

Computer-aided Diagnosis for the Early Breast Cancer Detection

Miran Hakim Aziz¹, Alan Anwer Abdulla^{2,3}

¹Applied Computer, Collage of Medicals and Applied Sciences, Charho University, Chamchamal, Sulaimani, Kurdistan Region, Iraq, ²Department of Information Technology, College of Commerce, University of Sulaimani, Sulaimani, Iraq,

³Department of Information Technology, University College of Goizha, Sulaimani, Iraq



ABSTRACT

The development of the use of medical image processing in the healthcare sector has contributed to enhancing the quality/accuracy of disease diagnosis or early detection because diagnosing a disease or cancer and identifying treatments manually is costly, time-consuming, and requires professional staff. Computer-aided diagnosis (CAD) system is a prominent tool for the detection of different forms of diseases, especially cancers, based on medical imaging. Digital image processing is a critical in the processing and analysis of medical images for the disease diagnosis and detection. This study introduces a CAD system for detecting breast cancer. Once the breast region is segmented from the mammograms image, certain texture and statistical features are extracted. Gray level run length matrix feature extraction technique is implemented to extracted texture features. On the other hand, statistical features such as skewness, mean, entropy, and standard deviation are extracted. Consequently, on the basis of the extracted features, support vector machine and K-nearest neighbor classifier techniques are utilized to classify the segmented region as normal or abnormal. The performance of the proposed approach has been investigated through extensive experiments conducted on the well-known Mammographic Image Analysis Society dataset of mammography images. The experimental findings show that the suggested approach outperforms other existing approaches, with an accuracy rate of 99.7%.

Index Terms: Computer-aided Diagnosis, Medical Image, Breast Cancer, Gray Level Run Length Matrix, Classifier Technique

1. INTRODUCTION

Digital image processing (DIP) is significant in many areas, particularly medical image processing, image in-painting, pattern recognition, biometrics, content-based image retrieval, image de-hazing, and multimedia security [1], [2]. It is becoming more important for analyzing medical images and identifying abnormalities in these images. Computer-aided diagnosis (CAD) systems based on image processing have emerged as an intriguing topic in the field of medical image processing research. A CAD system is a computer-based system that assists medical professionals in diagnosing

diseases, in particular cancers, using medical images such as X-ray, magnetic resonance imaging (MRI), computed tomography (CT), ultrasound, and microscopic images [3]. The aim of developing autonomous CAD systems is to extract the targeted illnesses with a high accuracy and at a lower cost and time consumption. Preprocessing, segmentation, feature extraction, and classification are the four basic phases of each CAD system. A feature is an important factor to categorize the disease in the cancer detection systems. Feature extraction is the process of transforming raw data into a set of features [4]. There are numerous types of cancers such as breast cancer, brain tumors, lung cancer, skin cancer, and blood cancer. This paper focuses on the early detection of the cancerous cells in the breast. Breast cancer is one of the most frequent kinds of cancer among females worldwide. There are currently no strategies for preventing breast cancer. The difficulty of radiologist interpretation of mammogram images can

Access this article online

DOI: 10.21928/uhdjst.v7n1y2023.pp7-14

E-ISSN: 2521-4217

P-ISSN: 2521-4209

Copyright © 2022 Aziz and Abdulla. This is an open access article distributed under the Creative Commons Attribution Non-Commercial No Derivatives License 4.0 (CC BY-NC-ND 4.0)

Corresponding author's e-mail: Dr. Alan Anwer Abdulla, Assistant Prof., Department of Information Technology, College of Commerce, University of Sulaimani, Sulaimani, Iraq, Department of Information Technology, University College of Goizha, Sulaimani, Iraq. E-mail: Alan.abdulla@univsul.edu.iq

Received: 17-09-2022

Accepted: 11-12-2022

Published: 12-01-2023

be alleviated by employing the early-stage breast cancer detection method. Thus, early diagnosis of this condition is critical in its treatment and has a significant influence in minimizing mortality. The most effective way of detecting breast cancer in its early stages is to analyze mammography images [5]. Breast cancer is a disorder in which the cells of the breast proliferate uncontrollably. The kind of breast cancer is determined by which cells in the breast develop into cancer. Breast cancer can start in any part of the breast. It can spread outside of the breast through blood and lymph arteries. Breast cancer is considered to have metastasized when it spreads to other regions of the body [6]. In general, a breast is composed of three major components: lobules, ducts, and connective tissue (Fig. 1) [6].

