❒   121

Classification of mammograms based on features extraction techniques using support vector machine

Enas Mohammed Hussein Saeed[1], Hayder Adnan Saleh[2], Enam Azez Khalel[3]
[1,2]Department of Computer Science, Mustansiriya University, Baghdad, Iraq
[3]Specialist Radiology Department-Oncology Teaching Hospitals, Medical City Complex, Baghdad, Iraq

Article Info

ABSTRACT

Now mammography can be defined as the most reliable method for early breast cancer detection. The main goal of this study is to design a classifier model to help radiologists to provide a second view to diagnose mammograms. In the proposed system medium filter and binary image with a global threshold have been applied for removing the noise and small artifacts in the pre-processing stage. Secondly, in the segmentation phase, a hybrid bounding box and region growing (HBBRG) algorithm are utilizing to remove pectoral muscles, and then a geometric method has been applied to cut the largest possible square that can be obtained from a mammogram which represents the region of interest (ROI). In the features extraction phase three method was used to prepare texture features to be a suitable introduction to the classification process are first Order (statistical features), local binary patterns (LBP), and gray-level co-occurrence matrix (GLCM), Finally, support vector machine (SVM) has been applied in two-level to classify mammogram images in the first level to normal or abnormal, and then the classification of abnormal once in the second level to the benign or malignant image. The system was tested on the MAIS the mammogram image analysis society (MIAS) database, in addition to the image from the Teaching Oncology Hospital, Medical City in Baghdad, where the results showed achieving an accuracy of 95.454% for the first level and 97.260% for the second level, also, the results of applying the proposed system to the MIAS database alone were achieving an accuracy of 93.105% for the first level and 94.59 for the second level.

Corresponding Author:

Enas Mohammed Hussein Saeed
Department of Computer Science
Mustansiriya University
Baghdad, Iraq
Email: enas.saeed38@gmail.com

1. INTRODUCTION

Breast cancer is the most widespread disease that affects females during their lifetime, and the 2nd leading death cause for women globally [1] according to the most recent statistical data that has been released in 2015 in the USA, breast cancer occupies 29% of the new cases of cancer and 15% of the cancer death cases [2]. Early detection is the optimal option for increasing treatment options, as there are many screening methods for breast cancer detection such as biopsy, magnetic resonance imaging (MRI), ultrasound, and mammography [3]. Mammography is considered a common method for detecting abnormalities in the early stages. Mammography is a low-dose X-ray application allowing to visualize the inner breast structure [4]. It's usually contained many artifacts and noises that make them difficult to understand in the initial stages. Therefore, standardizing image quality and extracting region of interest (ROI) is necessary to reduce the search

for distortions [5]. Visual examination of mammograms by a radiologist to detect breast cancer very command but sometimes leads to less accurate diagnosis, due to the stress of the radiologist and low image quality. The studies indicate that the error rate of 10% to 30% for diagnoses of malignant masses by a radiologist who utilizes visual inspection. The error rate decreases significantly by utilizing the computer-aided detection (CADe) system, which provides the support of the final decision and is considered as a second opinion with radiology specialists diagnosis to classify breast tumors [6], [7]. Normally, there are five phases of the CADe system for breast cancer detection as can be seen in Figure. 1 [8].
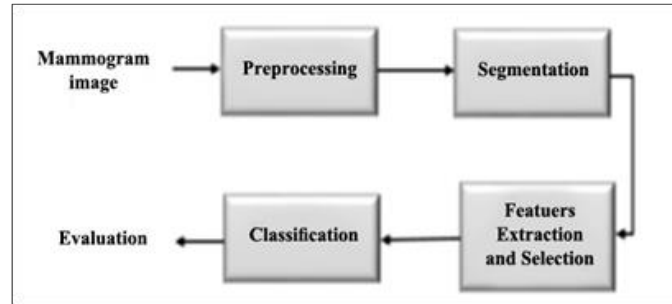

Figure 1. General phase for CADe system [8]

To achieve this goal, the medium filter was applied to remove noise, and local contrast, and then the binary image with a global threshold applied for small artifacts removal. In the segmentation phase, a hybrid bounding box and region growing (HBBRG) algorithm is utilizing to remove the pectoral muscles [9], and then we find the largest possible square area that can be obtained from the breast, which represents the ROI [10]. In the features extraction three feature extraction techniques were used are first Order, gray-level co-occurrence matrix (GLCM), and local binary patterns (LBP), to acquire strong texture features that were entered into the support vector machine (SVM) classifier at the classification phase. To conduct this search, investigate it, and evaluating all its phases, a special database was established that relies mainly on MIAS which has been proposed by the U.K. national program of breast screening. This database includes 322 digitized mammograms, 114 abnormal, and 208 normal which sub-dividing to 51 malignant and 63 Benign. A 1024x1024 pixel image with a "PGM" format [11]. The database is available on the website http://peipa.essex.ac.uk/info/mias.html [12]. Also, in coordination with the Teaching Oncology Hospital / Medical City / Baghdad, a set of images was obtained and added to the database. The pre-processing of these images has also been performed to be in the same mammogram image analysis society (MIAS) database image specification.

Theoretical consideration: Some many technologies and algorithms have been included in the stages of the computer-aided design (CAD) system in this part. We will look at some of the techniques that were used in our proposed method to extract and select features.

