223
Malignant	6.47	0.9805	0.1857	0.9669	0.0671
Malignant	6.48	0.9902	0.1518	0.9716	0.0577
Benign	6.53	0.0721	0.9643	0.1826	0.9781
Benign	6.53	0.0394	0.9807	0.2216	0.9912
Benign	6.50	0.0224	0.9894	0.4006	0.988
Benign	6.53	0.0485	0.9763	0.2729	0.9702
Benign	6.48	0.0393	0.9809	0.222	0.9876

Features in 135 direction					
Cancer type	Mean pdf4 E+04	Correlation	Energy	Homogeneity	Contrast
Malignant	6.52	0.9716	0.3191	0.9767	0.0474
Mali0gnant	6.53	0.953	0.4635	0.9821	0.0368
Malignant	6.51	0.9953	0.3032	0.9763	0.0648
Malignant	6.49	0.9725	0.1761	0.9531	0.0941
Malignant	6.50	0.9869	0.1462	0.9618	0.0767
Benign	6.51	0.0961	0.952	0.1734	0.9708
Benign	6.52	0.0476	0.9764	0.219	0.9894
Benign	6.48	0.0349	0.9828	0.3961	0.9812
Benign	6.52	0.0631	0.9687	0.2644	0.9611
Benign	6.45	0.0604	0.97	0.2136	0.9809

V. CLASSIFICATION USING MULTI-LAYER PERCEPTRON (MLP)

This algorithm has been used in many breast tumor research studies or other medical images, whether to classify images into benign and malignant images, multi-layer perceptron (MLP) net are used to perform the desired classification task. Three layered artificial neural networks are used in this experiment. It has an input layer as well as a hidden and output layer. Figure 4 shows features that used as input in the multilayer perceptron neural network and their connection with the hidden layer and the output layer.
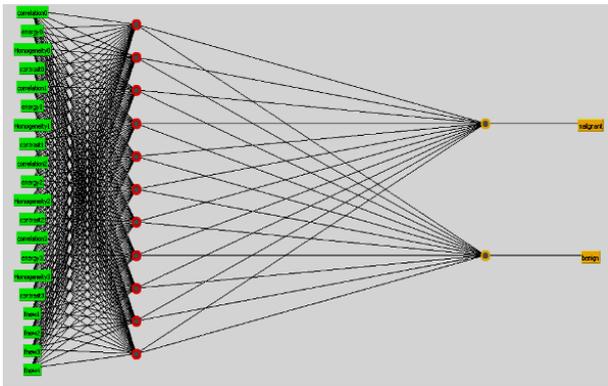

Figure 4. Multilayer Perceptron Neural Network

As part of the classification process, the phases involved are the training phase and the testing phase. As a part of the training phase, in the presence of known data which is given, the classifier is trained. In contrast to this, unknown data is available in the testing phase, and the trained classifier is used to perform classification. The neural classifier is then fed with the chosen features as input. The neural network uses a hidden layer as well. The output layer uses a node, and this is trained as a presentation of the malignant as well as the benign. The problem is very efficiently solved by the machine learning Weka with the multilayer perceptron classifier and gives us good results.

Weka (Waikaato Environment for Knowledge. Acquition) is a collection of machine learning algorithms using the JAVA-language for data mining tasks. The algorithms are applied directly to a dataset. WEKA-.implements algorithms for data preprocessing, classification, regression, clustering, and association-rules; It also includes visualization tools. The data file used by Weka is in ARFF file format(Attribute Relation file-format (.arff)), which consists of special tags to indicate different things in the data file (foremost: attribute names, attribute types, attribute values and the data). The main-.interface in Weka is the Explorer. It has a set of panels, each of which can be used to perform a particular task. Once a dataset has been loaded, multilayer perceptron can be used to perform further analysis. WEKA applications are currently used to classify different mammogram images, to process a medical dataset and identify the types of breast cancer.

A. Classification Evaluation

The supervised classifiers are employed with known data or tags. The classifier processes using these classifiers are categorized into a test phase and a training phase. In the training phase, the system has been trained regarding the identification of cancer and normal cases. During the phase of testing, the performance of the system is checked by providing input test images for calculation of system accuracy by using known cancer and normal images. Due to the limited available data during training, the use of auxiliary data testing systems becomes significant [13].