The lobules are the milk-producing glands. Ducts are tubes that transport milk to the nipple. The majority of breast cancers start in the lobules or ducts [6]. Connective tissue joins or separates and supports all other forms of bodily tissue. It contains of cells surrounded by a fluid compartment termed the extracellular matrix (ECM), as do all other forms of tissue. However, connective tissue varies from other kinds in that its cells are loosely instead of densely packed inside the ECM [7].

The aim of this study is developing a CAD system for the early detection of breast cancer. The developed CAD system has the advantages of increasing accuracy rate, reducing time consumption, and reducing cost in comparison with manually detecting system. The main contributions of the proposed approach are segmenting the breast region properly as well as extracting the most significant features, and this leads to increase the accuracy rate and reduce mistake rate of wrongly treating patients. The proposed system includes the following

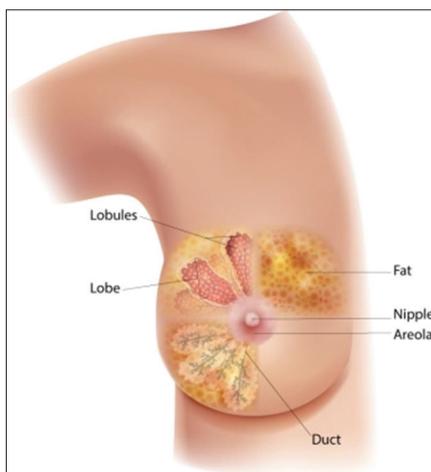


Fig. 1. Major components of the breast [6].

steps: A pre-processing step for enhancing the image quality, a segmentation step for segmenting the breast region from the other components of mammography images, and a feature extraction step for extracting the most influential features. Finally, the classification step is conducted, which helps the system decide whether a cell is cancerous or non-cancerous. The rest of the paper is structured as follows. Section 2 provides a summary of past efforts from the literature. Section 3 presents the proposed CAD system. Section 4 shows the results of experiments. Finally, Section 5 gives the conclusion.

2. LITERATURE REVIEW

In medical image processing, the CAD system is a computer-based system that helps clinicians in their last decision about different diseases, especially cancers. The whole process is about extracting significant information from medical images such as: MRI, CT, and ultrasounds. Several CAD systems have been developed for identifying different diseases including: Breast cancer, tumor detection, and lung cancer. This study concentrates on breast cancer.

The processing and analysis of breast mammogram images plays a significant role in the early diagnosis of breast cancer. This section reviews the most influential as well as relevant current efforts on the early breast cancer detection using DIP. The main obstacle in this field of research is reducing the rate of breast cancer detection errors. In general, most of the CAD systems for the early breast cancer detection consist of the following steps: Image enhancement, image segmentation, feature extraction, feature selection, and classification.

In 2010, Eltoukhy *et al.* suggested an algorithm for the breast cancer detection using a curvelet transform technique at multiple scales [8]. Different scales of the largest curvelet coefficients are extracted and investigated from each level as a classification feature vector. This algorithm is reached an accuracy rate of 98.59% at Scale 2. Srivastava *et al.*, in 2013, introduced a CAD system for the early breast cancer diagnosis using digital mammographic images [9]. Contrast-limited histogram equalization technique is utilized for the enhancement purposes. Consequently, three-class fuzzy C-means is used for the segmentation process. The texture features such as geometric/shape, wavelet-based, and Gabor were extracted. The minimum redundancy maximum relevance feature selection method was utilized to select the fewest redundant and most relevant characteristics. Finally, Support Vector

Machine (SVM), K-Nearest Neighbor (kNN), and Artificial Neural Network (ANN) classifier techniques were used for classifying cancerous and non-cancerous cells. Furthermore, SVM provides better results in comparison to the kNN and ANN. This technique is achieved an accuracy rate of 85.57% for the 10-fold cross-validation using Mammographic Image Analysis Society (MIAS) dataset of images.

Vishrutha *et al.*, in 2015, developed a strategy for combining wavelet and texture information that leads to increase the accuracy rate of the developed CAD system for the early breast cancer diagnosis [10]. The mammogram images were pre-processed using median filter. In addition, the label and the black background are removed on the bases of sum of each column's intensities. Consequently, if the total intensity of a column falls below a certain level/threshold, the column will be removed. The resulted images from the pre-processing step were utilized as input for the region growth technique used to determine the region of interest (ROI) as a segmentation step. Discrete Wavelet Transform technique was used to extract features from the segmented images/regions. Finally, SVM classifier technique was utilized to categorize the mammogram images as benign or malignant with an accuracy rate of 92% using Mini-MIAS dataset of images.