First order (statistical) features: Statistical texture analyses have been based upon the statistical characteristics of the intensity histogram with no consideration of the spatial dependences. The image histogram has given a statistical information summary concerning this image. The 1st order statistical image information may be produced with the use of an image histogram [13]. It is a group of the useful features that may be directly extracted from the spatial domain of the image histogram based on pixel values only, including mean, standard deviation (SD), variance, kurtosis, and skewness [14]. These features have been calculated using the following equations [15], [16]:

$$\text{Mean} = g \sum_{g=0}^{L-1} P(g) \tag{1}$$

where g is grey level in the image (0 to 255)

$$\text{Standard Deviation} = \sqrt[2]{\sum_{g=0}^{L-1} (g - \bar{g})^2 P(g)} \tag{2}$$

where $\bar{g}$ is mean of gray level in the image

$$\text{Variance} = \sum_{g=0}^{L-1}(g - \bar{g})^2 P(g) \tag{3}$$

$$\text{Skewness} = \sum_{g=0}^{L-1}(g - \bar{g})^3 P(g) \tag{4}$$

$$\text{Kurtosis} = \sum_{g=0}^{L-1}(g - \bar{g})^4 P(g) \tag{5}$$

LBP features: This texture descriptor is very interesting it is texture descriptor is considered very interestingly because particularly suitable for real-time quality controlling applications because it is fast and easy to implement. It was expanded for the color image by Maenpaa and Pietik¨ainen and used in numerous applications to classify problems based on color images [17], [18]. LBP indicates a relationship between a central pixel and its adjacent pixels in Micro pattern LBP examined window to cells (for example 8x8 and 16 x 16 pixels for each cell). or every one of the pixels in a cell, compression to all pixels of its eight neighbors (on the upper left, middle left, lower left, right top, and so on) as shown in Figure 2 which shows an example of the original operator of the LBP. LPB follow pixels along the circle, that is, counterclockwise or clockwise.
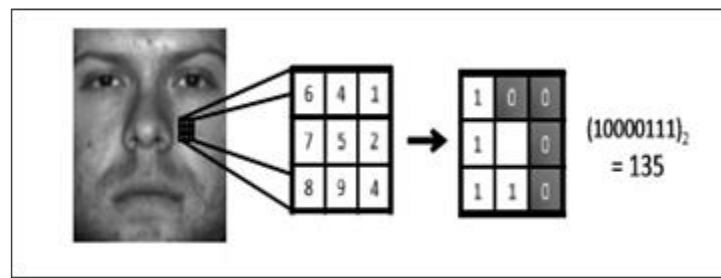

Figure 2. An example of the original operator of the LBP [19]

The string '10000111' is getting for 3*3 block with the central pixel 5. The binary form is transformed to its 135 duplicates in a decimal form. LBP histograms are created from all micro patterns depend on a decimal value. Assume that I is an image intensity and $r = (x, y)^T$ is a vector of position in I. The LBP b($r \in R^{Nn}$) is known as in the following description:

Bi(r) =1: if I(r) <I (r +Δ si), 0: otherwise, (i =1 Nn). Nn represents the number of the neighboring pixels, and Δsi is vectors of displacement from the position of center pixel r to neighboring pixels [19].

GLCM features: GLCM is also a statistical method that takes into account the spatial relationship between pixels. It has been employed widely in numerous applications based on texture analysis in the area of texture analysis [20]. The probability of a pixel with a gray level of I occurring in terms of a particular spatial relationship to another pixel j can always be calculated with GLCM. The size of the GLCM is located by how many pixels are found in an image. The GLCM feature is applied based on four angles (0°, 45°, 90°, 135° degrees) and displacement distance as shown in Figure 3 [21].In this paper, four features extracted from GLCM are energy, contrast, correlation, and homogeneity. Those features have been calculated utilized by the following (6).

$$\text{Energy} = \sum_{i,j=1}^{n} p(i,j)^2 \tag{6}$$

where n is the Co-Occurrence matrix dimension, p is the probability of GLCM

$$\text{Contrast} = \sum_{i,j=1}^{n}|i - j|p(i,j) \tag{7}$$

$$\text{Correlation} = \sum_{i,j=1}^{n} \frac{(i-\mu_i)(j-\mu_j)\,p(i,j)}{\sigma_i\sigma_j} \tag{8}$$

where μx, μy, σx and σy represent the mean and standard deviation values [22].

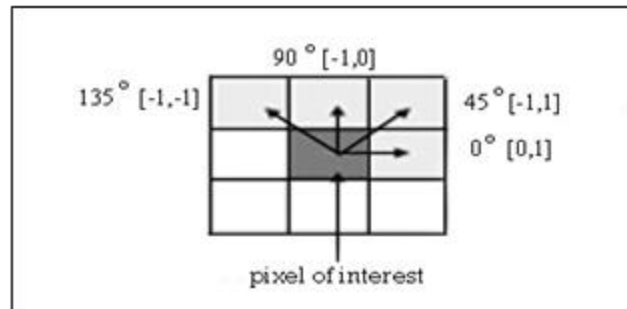$$\text{Homogeneity} = \sum_{i,j=1}^{n} \frac{p(i,j)}{1+(i-j)^2} \tag{9}$$

Figure 3. Show adjacency of pixel in four directions

Related works: Many researchers completed early to diagnose breast cancer in an attempt to help radiologists to detect abnormal tissue in the form of mammograms; we will review the most significant studies in this area below: In 2017 Harefa *et al.* applied to pre-process for improving the image quality, and then the segmentation stage has been applied depending on the database to obtain the ROI. Features are extracted by using GLCM at 0º, 45º, 90º, and 135º with a 128 x 128 block size. In the procedure of the classification, this study attempted at comparing the k-nearest neighbors (KNN) and SVM classifiers for achieving a higher level of accuracy. The result shows that SVM outperforms KNN in breast cancer abnormalities classification with 93.88% accuracy [23].