However, the problem is the usage of limited data during testing and training. The improved system design is the result of using more data during the training process. Moreover, the usage of more data during the tests helps in more reliable system evaluation.[12]

According to the successful evaluation system obtained on the training sets, there is a threat of data memory and over-optimized error rate. In order to avoid memory problems, the evaluation of a system should be done on individual data sets that are not used during the training process. To do this, one method is dividing the data into two disjoint sets and using them in training

and testing the system. K-fold cross-validation can be used when most data is not available as a test set.

Since the capability of future systems for working on unidentified samples should be measured in the testing phase. The testing process should also involve those samples which are not relevant to used examples in training. The performance of classifier's detection is quantitatively measured through the calculation of specificity and sensitivity of data. Before the description of specificity and sensitivity, the confusion matrix has been shown in Figure (5), in which four probable results of the assessment are illustrated. The outcomes of tests which are performed with perfect specificity and sensitivity will be TN or TP and never FN or FP [8].

Figure 5. Confusion Matrix

Generally, an image region can be termed as normal (negative) or cancerous (positive), and the decision based on a detection outcome can be either true (correct) or false (incorrect). Therefore, for any detection result, the decision will be one of the following possible options:

- False positive (FP)
- False negative (FN)
- True positive (TP)
- True negative (TN)

FP and FN are two types of errors.
- False Positive: The system of diagnosis indicates test report as positive for a sample, but it does not have that disease
- False Negative: The investigative system produces a negative result of a test for a sample. However, the disease is essentially present in the sample
- True Positive: The disease is present in the sample, and the same is indicated by the diagnostic system as well
- True Negative: Negative test result is yielded by the diagnostic system, and the sample also doesn't have the disease in actual

The specificity and sensitivity are defined with the help of samples falling in this category, for the assessment of diagnostic system success. Sensitivity is termed as a conditional probability for cancer detection when it is actually present in the image. Using the samples belonging to these classes, the sensitivity is determined for the privacy of the evaluation of diagnostic system success.

$$\text{Sensitivity} = \frac{TF}{TF+FN} \qquad 8$$

The conditional probability in the case of normal breast detection whilst it is having a true normal state termed as specificity:

$$\text{Specificity} = \frac{TN}{TN+FF} \qquad 9$$

In terms of FP and FN rates it can be shown as:

$$\textit{Sensitivity= 1 – false negative rate} \qquad 10$$

False Negative Rate: is defined as the likelihood where the classification result indicates a normal breast, but the actual diagnosis is positive for breast disease. It is must to avoid such case as it is dangerous from patients' point of view.

False Positive Rate: is defined as the probability of indication of breast disease by the classification result when the real diagnosis is without a doubt a usual breast presence, i.e. negative result.

On the other hand, precision or positive predictive value is defined as the proportion of the true positives against all the positive results (both true positives and false positives).

$$precision = \frac{TF}{TP+FP} \qquad 11$$

Accuracy of a Result The accuracy of a result can be quantified by calculating the percent error. The percent error can only be found if the true value is known. Although the percent error is usually written as an absolute value, it can be expressed a negative or positive sign to indicate the direction of error from true value.

$$\% \text{ Error} = \frac{(\text{true value} – \text{experimental value})}{\text{true value}} \times 100 \qquad 12$$

For each image , compute 2nd order feature from Gray-Level Co-occurrence Matrices to extract texture feature (Correlation, Homogeneity , Contrast, Energy) In four direction (0 , 45 , 90 , 135), and suggested feature (mean(μ) from Gray-Level Difference Matrices In four direction (0 ,45 ,90 , 135). The evaluation of the proposed system based on dividing the dataset into two parts (training and testing). The training images in this system consist of two classes (malignant, benign); each class consists of a number of images. The total number of images is (115), the proposed system will be tested under three modes, and these modes include (50%, 60%, and 70%). The mode 70% means that 70% of the dataset was used for training phase and 30% of the dataset was used for the testing phase, where (81) images are selected for training phase and (34) images are selected for the testing phase. The Accuracy of this stage shown in the table (1). Figure (6) visualizes the process results of mode 70% under WEKA program. Figure (7) showed the accuracy of each type of breast cancer under mode 70%.