In 2017, Pashoutan *et al.* developed a CAD system for the early breast cancer diagnosis [11]. For the pre-processing step, cropping begins by employing coordinates and an estimated radius of any artifacts introduced into images to get to the ROI where bulk and aberrant tissues are found. Moreover, histogram equalization and median filter were used to enhance the contrast of the images. Edge-based segmentation and region-based segmentation methods are that the two main methods were used for the segmentation purposes. Furthermore, four different techniques were utilized for extracting features, such as Wavelet transform, Gabor wavlet transform, Zernike moments, and Gray-Level Cooccurrence Matrix (GLCM). Eventually, using the MIAS dataset, this technique reached an accuracy rate of 94.18%.

Hariraj *et al.*, in 2018, developed a CAD system for the breast cancer detection [12]. In the pre-processing step, Fuzzy Multi-layer was used to eliminate background information such as labels and wedges from images. Moreover, thresholding was used to transform the grayscale image to the binary image. Furthermore, morphological technique was implemented on the binary image to remove undesirable tiny items. Regarding to the segmentation step, K-means clustering was utilized. For the feature extraction purposes, certain

shape and texture features were extracted such as: diameter, perimeter, compactness, mean, standard deviation, entropy, and correlation. Finally, the Fuzzy Multi-Layer SVM classifier technique provides better accuracy rate of 98% out of other tested classifier techniques using Mini-Mammographic MIAS dataset of images.

Sarosa *et al.*, in 2019, designed a breast cancer diagnosis technique by investigating GLCM and Backpropagation Neural Network (BPNN) classification technique [13]. Histogram equalization was utilized for the pre-processing and enhancing the images. Consequently, GLCM was used to extract features from the pre-processed images. Finally BPNN was used to determine whether the input image is normal or abnormal. The suggested approach was evaluated using a MIAS dataset of images and it achieved an accuracy rate of 90%.

In 2019, Arafa *et al.* introduced a technique for the breast cancer detection [14]. In the pre-processing step, just the area including the breast region is automatically picked and artifacts as well as pectoral muscle were removed. The Gaussian Mixture Model (GMM) was utilized to extract the ROI. Moreover, texture, shape, and statistical features were extracted from the ROI. For the texture feature, GLCM was utilized. Furthermore, the following shape features such as circularity, brightness, compactness, and volume were extracted. Regarding to the statistical features, mean, standard deviation, correlation, skewness, smoothness, kurtosis, energy, and histogram were extracted. Finally SVM classifier technique was used to classify segmented ROI into normal, abnormal, benign, and malignant. This proposed technique was evaluated using MIAS dataset of images and it achieves an accuracy of 92.5%.

Farhan and Kamil developed a CAD system for classifying the input mamogram images into normal or abnormal, in 2020, [15]. At the beginning, contrast limited adaptive histogram equalization (CLAHE) method was used to improve all mammogram images. In addition, the histogram of oriented gradient, GLCM, as well as the local binary pattern (LBP) techniques was used to extract features. Finally, SVM and kNN classifier techniques were used for classifying cancerous and non-cancerous cells. The best accuracy rate of 90.3%, using Mini-MIAS dataset, was obtained when GLCM and kNN were used.

In 2020, Eltrass and Salama developed a technique for breast cancer diagnosis [16]. As a pre-processing step, the mammography image was translated into a binary image,

and then all regions are sorted to identify the mammogram's greatest area, that is, breast region. In addition, all artifacts and pectoral muscle were eliminated. This CAD system utilized the expectation maximization technique for the segmentation purposes. Wavelet-based contourlet transform technique was used to extract features. Finally, SVM classifier technique was used and an accuracy rate of 98.16% was achieved using MIAS dataset.

Saeed *et al.*, in 2020, designed a classifier model to aid radiologists in providing a second opinion when diagnosing mammograms [17]. In the pre-processing step, median filter was used to remove noise and minor artifacts. Hybrid Bounding Box and Region Growing algorithm was used to segment the ROI. For the features extraction, two types of features were extracted which are: (1) Statistical features such as mean, standard deviation skewness, and kurtosis and (2) texture features such as LBP and GLCM. Consequently, SVM was used to categorize mammography images as normal or abnormal in the first level, and benign or malignant in the second level. This proposed technique used MAIS dataset to evaluate the performance, and an accuracy of 95.45% was obtained for the first level and 97.26% for the second level.