In 2018 Sheba and Raj, in the proposed methodology for the pre-processing median filter, was utilized for the noise filtering, global thresholding for removing the small artifacts [24]. BB is utilized for removing pectoral muscles, and adaptive fuzzy logic based bi-histogram equalization to enhance the quality of the mammograms for better perception. The ROI is automatically selected and segmentation from mammograms image with the use of morphological operations and global thresholding. Shape, GLCM, and texture features have been obtained from ROI, and then optimum features have been chosen with the use of classifier and regression tree (CART). Finally, the classification step has been carried out with the Feed-forward artificial neural network (ANN) utilizing the backpropagation. The proposed approach achieved 96% accuracy. In 2019 Mostafa *et al.* utilized few features than other previous research that used many feature sets, many techniques have been used to reduce dimensions [25]. The KNN and ANN classifiers are used to classify these few features. 50 cases of the 'BAHEYA Foundation to Early Detection and Treatment of Breast Cancer by doctors and radiologists in the hospital have been utilized for the proposed system. The images used are contrast-enhanced spectral mammograms (CESMs) that have clearer and more contrasting images compared to the typical mammals. The KNN and ANN classifiers were used and the outcomes indicate to achieve accuracy percent with 92 percent with ANN.

In 2019 Salman *et al.* have proposed a system for detect potential cancer tumors in mammograms, the detection is made through automatically dividing breast images by combining hybrid density slicing technique with the adaptive k-means algorithm, also by dividing breast images and extracting areas of cancer [26]. GLCM have been used with proposed features that are gray level density matrices (GLDM) to detect abnormal tissue using MLP classifiers. Experimental results showed a significant improvement in breast cancer diagnosis accuracy with more than, 91.17%. In 2019 Vijayarajeswari *et al*, present a CAD system with the features obtained with the use of the Hough transformation method, it is a 2-D transformation [27]. Which is utilized for isolating the feature of a specific shape in the image. This study discusses strategies for the process of classification and feature extraction. Here, it is utilized for the detection of the mammogram image features and has been classified with the use of the SVMs. The results have shown that the suggested approach has been successful in classifying the abnormal mammogram classes with an accuracy of 94%.

2. RESEARCH METHOD

The main goal of the present search is to build a classifier model to helps physicians and diagnostic experts by providing a second diagnostic view for a more reliable diagnostic decision. Figure 4 illustrates the diagram of the stages of the proposed classifier model.
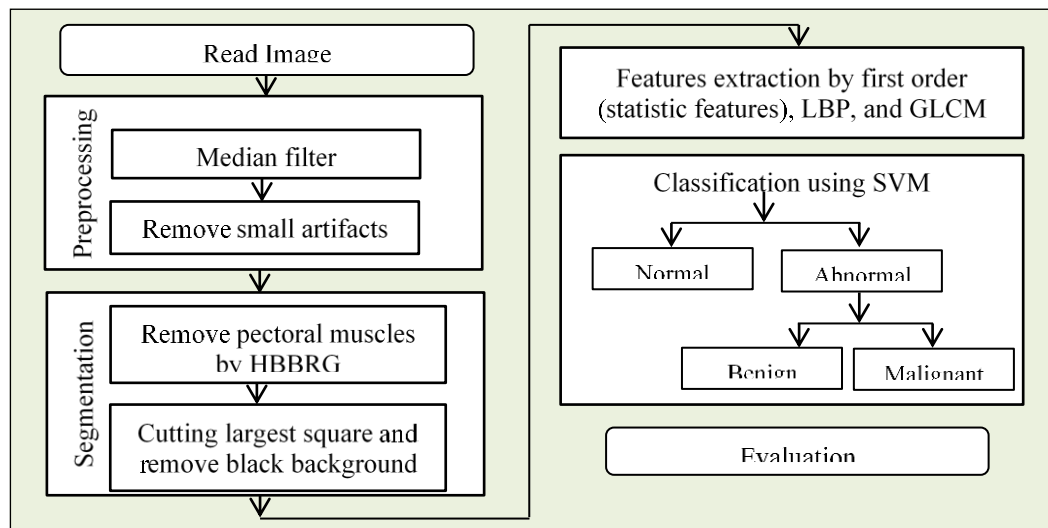
Figure 4. Show diagram of the suggested system