TABLE I. RESULTS OF MODE 70%

	No. of Training Images	No. of Testing Images	Recognized		Classification Rate in %
			yes	no	
Malignant	15	33	15	0	100%
Benign	19	48	16	3	84.2%

```
=== Evaluation on test split ===
=== Summary ===

Correctly Classified Instances          31              91.1765 %
Incorrectly Classified Instances         3               8.8235 %
Kappa statistic                         0.8235
Mean absolute error                     0.1426
Root mean squared error                 0.2822
Relative absolute error                 28.8582 %
Root relative squared error             55.939  %
Total Number of Instances               34

=== Detailed Accuracy By Class ===

            TP Rate   FP Rate   Precision   Recall   F-Measure   ROC Area   Class
            0.938     0.111     0.882       0.938    0.909       0.955      malignant
            0.889     0.063     0.941       0.889    0.914       0.955      benign
Weighted Avg.  0.912  0.085     0.913       0.912    0.912       0.955

=== Confusion Matrix ===

  a  b    <-- classified as
 15  1 |  a = malignant
  2 16 |  b = benign
```

Figure 6. MLP Classifier Results under WEKA with Mode of 70% with GLCM   GLDM Feature
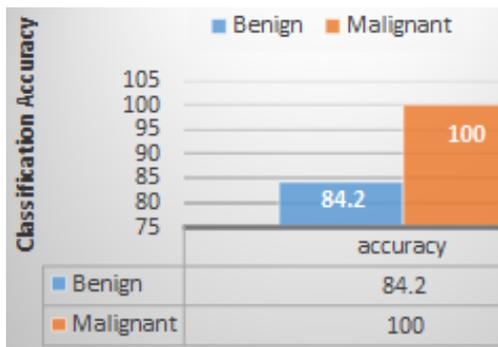

Figure 7. Chart of the Accuracy of Breast Cancer Types under WEKA with Mode of 70%
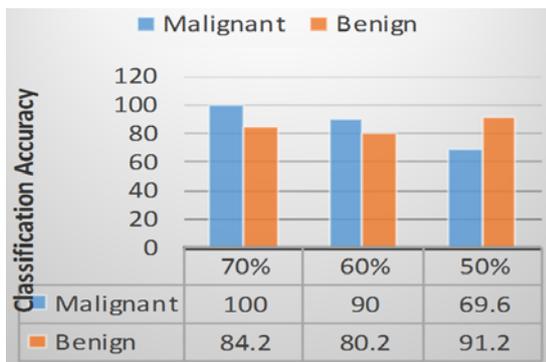

Figure 8. Chart of the Accuracy of Breast Cancer Types under WEKA with Mode of 50%, 60% and 70% with GLCM   GLDM Feature