Mu'jizah and Novitasari in 2021, developed a CAD system for the breast cancer diagnosis [18]. At the beginning, certain pre-processing techniques, such as Gaussian filter and Canny edge detection technique, were implemented to enhance the visual quality of the input images. The thresholding method was also used for the segmentation purposes. To extract features, GLCM was used as texture feature, and area, perimeter, metric, as well as eccentricity were extracted as shape feature. Finally, for the classification step, SVM was used and an accuracy rate of 98.44% was obtained using Mini-MIAS dataset of images.

Recently, in 2022, Holi produced a breast cancer detection system [19] which used a median filter and CLAHE for enhancing the input image. Then, Chebyshev Distanced-Fuzzy C-Means Clustering was used to segment the pre-processed image. The augmented local vector pattern, shape features, and GLCM were used to extract features. The classification step was conducted using kNN classifier technique. This proposed technique was achieved an accuracy rate of 97% using MIAS dataset of images.

The remainder of this paper concerns with the extension and further refinement of the strategy of using DIP to increase the accuracy rate for the early breast cancer detection.

3. PROPOSED APPROACH

The microscopic image of breast is called a mammogram, which consists of three parts/regions. The breast part appears on a mammogram in colors of gray and white, while the mammogram backdrop is often black. In addition, a lump or tumor appears as a concentrated white area. Tumors may be either malignant or benign [20]. The most significant step of each CAD system for the breast cancer detection is extracting/cropping the ROI from the other parts of the mammogram image. This section describes the proposed approach which involves the following steps:

1. Pre-processing: In this step, certain techniques are applied such as region-props to delete the label from the mammogram images, and median filter as well as adaptive histogram equalization to enhance the image quality (Fig. 2).
2. Segmentation: To segment the ROI from other parts of the input image, the thresholding segmentation technique is applied on image (d) in Fig. 2, and the resulted image is a binary image, see image (a) in Fig. 3.

The threshold-based segmentation approach is an effective segmentation technique that divides an image based on the intensity value of each pixel. It is used to segment an image into smaller portions using a single color value to generate a binary image, with black representing the background and

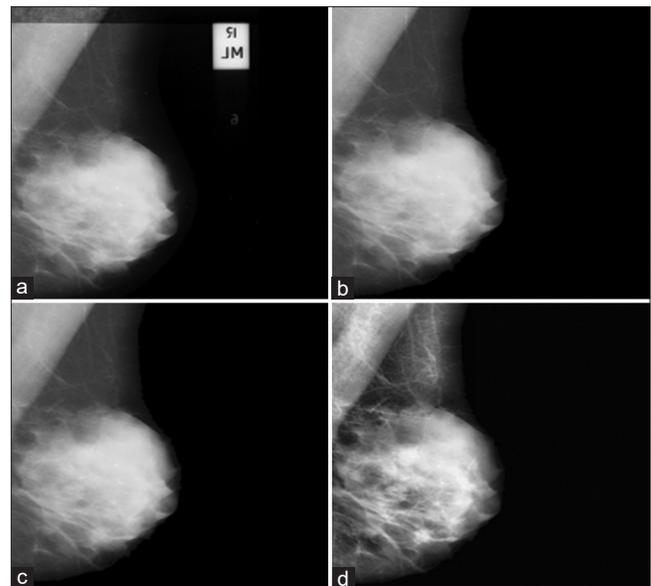


Fig. 2. Pre-Processing Step:(a) Original Mammogram Image, (b) Label Removed, (c) Resulted Image After the Median Filter has been applied on Image (b), and (d) Resulted Image after Histogram Equalization has been applied on Image (c).

white representing the objects [21]. The threshold T value can be selected either manually or automatically based on the characteristics of the image. In the proposed approach, $T = 0.7$ was used, which provides the optimum accuracy results. In the next section, all the tested values for the T are illustrated in Table 5.

