❏ 2923

Texture features analysis technique to detect mass lesion in digitized mammogram images

Ayman A. AbuBaker[1], Yazeed Yasin Ghadi[2]

[1]Department of Electrical Engineering, Faculty of Engineering, Applied Science Private University, Amman, Jordan
[2]Department of Software Engineering and Computer Science Programs, Al Ain University, Abu Dhabi, United Arab Emirates

Article Info	ABSTRACT
	Mass lesions are one of the breast cancer tumors. Mammogram images are the first screening tool to detect tumors in the women breast, but due to radiologist fatigue, number of false positive (FP) and false negative (FN) rates are increased. The main objective of this paper is to develop an intelligent computer aided diagnosis (CAD) system that can accurately detect mass lesions in digitized mammogram images. The proposed method has three stages. The first stage is a preprocessing stage, where the mass lesion is enhanced using a customized Laplacian filter. Then, multi-statistical filters are implemented to detect a potential mass lesion in the mammogram images. In the final stage, the number detected FP regions are reduced using five texture features. The proposed algorithm is evaluated using 45 mammogram images and the algorithm achieved an accuracy rate of 97% in detecting mass lesion with 83% sensitivity rate and 98% specificity rate.
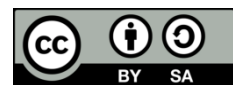
Corresponding Author:

Ayman A. AbuBaker
Department of Electrical Engineering, Faculty of Engineering, Applied Science Private University
Al-Arab St. 21, Amman, Jordan
Email: a_abubaker@asu.edu.jo

1. INTRODUCTION

One of the most common causes of death in women is the breast cancer. In the past years, number of breast cancer cases are increased and specially in young women [1]. Statistically, India nowadays has the larger breast cancer mortality in the world [2]. Breast cancer can be detected using different modalities such as mammography machine, positron emission tomography (PET), magnetic resonance imaging (MRI), and ultrasonography.

Mammography machine is the most reliable and economic as a first stage in detecting breast cancer [3]. Mammogram images can screen mainly three types of breast cancer tumors which are masses, architectural distortion, and bilateral asymmetry of the breast [4]. In mammogram images, the radiologist can distinguish between benign and malignant mass lesion based on their shape, margin, and density properties [5]. Masses that are round in shape, oval and slightly labiate shape are considered as a benign masse region. Whereas, irregular shape, and multi lobular mass may considered as malignancy mass region [6]. Margin is another way to distinguish between benign and malignant regions, where the detected region is considered as benign when it is circumscribed margin but the spiculated or micro lobulated margins are highly considered as malignancy regions [7].

A computer aided diagnosis (CAD) platform has been used to help radiologists in detecting mass lesion accurately. Therefore, many of the authors investigate different intelligent techniques in enhancing the performance of CAD system by increasing sensivity of detecting tumors [8]. Detection of mass clusters in

mammogram images is more difficult than microcalcification since they are highly connected to surrounding parenchymal tissues and they are usually surrounded by non-uniform tissue background with similar characteristics [9], [10]. Therefore, in this paper, the authors develop a novel CAD system that can accurately detect the mass lesion in the mammogram images. The proposed method is divided into three main stages. In the first stage, the mammogram image is preprocessed using Laplacian filter. The preprocessing process will enhance the contrast of mass region in mammogram image. After that as a second stage, the multi-statical filters are implemented to detect potential mass lesion clusters (PMC). Finally, five texture features are implemented to reduce the detected false positive (FP) regions. The paper is organized as follows. Literature review is presented in section 2. Section 3 explains the proposed method. Section 4 discusses the experimental results and section 5 concludes the paper.