2.1. Preprocessing

Mammography sometimes contains many errors like noise, small artifacts, and pectoral muscles. These effects should be removed because they greatly affect the results of the following stages, such as feature extraction. A median filter is used for noise and local contrast removal; it represents a filter of the spatial domain, which works through replacing the central pixels in a certain block with block pixel median values. The small artifacts have been eliminated through the conversion of the image into a binary format with the use of a suitable threshold and after that, the arrangement of those components through the area for the isolation of small spaces, which include the numbers and labels. The results which have been obtained from the application of this stage have been illustrative in Figure 5.
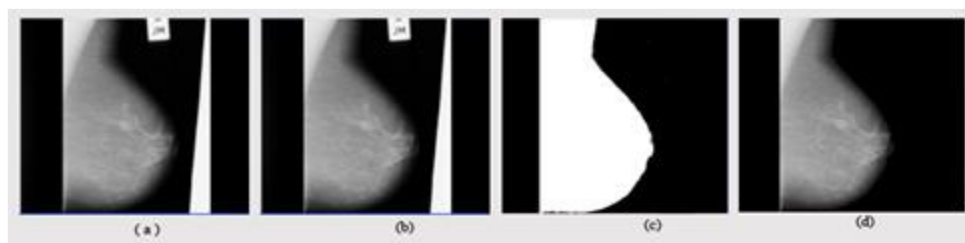

Figure 5. Show results the pre-processing: (a) Input image; (b) After apply median filter; (c) Cut breast only; (d) Image without noise and label

2.2. Segmentation

Many segmentation algorithms have been used on medical images. In this paper, this stage was applied to remove pectoral muscles and cut the largest possible square from a mammogram, which represents the ROI. Firstly, the HBBRG algorithm was used for the pectoral muscle removal by combining BB and RG. Where BB algorithms were applied according to the fact that pectoral muscles are almost triangular and are appearing in the breast contour's upper left or corner according to whether it is the right or the left breast, then the region's growing algorithm applied by selecting a seed point that will be automatically characterized to be in the pectoral muscle limits. In addition to that, this function requires locating the distance of the maximal intensity between the seed point and neighbor pixels, finally, the 2 methods have been combined to obtain the HBBRG mask as shown in Figure 6.
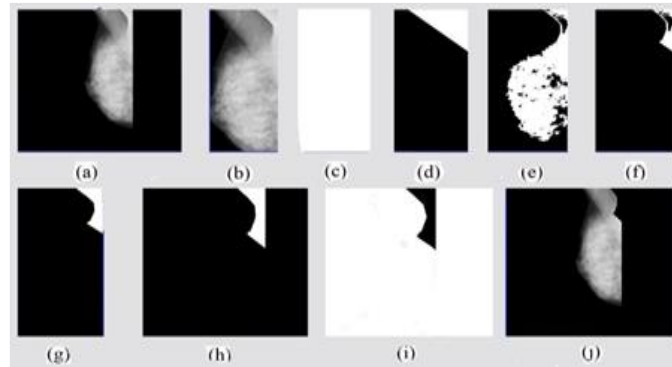
Figure 6. Results of the HBBRG algorithm: (a) Preprocessing image; (b) Cut breast only; (c) One's value for breast; (d) Mask BB €. mask RG; (f) Merging between BB, and RG mask (HBBRG mask); (g) Improvement mask HBBRG algorithm; (h) Integrates with the zero matrix's; (i) Inverse mask; (j) Output image

In the second phase, the breast image has been segmented to obtain the largest possible square area that can be cut from the image, as this area is square, This process was applied using a geometrical method by converting the image into the binary and finding the mask that contains only the breast with one's value and then perform a reverse search process starts from the penultimate pixel in the lower right corner and compares it with its three neighbors right, bottom and diagonal to find the least value between them and then increase its value in one measure, this process continues on all the mask, then a square is drawn with the coordinates starting from the smallest pixel to the largest value. Finally, to remove the black background all columns and rows with a total sum equal to zero are excluded. Algorithm1 describes the process and Figure 7 illustrates results.
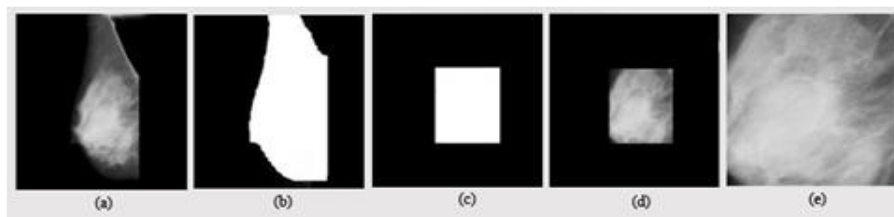

Figure 7. Show results segment mammogram image: (a) Mammogram image; (b) Binary mask; (c) Largest square mask; (d) Output largest square; (e) Remove black background (ROI)

Algorithm (1): Cut the largest possible square to find the ROI
Input: Mammogram image without pectoral muscles
Output: Image with the largest possible square
Begin
Step1: IM = read image.
Step2: Convert input image to binary
Step3: assign zero arrays (Mask1) with the size of the mammogram image
Step4: find a mask with one's value for breast only (mask2) from segmentation of the mammogram image
Step5: find pixels $p(r, c)$ before the last for (Mask2), which is in the lower right corner, then apply
 The following operations: $a = p(r, c+1)$; $b = p(r+1, c)$; $d = p(r+1, c+1)$
 Temp = min ([a b d]); $p(r, c)$ = Temp + 1.
Step6: Find the pixel that contains the maximum value in the formed array: max ($p(r, c)$). Then apply
 A following equation to find a square with white values, Mask3 (c: c +p(c, r), r: r+ p(c, r)) =1
Step7: for i= r to r +p (c, r)
Step8: for j= c to c + p(c, r)
Step9: IM= Multiply Mask3 (i, j) × IM (i, j)
Step10: end for
Step11: end for
Step12: counting the sum of pixels values for all rows and column and then find the min and max
For rows and columns containing a sum greater than zero
Step13: ROI = IM (min row: max row, min column: max column)
End

where r is the row, c is the columns, and Temp is the temporary tank.