3. Feature Extraction: Texture features and statistical features are extracted from the segmented image, that is, image (b) in Fig. 3. The extracted features are summarized in Table 5. Furthermore, all the extracted features are fused for the classification purposes.
4. Classification: SVM and kNN classification techniques were applied on the extracted features to distinguish normal cells from abnormal cells. The reason behind using SVM and kNN is because these two classifier techniques are the most common used in this field of research. For the both classifiers, the k-fold cross-validation with $k = 5, 10, 15,$ and 20 was investigated.

Fig. 4 illustrates the block diagram of the proposed approach.

4. EXPERIMENTAL RESULTS

The primary goal of the proposed CAD approach is classifying the breast cancer cells into normal or abnormal. Experiments are carried out in a thorough manner in this part of the study to evaluate how well the suggested approach works in terms of accuracy rate. In addition, the proposed approach is assessed alongside the findings of the earlier research.

4.1. Dataset

The MIAS dataset provides the tested input images, which are taken from the public domain and are quite well recognized. The MIAS dataset contains the original 322 images, 206 normal and 116 abnormal, in the PGM format [22]. All of the images have the same resolution which is 1024 by 1024 pixels. The MIAS dataset has been taken into consideration

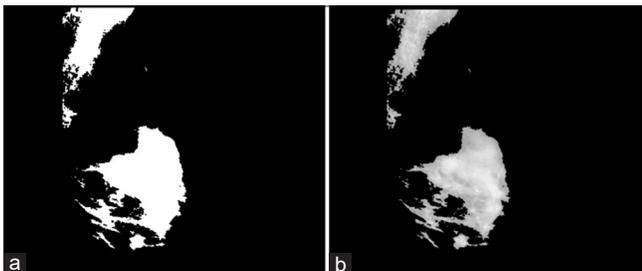


Fig 3. Segmentation step: (a) Binary image, (b) Based on the binary image in (a), the ROI is selected in the original image.

in order to assess the performance of the proposed CAD approach.

4.2. Results

Using several classifier techniques, such as SVM and kNN, the accuracy rate for each the extracted features is assessed. Tables 2 and 3 present the accuracy rate of statistical and GLCM separately using SVM and kNN respectively. In all the evaluation tests, different values of k -fold have been considered. In addition, the accuracy rate has been calculated using the following formula [23]:

$$\text{Accuracy rate} = \frac{TP + TN}{(TP + TN + FP + FN)} \quad (1)$$

Where: TP, TN, FP, and FN refer to true positive, true negative, false positive, and false negative, respectively.

More investigation has been conducted by fusing the extracted features, namely statistical and GLCM. Meanwhile, the 11 retrieved features are utilized to evaluate the effectiveness

TABLE 1: Extracted features

Type of Features	Name of the Feature
Statistical Features	Skewness
	Mean
	Entropy
	Standard deviation
Texture feature: Gray level run length matrix	Short run emphasis
	Long run emphasis
	Gray level non-uniformity
	Run percentage
	Run length non-uniformity
	Low gray level run emphasis
	High gray level run emphasis

TABLE 2: SVM-based accuracy rate for the extracted features separately

Features	5 K-fold	10 K-fold	15 K-fold	20 K-fold	Average (%)
Statistical	99.1	99.2	98.8	99.3	99.1
GLRLM	98.1	99.3	99.1	99.1	98.1

SVM: Support vector machine, GLRLM: Gray level run length matrix

TABLE 3: kNN-based accuracy rate for the extracted features separately

Features	5 K-fold	10 K-fold	15 K-fold	20 K-fold	Average (%)
Statistical	94.4	94.7	97.5	97.5	96
GLRLM	97.2	98.1	97.6	98.3	97.8