2.　LITERATURE REVIEW

Many authors developed different techniques to enhance the performance of CAD system such as [11]. The paper used classification and extreme learning machine (ELM) classifier to detect tumors in mammogram images. The algorithm used five textural and morphological features in detecting regions of interest. Then support vector machine (SVM) is implemented for classification process. Other authors used texture features in CAD system who are [12]. Where Zirinke moments are extracted from the ROI of the images. Then, combination between Zirinke of moments and texture features are implemented to classify FP and true positive (TP) regions. The classification accuracy for their algorithm reaches 94.11%. Genetic algorithm was another tool to accurately detect mass lesions in mammogram images such [13]. They accurately classify regions using genetic algorithm by 10.53% comparing with traditional multi-layer perceptron neural network (MLP-NN). Contourlet transform algorithm and SVM is proposed by Gedik [14] to classify mammogram images. The accuracy for their algorithm reaches 98.467%. Moreover, Hu *et al*. [15] proposed adaptive thresholding segmentation technique. They combine both local threshold and an adaptive global threshold segmentation in detection tumors in mammogram images. The algorithm sensitivity was 91.3% with 0.71 FP per image. According to Surinderan and Vadivel [16], shape features have also been used in classification mass lesions. They used classifier and regression tree (CART) classifier to accurately detect mass lesions in digitized mammogram images. The proposed classifier successfully detects 93.62%. Minavathi *et al*. [17] proposed another contour algorithm to detect mass lesion. They measure curvature angle of each pixel connected in mass lesion boundaries. Their method detect the TP regions with accuracy 92.7% sensitivity with 0.88. Genetic programming (GP) filter is proposed by Uppal [18] to enhance the performance to the classifier of CAD system. They customized a new GP filter to accurately detect mass lesion and they successfully detected them with 96.97% with 98.39% sensitivity and 94.59% specificity. Mehdy *et al*. [19] proposed artificial neural network (ANN) to detect mass lesions in mammogram image. They investigate different types NN and hybrid NN like SOM to build a novel classifier to detect mass lesions. Wang *et al*. used NN [20] to detect mass lesion. They used convolutional neural network (CNN) to which the input consisted of large image window for computerized detection of clustered microcalcifications. The proposed CNN classifier has an accuracy rate of 97.1% in detecting mass lesions. Also, Abdel-Zaher and Eldeib [21] proposed an intelligent classifier using supervised NN in detection breast cancer. Finally, latent dirichlet allocation (LDA) model classifier was proposed by Wang *et al*. [22] to detect breast cancer. The classifier has achieved an accuracy 92.74% in detecting tumors in mammogram images.

3.　METHOD

The purpose of this section is to discuss the methodologies used to detect and enhance the contrast of a mass lesion on mammographic images. According to Figure 1, the proposed algorithm consists of three main stages. In the first stage, Laplacian filter is used to enhance the contrast of the mass lesion. Next, a potential mass lesion region is detected. As a final step, a texture feature algorithm is applied to mammogram images to reduce the total number of detected false positive regions.

3.1. Pre-processing stage

As a first stage, the contrast of the mass lesion is enhanced as a prepossessing stage. Laplacian filter is one of simplest and most effective filter in enhancing the intensity of mammogram images as in (1):

$$S(x,y) = f(x,y) + c[\nabla^2 f(x,y)] \tag{1}$$

where S(x, y) is the intensity value for the processed image, *f(x, y)* the intensity value for the input image and *c* is consider as one in this paper.

The appearance characteristics of mass lesion is a small white region with diameter less than 50 mm appears in the mammogram image [23]. Also, the intensity value of the mass lesion is higher than all the surrounding regions (background). So, mass lesion looks like a peak corresponding to surrounding regions. Therefore, 45 mammogram images are tested to find intensity value for center of the mass lesions and intensity of surrounding regions. It was found that the range between mass lesion and surrounding area is between 40 to 250 grey levels. In accordance with these observations, each mammogram image is processed using the modified average filter that represent on (2):

$$S_K = \frac{1}{mn} \Sigma_{j=40}^{240} r_j \tag{2}$$

Where S_k is the intensity value for the processed image, r_j the intensity value for the input image and m and n are the mask size. Both Laplcian and modified average filters are implemented on 45 mammogram images. The result shows that the contrast of mass lesion is slightly improved as shown in Figures 2(a) and (b).
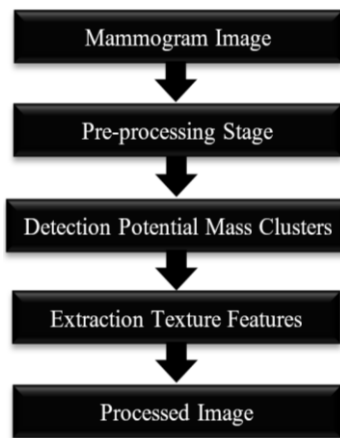

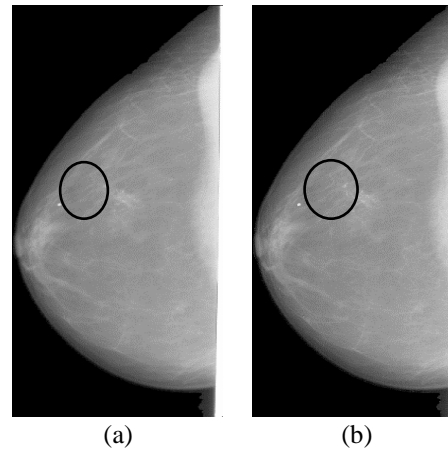Figure 1. Mass detection and classification stages

(a)   (b)

Figure 2. Mass contrast enhancement in mammogram image
(a) original image and (b) enhanced mass image