2.3. Features extraction

In this phase, the ROI is assigned a set of features that represent the properties of the tissue. These features can be a set of real numbers through which the normal tissue can be distinguished from the abnormal and malignant from benign tissue. In this paper, after finding the ROI for each image the vector features consist of 73 features that were calculated based on three techniques. Firstly, ten features of the first-order features are (mean and SD of the mean, mean and SD of SD, mean and SD of variance, mean and SD of skewness, mean and standard deviation of kurtosis), Secondly the fifty-nine feature of the LBP method which is listed from eleven to sixty-nine features. Finally, four features of the GLCM are contrast, correlation, energy, and homogeneity. All of those features were calculated with the aim of creating powerful texture features. The steps for extracting these features are described in Algorithm 2.

Algorithm (2): Feature Extraction for the ROI

Input: ROI for images
Output: Array of features
Begin
Step1: Read ROI image I and Get row and column of the image I [r c].
Step2: For j =1to c
Step3: Find the mean, Standard deviation, variance, skewness, and kurtosis.
Step4: end for
Step5: features = [M (mean), SD (mean), M (standard deviation), SD (standard Deviation, M (variance),
    SD(variance),M(skewness),SD(skewness), M(kurtosis),and SD(kurtosis)].
Step6: Find features LBP using it extractLBPFeatures (I)
Step7: Create a GLCM from Image I with (0°) and displacement d=1
Step8: Find features GLCM using its features = graycoprops (GLCM).
Step9: Find the Contrast of ROIs.
Step10: Find the Correlation of ROIs
Step11: Find the Energy of ROIs.
Step12: Find the Homogeneity of ROIs.
Step13: Save features
End

2.4. Classification

Once extracted the features, choose the appropriate ones to enter it into the SVM classifier. SVM has been applied in two levels of binary classification, the first level representing the classification of image features to a normal or abnormal image, then if the results of the first level are abnormal, the second level of binary classification is applied, which classifies features of benign or malignant images. SVM can be defined as a supervised ML classifier, where a reduced feature vector from the step of the features selection has been provided as input data to SVMs classifier. It produces support vectors for the identification of boundaries between both classes. This support-vector is utilized for the determination of the hyperplane position where it has been tested with a variety of kernel functions. There is an infinite number of separating lines which may be drawn, the objective is finding the "optimal" one, which means, one which has a minimal classification error on the previously unseen tuples. The SVM has approached this issue by searching for maximal marginal hyper-plane. The optimal splitting of a hyperplane is shown in Figure 8.
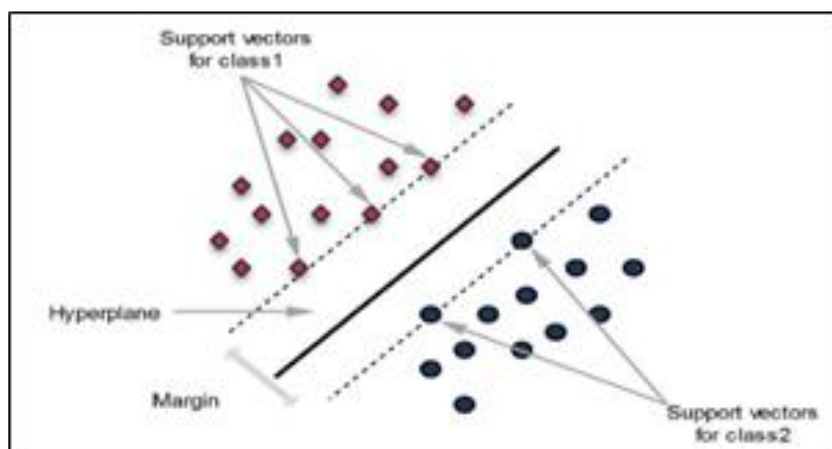

Figure 8. Optimal hyper plane for support vector machine

3.   RESULTS AND DISCUSSIONS

The efficiency of the approaches of machine learning is evaluated based on some of the indices of the performance measures. The result, false positive (FP), may put the patient in a fragile position however, using complementary exams, the result may be excluded. While in a case where the results the false negative (FN) it is a more worrying case if an individual has the lesion but the algorithm does not detect [28]. A confusion matrix for the predicted and actual classes is carried out comprising FP, true positive (TP), FN, and false negative (TN). In this classification, positive/negative indicates the decision which has been made by the algorithm, and true/false indicates the way by which the decision agrees with the actual clinical state. Where in case the two classes there are only four possible outputs represented elements of the confusion matrix of (2*2) for a binary classifier see Figure 9 [29], [30].
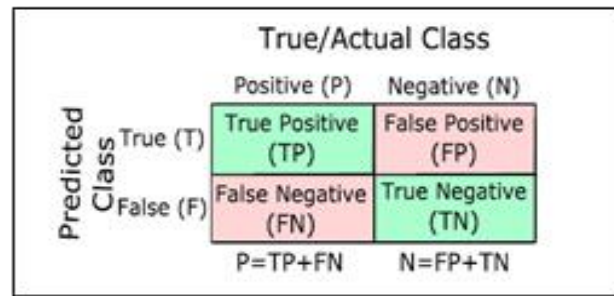