kNN: K-nearest neighbor, GLRLM: Gray level run length matrix

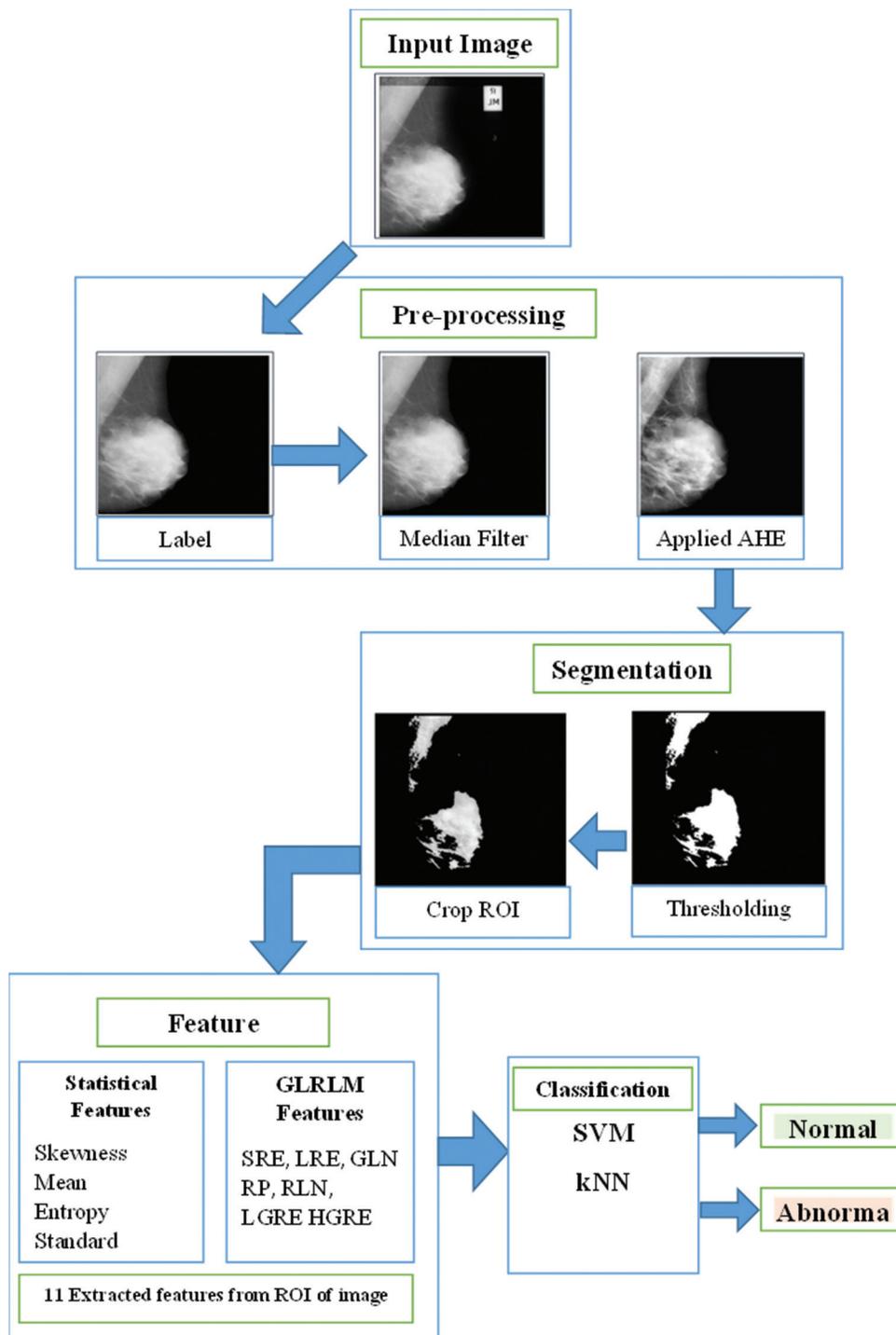


Fig. 4. Block diagram of the proposed approach.

of the proposed CAD approach in distinguishing between normal and abnormal cells. Those 11 features are previously mentioned in Table 1. Moreover, k -fold cross-validation with various values of k is used in the evaluation process

to measure the accuracy. Training and testing have been done using k -fold cross-validation, which divides data automatically into training and testing depending on the value of k . Based on the investigation conducted in this study,

the SVM classifier technique provides a higher accuracy rate (Table 4).

Tables 5 and 6 illustrate the findings of further tests done by comparing the obtained results of the proposed approach to

TABLE 4: Accuracy rate of the proposed CAD approach

	Cross validation				Average (%)
	5K	10K	15k	20k	
SVM	99.7	99.8	99.4	99.7	99.7
kNN	98.4	99.1	98.8	98.8	98.9

CAD: Computer-aided diagnosis, SVM: Support vector machine, kNN: K-nearest neighbor

TABLE 5: Accuracy rate of the tested approaches using SVM

	Accuracy Rate (%)
Proposed	99.7
Mu'jizah and Novitasari[18]	98.4
Eltrass and Salama [16]	98.1