B. Receiver Operating Characteristic (ROC) Curve

The space of possible tradeoffs between sensitivity and specificity are summarized in a Receiver Operating Characteristic (ROC) curve. The ROC curve is typically plotted with the TPF (sensitivity) on the Y-axis and the FPF (1- specificity) on the X-axis. Thus, the ROC curve produced allows the detection of massive lesions with predictable performance. In the ROC, the ideal operating point is the upper left corner, with TPF=1 and FPF=0. An indicator of the resulting ROC area of the class is after the prediction process. However, to understand the concept properly, we get different ROC area value in the Classifier output for each class. Each instance in the ROC curve has a threshold value of the class. If the instance highly belongs to this class, its threshold will be close to 1, so it will have a red color, the higher threshold of instance the darker color it will have in the ROC curve. Figure (9,10) shows a ROC curve for multi-layer perceptron classifiers, show the best results because curve is very close to the left top corner and then curves expand to the X-axis.
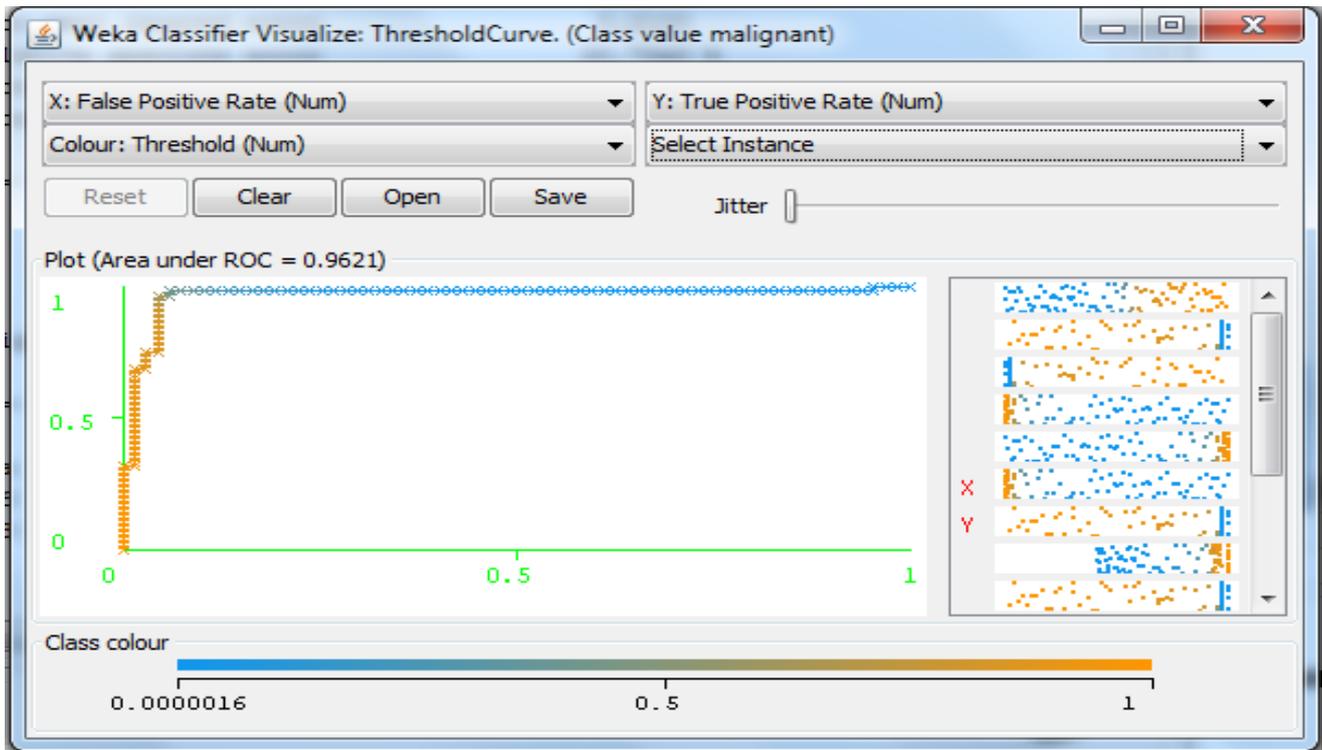
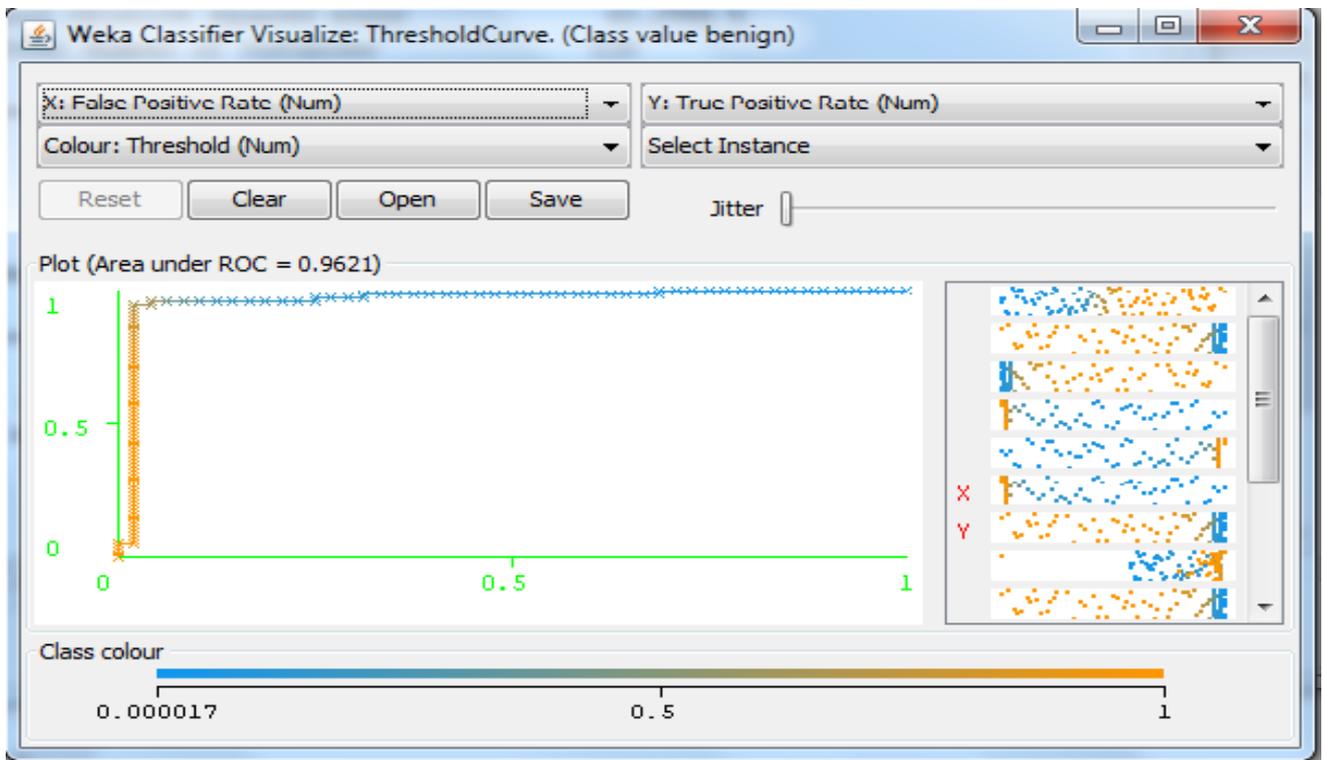Figure 9. ROC Curve for tested malignant Class using MLP Classifier