Figure 9. An illustrative example of the 2*2 confusion matrix [29]

There are six statistical metrics utilized for the evaluation of the efficiency of the proposed system based on the confusion matrix are accuracy (ACC), error rate (ERR), sensitivity (SN), false-positive rate (FPR), specificity (SP), and precision (P). In the present research, the proposed system has been applied to all images in the MIAS database, where 70% of the image was used for the training phase and 30% testing phase of random instants of image features from the dataset with 100 iterations. The results show that SVM for the first level has achieved the average, best, and worse accuracy they are 89.171%, 95.454%, and 79.293%, respectively, see Table 1. Also, the results show that SVM for the second level has achieved the average, best, and worse accuracy they are 90.493%, 97.60%, and 80.342%, respectively, as shown in Figure 10.

Table 1. Confusion matrix result of SVM for first-level

#	ACC	ERR	SE	FPR	SP	P
The average results	89.171	10.828	99.45	33.359	66.64	83.519
The best results	95.454	4.545	100	11.543	88.456	93.377
The worst results	79.281	20.717	86.301	47.058	52.941	73.856

Analysis of the performance of results of the evaluation metrics for a first-level classification which classifies the image to normal or abnormal classes is shown in Figure 10, where it can be observed the best and worst value obtained for this classifier.
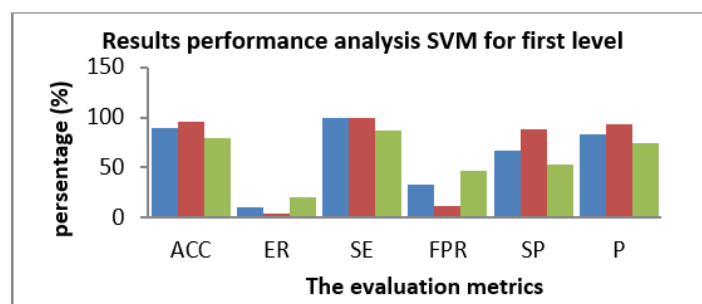

Figure 10. Displays the performance analysis of the SVM for first level from Table 1

Table 2 shows Confusion matrix result of SVM for second-level. Analysis of the performance of results of the evaluation metrics for a second-level classification which classifies the abnormal images only to benign and malignant classes is shown in Figure 11, where it can be observed the best and worst value obtained for this classifier.

Table 2. Confusion matrix result of SVM for second-level

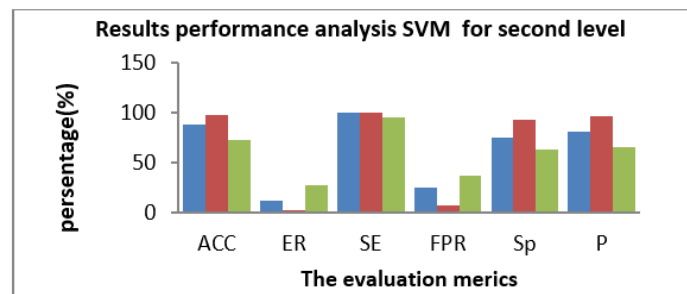| #	ACC	ERR	SE	FPR	SP	P
The average results	89.171	10.828	99.45	33.359	66.64	83.519
The best results	95.454	4.545	100	11.543	88.456	93.377
The worst results	79.281	20.717	86.301	47.058	52.941	73.856

Figure 11. Displays the results and performance analysis of the SVM for second level

In this part, in order to better evaluate our proposed system, our proposed system was compared with a set of previous works by comparing all the main stages of the system, algorithms, techniques, and the level for which the system was designed without acres and the results obtained. These comparisons are shown in Table 3.

Table 3. Shows the comparison of the proposed system with the related works

Related work	Pre-Processing and Enhancement	Segmentation	Feature extraction	Classifier	Accuracy (%)
(Sheba, and Raj,2018) [22]	median filter global thresholding	BB, global thresholding, and morphological	Shape, GLRLM, and GLCM.	Feedforward neural networks (FFNN)	96 for the first and second levels
(Salman, and Semaa,2019) [24]	median filter and histogram equalization	Segmentation based dataset, K-mean, and RG	GLCM, and GLDM Feature	MLP	91.17 for the second level
Vijayarajeswari etl,2019) [25]	Remove all unwanted part	maximization estimation	Hough transform	SVM	94% for second level
Proposed method	median filter global thresholding	HBBRG algorithm and cutting largest possible square	First-order, LBP, and GLCM.	SVM	95.454 For first level, 97.260 for second level