TABLE 6: Accuracy rate of the tested approaches using K-nearest neighbor

	Accuracy Rate (%)
Proposed	98.9
Farhan and Kamil[15]	90.3
Holi [19]	97

Table 7: Time consumption of the proposed computer-aided diagnosis system

Stage	Times in second
Pre-processing	0.371
Segmentation	0.298
Feature extraction	0.061
Classification	1.274
Total	2.004 s

TABLE 8: Investigating the optimum value for thresholding T

Thresholding values	kNN (%)	SVM (%)
0.1	89	90.8
0.2	89.7	91.1
0.3	89.8	91.9
0.4	94.1	94.7
0.5	96.3	96
0.6	97.6	99.1
0.7	98.9	99.7
0.8	98.4	99.3
0.9	98.6	99.5

kNN: K-nearest neighbor, SVM: Support vector machine

results of four existing approaches. Two of the existing works were used SVM classifier techniques and the remained two works were used kNN. All of the four tested CAD systems used only 5k fold to evaluate the performance of their approaches and also Tables 7 illustrate the Time consumption of the all process in our system.

According to the results presented in Tables 5 and 6, the best accuracy rate is achieved by the proposed approach and it outperforms all the tested existing approaches. Moreover, in Eltrass and Salama [16], the total time consumption is highlighted which is (2.26267) second, while the time consuming of our proposed approach is (2.004) second. The time consumption of the proposed approach is calculated as follows:

More investigations have been done for testing the optimum value for the thresholding T that used for the segmentation purposes. Based on the results presented in Table 8, it is quite obvious that the best accuracy rate was achieved when $T = 7$.

5. CONCLUSIONS

Since detecting a disease/cancer and identifying treatments manually is costly, time consuming, and requires professional staff, the evolution of the application of medical image processing in the healthcare field has contributed in an improvement in the quality/accuracy of disease diagnosis (or early detection). Meanwhile, medical image processing techniques can accurately extract target diseases/cancers at higher accuracy and lower cost. Breast cancer is one of the leading causes of mortality among women, compared to all other cancers. Therefore, early detection of breast cancer is necessary to reduce fatalities. Thus, early detection of breast cancer cells may be anticipated using recent machine learning approaches. The primary objective of developing CAD system for mammogram images is to aid physicians and diagnostic experts by providing a second perspective, this increases confidence in the diagnostic process. This study was focused on the development of an efficient CAD system for the early breast cancer detection. The testing findings reveal that the proposed CAD approach obtained an accuracy rate of 99.7% and outperforms the existing approaches.

To improve the performance of the proposed approach, the following are points of potential plans that extend our work in the future: (1) More filters and image processing techniques will be tested for pre-processing purposes to

enhance the image quality, (2) different techniques will be tested to improve segmenting purposes, and (3) different kinds of features should be tested and investigated.