Figure 10. ROC Curve for tested_benign Class Using MLP Classifier

Figure 11 below shows the classification summary of our proposed system. As it shows the confusion matrix of the classifier, in addition to classification accuracy.
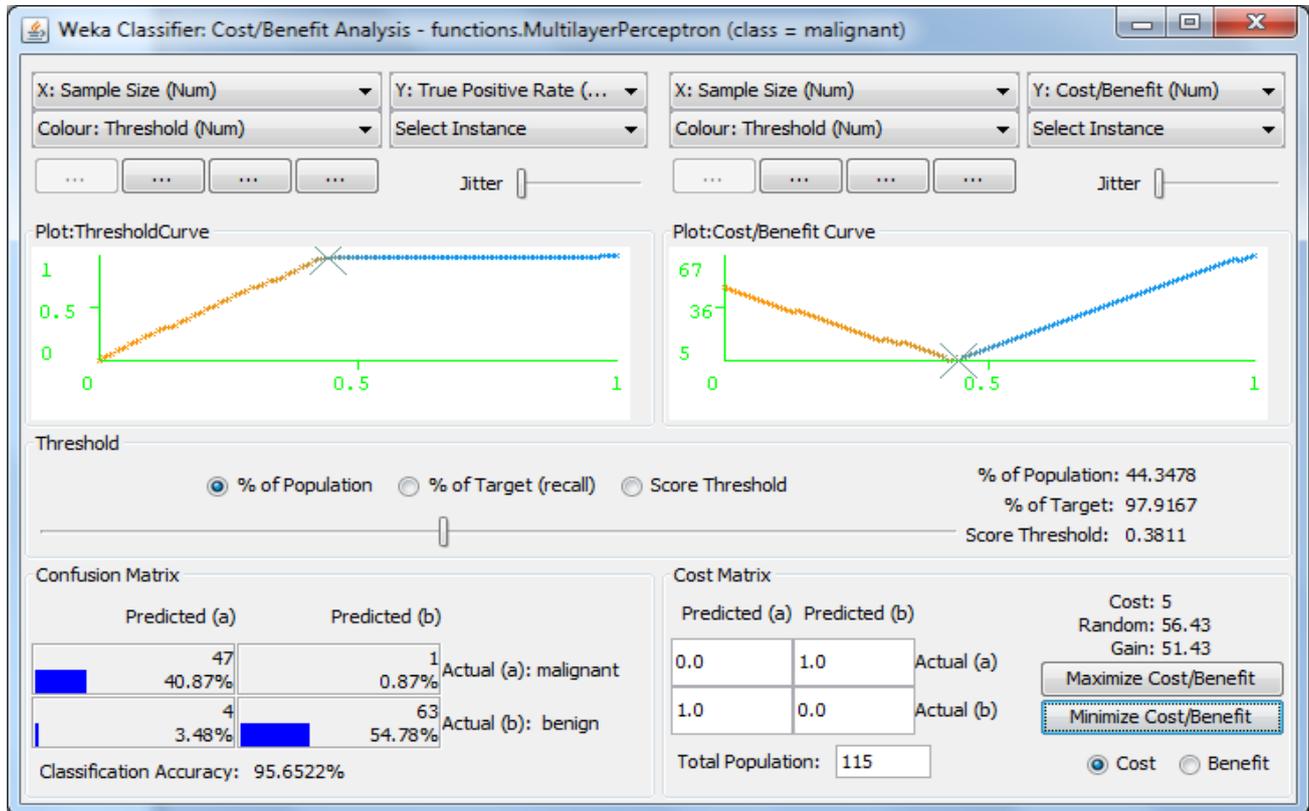