4. CONCLUSION

Mammography is the most effective method that is used in early detections of breast cancer, The main objective is developing the CAD system for diagnosis mammogram images to assist doctors and diagnostic experts by providing a second viewpoint, it gives more confidence to the diagnostic process. In this paper, The results proved that the median filter is an ideal filter to remove noise and local contrast found in the mammogram. The proposed algorithm for removing small artifacts has achieved 100% results for this purpose. Removal of the pectoral muscles represents the biggest obstacle in the treatment of mammograms because they closely resemble tumors and the rest of the breast tissue, particularly in the types of fatty and adenocarcinomas, as well as their presence, has a great impact on the results of the following stages. A geometric segmentation method was proposed to cut the largest possible square representing a sample of breast tissue, it achieving 100% success with normal images and more than 98% with abnormal images, where the tumor is within the crop area. The proposed texture features are based on three technologies are first-order, LBP, and GLCM, these features give the algorithm more robustness due to its resistance to many image situation variations that lead

to the best discrimination potential for classification the type of image. Finally, several future research work can be done for the production of our paper such as model development from the diagnostic model to the diagnostic and prediction models, and tests of new segmentation methods that will provide better results to identify the damage and insulation from the rest of our breast tissue, particularly in fatty and glandular photos, during this stage, also apply the proposed breast diagnosis model to other breast imaging and examination methods, like MRI and CT.

REFERENCES

[1]   F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: a cancer journal for clinicians*, vol. 68, no. 6, pp. 394-424, 2018, doi: 10.3322/caac.21492.

[2]   Y. Qiu, *et al*., "An initial investigation on developing a new method to predict short-term breast cancer risk based on deep learning technology," *Proceedings, Medical Imaging 2016: Computer-Aided Diagnosis*, 2016, vol. 978, doi: 10.1117/12.2216275.

[3]   A. Jalalian, S. B. T. Mashohor, H. R. Mahmud, M. I. B. Saripan, A. R. B. Ramli, and B. Karasfi, "Computer-aided detection/diagnosis of breast cancer in mammography and ultrasound: a review," *Clinical imaging*, vol. 37, no. 3, pp. 420-426, 2013, doi: 10.1016/j.clinimag.2012.09.024.

[4]   I. K. Maitra, S. Nag, and S. K. Bandyopadhyay, "Technique for preprocessing of a digital mammogram," *Computer methods and programs in biomedicine*, vol. 107, no. 2, pp. 175-188, 2012, doi: 10.1016/j.cmpb.2011.05.007.

[5]   A. Makandar and B. Halalli, "Pre-processing of mammography image for early detection of breast cancer," *International Journal of Computer Applications*, vol. 144, no. 3, pp. 11-15, 2016.

[6]   M. A. Berbar, "Hybrid methods for feature extraction for breast masses classification," *Egyptian informatics journal*, vol. 19, no. 1, pp. 63-73, 2018, doi: 10.1016/j.eij.2017.08.001.

[7]   M. Tembey, "Computer-aided diagnosis for mammographic microcalcification clusters," Ph.D. dissertation, Dept. Comp. Sci. and Eng., Univ. of South Florida, Scholar Commons, 2003.

[8]   H. D. Cheng, J. Shan, W. Ju, Y. Guo, and L. Zhang, "Automated breast cancer detection and classification using ultrasound images: A survey," *Pattern recognition*, vol. 43, no. 1, pp. 299-317, 2010, doi: 10.1016/j.patcog.2009.05.012.

[9]   E. M. H. Saeed and H. A. Saleh, "Pectoral Muscles Removal in Mammogram Image by Hybrid Bounding Box and Region Growing Algorithm," *2020 International Conference on Computer Science and Software Engineering (CSASE)*, Duhok, Iraq, 2020, pp. 146-151, doi: 10.1109/CSASE48920.2020.9142055.

[10]   P. S. Duth, V. Viswanath, and P. Sreekumar, "Brain image segmentation using level set: An hybrid approach," *International Journal of Advances in Applied Sciences (IJAAS)*, vol. 6, no. 3, pp. 267-276, 2017, doi: 10.11591/ijaas.v6.i3.pp258-267.

[11]   S. Li, T. Hara, Y. Hatanaka, H. Fujita, T. Endo, and T. Iwase, "Performance evaluation of a CAD system for detecting masses on mammograms by using the MIAS database," *Medical Imaging and Information Sciences*, vol. 18, no. 3, pp. 144-153, 2001, doi: 10.11318/mii1984.18.144.

[12]   J. Suckling *et al.*, "The mini-MIAS database of mammograms," [Online] Available: http://peipa.essex.ac.uk/info/mias.html.

[13]   A. Hussain, A. N. Mazher, and A. Razak, "Classification of Breast Tissue for mammograms images using intensity histogram and statistical methods," *Iraqi Journal of Science*, vol. 53, no. 4, pp. 1092-1096, 2012.

[14]   K. N. N. Hlaing, "First order statistics and GLCM based feature extraction for recognition of Myanmar paper currency," *Proceedings of the 31st IIER International Conference*, Bangkok, Thailand. 2015, pp. 1-6.

[15]   S. K. M Hamouda, R. H. A. El-Ezz, and M. E. Wahed, "Enhancement accuracy of breast tumor diagnosis in digital mammograms," *Journal of Biomedical Sciences*, vol. 6. no. 4, pp. 1-8, 2017, doi: 10.4172/2254-609X.100072.