REFERENCES

- [1] A. A. Abdulla. "Efficient computer-aided diagnosis technique for leukaemia cancer detection". *The Institution of Engineering and Technology*, vol. 14, no. 17, pp. 4435-4440, 2020.
- [2] A. A. Abdulla and M. W. Ahmed. "An improved image quality algorithm for exemplar-based image inpainting". *Multimedia Tools and Applications*, vol. 80, pp. 13143-13156, 2021.
- [3] H. Arimura, T. Magome, Y. Yamashita and D. Yamamoto. "Computer-aided diagnosis systems for brain diseases in magnetic resonance images". *Algorithms*, vol. 2, no. 3, pp. 925-952, 2009.
- [4] G. Kumar and P. K. Bhatia. "A Detailed Review of Feature Extraction in Image Processing Systems". *International Conference on Advanced Computing and Communication Technologies ACCT*, pp. 5-12, 2014.
- [5] T. T. Htay and S. S. Maung. "Early Stage Breast Cancer Detection System Using GLCM Feature Extraction and K-Nearest Neighbor (k-NN) on Mammography Image". *2018-The 18th International Symposium on Communications and Information Technologies*, pp. 345-348, 2018.
- [6] Centers for Disease Control and Prevention. "What Is Breast Cancer?". Centers for Disease Control and Prevention, United States. 2021. Available from: https://www.cdc.gov/cancer/breast/basic_info/what-is-breast-cancer.html [Last accessed on 2022 Dec 18].
- [7] J. Vasković. "Overview and Types of Connective Tissue." Medical and Anatomy Experts, 2022. Available from: <https://www.kenhub.com/en/library/anatomy/overview-and-types-of-connective-tissue> [Last accessed on 2022 Dec 20].
- [8] M. M. Eltoukhy, I. Faye and B. B. Samir. "Breast cancer diagnosis in digital mammogram using multiscale curvelet transform". *Computerized Medical Imaging and Graphics*, vol. 34, no. 4, pp. 269-276, 2010.
- [9] S. Srivastava, N. Sharma, S. K. Singh and R. Srivastava. "Design, analysis and classifier evaluation for a CAD tool for breast cancer detection from digital mammograms". *International Journal of Biomedical Engineering and Technology*, vol. 13, no. 3, pp. 270-300, 2013.
- [10] S. C. Satapathy, B. N. Biswal, S. K. Udgata and J. K. Mandal. "Proceedings of the 3rd International Conference on Frontiers of Intelligent Computing: Theory and Applications (FICTA) 2014". *Advances in Intelligent Systems and Computing*, vol. 327, pp. 413-419, 2014.
- [11] S. Pashoutan, S. B. Shokouhi and M. Pashoutan. "Automatic Breast Tumor Classification Using a Level Set Method and Feature Extraction in Mammography." *2017 24th Iranian Conference on Biomedical Engineering and 2017 2nd International Iranian Conference on Biomedical Engineering ICBME 2017*, pp. 1-6, 2018.
- [12] V. Hariraj, W. Khairunizam, V. Vijejan and Z. Ibrahim. "Fuzzy multi-layer SVM classification". *International Journal of Mechanical Engineering and Technology (IJMET)*, vol. 9, pp. 1281-1299, 2018.
- [13] S. J. A. Sarosa, F. Utaminingrum and F. A. Bachtiar. "Breast cancer classification using GLCM and BPNN". *International Journal of Advances in Soft Computing and Its Applications*, vol. 11, no. 3, pp. 157-172, 2019.
- [14] A. Arafa, N. El-Sokary, A. Asad and H. Hefny. "Computer-aided detection system for breast cancer based on GMM and SVM". *Arab Journal of Nuclear Sciences and Applications*, vol. 52, no. 2, pp. 142-150, 2019.
- [15] A. H. Farhan and M. Y. Kamil. "Texture analysis of breast cancer via LBP, HOG, and GLCM techniques". *IOP Conference Series: Materials Science and Engineering*, vol. 928, no. 7, p. 072098, 2020.
- [16] A. S. Eltrass and M. S. Salama. "Fully automated scheme for computer-aided detection and breast cancer diagnosis using digitised mammograms". *IET The Institution of Engineering and Technology*, vol. 14, no. 3, pp. 495-505, 2020.
- [17] E. M. H. Saeed, H. A. Saleh and E. A. Khalel. "Classification of mammograms based on features extraction techniques using support vector machine". *Computer Science and Information Technologies*, vol. 2, no. 3, pp. 121-131, 2020.
- [18] H. Mu'jizah and D. C. R. Novitasari. "Comparison of the histogram of oriented gradient, GLCM, and shape feature extraction methods for breast cancer classification using SVM". *Journal of Technology and Computer Systems*, vol. 9, no. 3, pp. 150-156, 2021.
- [19] G. Holi. "Automatic breast cancer detection with optimized ensemble of classifiers". *International Journal of Advanced Research in Engineering and Technology (IJARET)*, vol. 11, no. 11, pp. 2545-2555, 2020.
- [20] V. R. Nwadike. "What Does Breast Cancer Look Like on a Mammogram?". 2018. Available from: <https://www.medicalnewstoday.com/articles/322068> [Last accessed on 2022 Dec 16].
- [21] K. Bhargavi and S. Jyothi. "A survey on threshold based segmentation technique in image processing". *International Journal of Innovative Research and Development*, vol. 3, no. 12, pp. 234-239, 2014.
- [22] J. Suckling, J. Parker, D. Dance, S. Astley, I. Hutt, C. Boggis, I. Ricketts, E. Stamatakis, N. Cermeaz, N, S. Kok, P. Taylor, D. Betal and J. Savage. "The mammographic image analysis society digital mammogram database". *International Congress Series*, vol. 1069, pp. 375-378, 1994.
- [23] R. Murtirawat, S. Panchal, V. K. Singh and Y. Panchal. "Breast Cancer Detection Using K-Nearest Neighbors, Logistic Regression and Ensemble Learning". *Proceedings of the International Conference on Electronics and Sustainable Communication Systems, ICESC 2020*, pp. 534-540, 2020.