Figure 11. Cost Analysis of MLP with Minimum Cost.

C. Performance Comparison between Multilayer Perceptron Classifiers and other Classifiers

Finally, when comparing the results of breast cancer masses classification with mode 70% using multilayer perceptron under weka with J48, SVM (Support vector machine) under weka program with same data set that used for multilayer perceptron classifier. The multilayer perceptron has achieved the highest accuracy On the other hand, SVM toolkit had the lowest accuracy measures that reached to (64.7059 %) as seen in figure (12), Meanwhile, J48 exhibited (70.5882 %) accuracy when using mode of 70% as seen in Figure (13).display the differential between the multilayer perceptron classifier, SVM and J48 with mode 70%. These results showed that multilayer perceptron classifier under weka program tool is better than the other to be used for a classification task, this is may be due to the kind of data sets used, or maybe there are some differences in the way the algorithms were implemented within the tools themselves. Besides that, it has also exhibited high accuracy and speed when applied to an extensive database. Moreover, it is swift for both learning and predicting. Its learning time is linear in the number of examples and its prediction time is independent of the number of examples. The reason for good performance is value of MSE between the target and calculated output is reduced by using the iterative gradient technique, which employs a back-propagation algorithm. The initialization of the training process is done through establishing a few small random weights. The presentation of the neural network is done along with some training data, and the adjustments of weights can be made till the reduction of MSE up to a satisfactory level.

```
=== Evaluation on test split ===
=== Summary ===

Correctly Classified Instances          22               64.7059 %
Incorrectly Classified Instances        12               35.2941 %
Kappa statistic                          0.2817
Mean absolute error                      0.3529
Root mean squared error                  0.5941
Relative absolute error                 71.4491 %
Root relative squared error            117.771  %
Total Number of Instances               34

=== Detailed Accuracy By Class ===

              TP Rate   FP Rate   Precision   Recall   F-Measure   ROC Area   Class
              0.5       0.222     0.667       0.5      0.571       0.639      malignant
              0.778     0.5       0.636       0.778    0.7         0.639      benign
Weighted Avg. 0.647     0.369     0.651       0.647    0.639       0.639

=== Confusion Matrix ===

  a   b    <-- classified as
  8   8 |  a = malignant
  4  14 |  b = benign
```

Figure 12. SVM classifier results under weka with a mode of 70%

```
=== Evaluation on test split ===
=== Summary ===

Correctly Classified Instances          24               70.5882 %
Incorrectly Classified Instances        10               29.4118 %
Kappa statistic                          0.4056
Mean absolute error                      0.2888
Root mean squared error                  0.525
Relative absolute error                 58.4583 %
Root relative squared error            104.0786 %
Total Number of Instances               34

=== Detailed Accuracy By Class ===

              TP Rate   FP Rate   Precision   Recall   F-Measure   ROC Area   Class
              0.625     0.222     0.714       0.625    0.667       0.729      malignant
              0.778     0.375     0.7         0.778    0.737       0.729      benign
Weighted Avg. 0.706     0.303     0.707       0.706    0.704       0.729

=== Confusion Matrix ===

  a   b    <-- classified as
 10   6 |  a = malignant
  4  14 |  b = benign
```

Figure 13. J48 classifier results under weka with a mode of 70%.