[16]   S. Maniar and J. S. Shah, "Classification of content based medical image retrieval using texture and shape feature with neural network," *International Journal of Advances in Applied Sciences (IJAAS)*, vol. 6, no. 4, pp. 368-374, 2017, doi: 10.11591/ijaas.v6.i4.pp368-374.

[17]   P. P. Dash, D. Patra, and S. K. Mishra, "Local binary pattern as a texture feature descriptor in object tracking algorithm," *Intelligent Computing, Networking, and Informatics*, vol. 243, pp. 541-548, 2014, doi: 10.1007/978-81-322-1665-0_52.

[18]   L. Nanni, A. Lumini, and S. Brahnam, "Local binary patterns variants as texture descriptors for medical image analysis," *Artificial intelligence in medicine*, vol. 49, no. 2, pp. 117-125, 2010, doi: 10.1016/j.artmed.2010.02.006.

[19]   R. Nosaka, Y. Ohkawa, and K. Fukui, "Feature extraction based on co-occurrence of adjacent local binary patterns," *Proceedings, Part II: Advances in Image and Video Technology - 5th Pacific Rim Symposium*, PSIVT 2011, Gwangju, South Korea, November 20-23, 2011, pp 82-91, doi: 10.1007/978-3-642-25346-1_8.

[20]   F. R. De Siqueira, W. R. Schwartz, and H. Pedrini, "Multi-scale gray level co-occurrence matrices for texture description," *Neurocomputing*, vol. 120, no. 1, pp. 336-345, 2013, doi: 10.1016/j.neucom.2012.09.042.

[21]   M. A. Al Mutaz, S. Dress, and N. Zaki, "Detection of masses in digital mammogram using second order statistics and artificial neural network," *International Journal of Computer Science & Information Technology (IJCSIT)*, vol. 3, no. 3, pp. 176-186, June 2011.

[22] F. H. Mahmood and W. A. Abbas, "Texture Features Analysis using Gray Level Co-occurrence Matrix for Abnormality Detection in Chest CT Images," *Iraqi Journal of Science*, vol. 57, no. 1A, pp. 279-288, 2016.

[23] J. Harefa, Alexander, and M. Pratiwi, "Comparison classifier: support vector machine (SVM) and K-nearest neighbor (K-NN) in digital mammogram images," *Jurnal Informatika dan Sistem Informasi*, vol. 2, no. 2, pp. 35-40, 2016.

[24] K. U. Sheba and S. G. Raj, "An approach for automatic lesion detection in mammograms," *Cogent Engineering*, vol. 5, no. 1, pp. 1-16, 2018, doi: 10.1080/23311916.2018.1444320.

[25] S. Mostafa, R. Mubarak, M. El-Adawy, A. F. Ibrahim, M. M. Gomaa and R. M. Kamal, "Breast Cancer Detection Using Polynomial Fitting Applied on Contrast Enhanced Spectral Mammography," *2019 International Conference on Innovative Trends in Computer Engineering (ITCE)*, Aswan, Egypt, 2019, pp. 11-16, doi: 10.1109/ITCE.2019.8646379.

[26] N. H. Salman and S. I. M. Ali, "Breast Cancer Classification as Malignant or Benign based on Texture Features using Multilayer Perceptron," *International Journal of Simulation--Systems, Science & Technology*, vol. 20, no. 1, pp. 1-11, 2019, doi: 10.5013/IJSSST.a.20.01.12.

[27] R. Vijayarajeswari, P. Parthasarathy, S. Vivekanandan, and A. A. Basha, "Classification of mammogram for early detection of breast cancer using SVM classifier and Hough transform," *Measurement*, vol. 146, no. 1, pp. 800-805., 2019, doi: 10.1016/j.measurement.2019.05.083.

[28] R. F. d. S. Teixeira, "Computer analysis of mammography images to aid diagnosis," MSc in Biomedical Engineering, Faculdade de Engenharia da Universidade do Porto 35, 2012.

[29] A. Tharwat, "Classification assessment methods," *Applied Computing and Informatics*, vol. 17. no. 1, pp. 168-192, 2020, doi: 10.1016/j.aci.2018.08.003.

[30] A. Sahu and S. Pattnaik, "Feature Selection Using Evolutionary Functional Link Neural Network for Classification," *International Journal of Advances in Applied Sciences (IJAAS)*, vol. 6, no. 4, pp. 359-367, 2017, doi: 10.11591/ijaas.v6.i4.pp359-367

BIOGRAPHIES OF AUTHORS

Enas Mohammed Hussein Saeed received the B.Sc. in computer science from Mustansiriya Univ., Baghdad, Iraq in 1998, M.Sc. in computer science from computer science from Mustansiriyah Univ., Baghdad, Iraq in 2004, and Ph.D. in computer science from Univ. of Technology, Baghdad, Iraq in 2014.
Current position & Functions: computer science from Mustansiriyah University.
Official Email: drenasmohammed@uomustansiriyah.edu.iq
enas.saeed38@gmail.com

Hayder Adnan Saleh, a master's student, College of Education, Department of Computer Science, Al-Mustansiriya University.
an employee in the Iraqi Ministry of Education from 2009
The e-mail: haider8883gmail.com

Enam Aziz Khalel
MBCHB-DMRD
Specialist radiology
Head of radiology department -oncology teaching hospital- medical city complex
Enam_azez@yahoo.com