VI. CONCLUSION

The possibilities of a hit final result of breast cancer commonly depend on its early detection and prognosis. This guide inspection Technique is gain time by Using 2nd order co-occurrence texture Homogeneity and Energy) from Gray-Level Co-occurrence Matrices with suggesting feature (Mean in four direction from GLDM gives the algorithm more robustness due to its resistance to many image situation variations that leads to best discrimination potential for identifying the type of masses. Multilayer perceptron classifier is also fast, consistent, easy to maintain and

accurate in the classification of attribute data. In addition, from the computation point of view, Multilayer perceptron is more efficient both in the learning and in the classification task as compared to SVM and J48 classifier. This system was implemented and gave very good results that provide accuracy 95.6% for all data set with sensitivity 95.8% and specificity 95.5% and accuracy 91.1% for 70% of data set with sensitivity 100% and specificity 84%.

REFERENCES

[1] K. Doi, "Computer-aided diagnosis in medical imaging : Historical review , current status and future potential," vol. 31, pp. 198–211, 2007.

[2] P. Smitha, "A Review of Medical Image Classification Techniques," no. Icvci, pp. 34–38, 2011.

[3] S. Dehghani and M. A. Dezfooli, "Breast Cancer Diagnosis System Based on Contourlet Analysis and Support Vector Machine," vol. 13, no. 5, pp. 1067–1076, 2011.

[4] M. T. L. S. P. Sjcit, "A Lossless Data Hiding Technique For Medical Images Using Lsb Substitution" Sunil Kumar M R 14th sem M . Tech (L . S . P) Sunil Kumar M R1 and Veena R2," pp. 1–6.

[5] M. Tuceryan, "Texture Analysis," pp. 1–41, 1998.

[6] J.Suckling,"The Mammographic Image Analysis Society Digital Mammogram Database ' ' Exerpta Medica," no. April, 2014

[7] A. M. Ahmad, S. A. Mahmud, A. M. Ahmad, G. Muhammad, S. A. Mahmud, and J. F. Miller, "Breast Cancer Detection Using Cartesian Genetic Programming evolved Artificial Neural Networks Breast Cancer Detection Using Cartesian Genetic," no. January 2014, 2013.

[8] Gonzalez, Rafael C., and Richard E. Woods. "Image processing." Digital image processing (2007).

[9] Haralick, Robert M., and Karthikeyan Shanmugam. "Textural features for image classification." IEEE Transactions on systems, man, and cybernetics 6 (1973): 610-621.

[10] C. Science and T. Khulna-, "Prediction of Breast Cancer Using Support Vector Machine and K-Nearest Neighbors," pp. 21–23, 2017.

[11] Sun Yajie, "Normal Mammogram Analysis," no. August, 2004.

[12] Sheba, K. U., and S. Gladston Raj. "An approach for automatic lesion detection in mammograms." Cogent Engineering 5, no. 1 (2018): 1444320.

[13] Biswas, Ranjit, Abhijit Nath, and Sudipta Roy. "Mammogram Classification Using Gray-Level Co-occurrence Matrix for Diagnosis of Breast Cancer." In Micro-Electronics and Telecommunication Engineering (ICMETE), 2016 International Conference on, pp. 161-166. IEEE, 2016

[14] Cheng, A.; Liu, H.; Zheng, B. Computer-aided classification of mammographic masses using the deep learning technology: A preliminary study. In Proceedings of the SPIE 9785, Medical Imaging 2016: Computer-Aided Diagnosis, San Diego,

[15] Qiu, Y.; Wang, Y.; Yan, S.; Tan, M.; Cheng, S.; Liu, H.; Zheng, B. An initial investigation on developing a new method to predict short-term breast cancer risk based on deep learning technology. In Proceedings of the SPIE 9785, Medical Imaging 2016: Computer-Aided Diagnosis, San Diego, CA, USA, 24 March 2016.