3.2. Detection potential mass cluster stage

As presented before, the visual characteristics of mass lesion is a small white region in range of 25 to 50 mm [24]. To detect a PMC, two concentric circular masks were designed as shown in Figure 3. So, when average the inner mask is greater than the average of the outer mask this will be considered as a PMC. The size of both inner and outer masks are design based on the image resolution of both databases (University of South of Florida (USF) and mammographic image analysis society (MIAS) databases) which are 45×45 µm [25]. This concentric circular mask has been tested on 45 mammogram images and it was found that all mass lesions are detected but with high number of FP regions. The PMC algorithm is designed based on the fact that intensity value of mass lesion is higher than the surrounding regions as shown in Figure 4. Therefore, when the average of inner mask is greater than the average of outer mask it will be considered as a PMC.

PMC algorithm is implemented on all mammogram images. All mass lesions are accurately detected but unfortunate with a very high number of detected FP regions. This huge number of detected FP regions will decrease the sensitivity of our proposed CAD systems as shown in Figures 5(a) and (b). So, we need another stage to reduce the number of detected FP regions which will be presented in the next section.
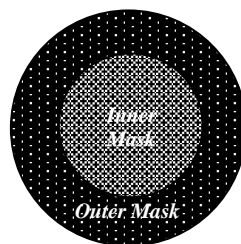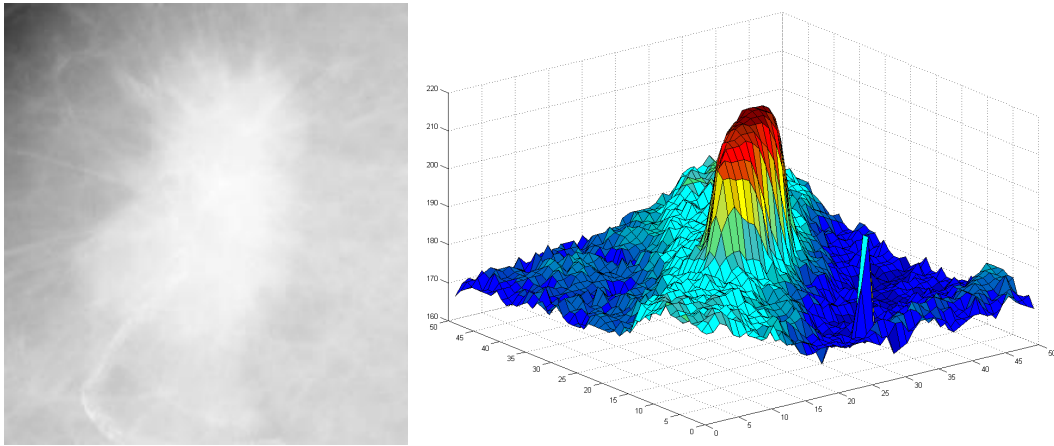

Figure. 3. Two concentric circular masks
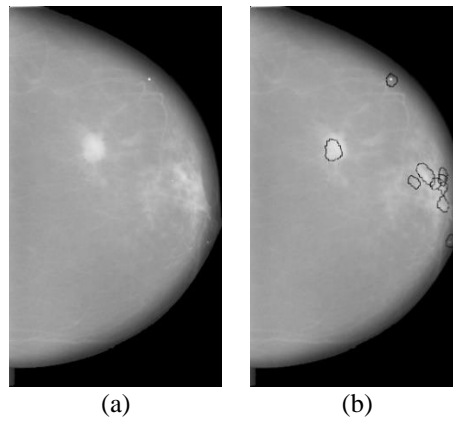
Figure 4. The mesh grid of mass

(a)   (b)

Figure. 5. PMC processing (a) the original dense breast image and (b) the processed dense breast image

3.3. Texture feature extraction

Mass lesions are accurately detected using the PMC algorithm but with a high number of detected FP regions. So, further stage is needed to reduce the number of detected FP regions. In this paper, texture features extraction technique is used to reduce the number of detected FP regions.

Initially, proper samples of malignant and benign datasets were collected from 45 mammogram images. Clusters of size 21×21 are manually cropped from these mammogram images. Therefore, a dataset of 611 expertly identified as an actual TP clusters and 906 an actual FP cluster were collected and organized in two categories (TP and FP clusters). In this paper the first order texture features base are calculated from the intensity histograms renormalized to give the histogram probability P(i) defined in (3) for each actual TP and FP cluster:

$$P(i) = \sum_{i=40}^{240} h(i)/MN \tag{3}$$

where *h(i)* is the intensity histogram and *M, N* are the image region's height and width respectively.

Then five features are calculated and modified comparing with AbuBaker [26] by considering the upper and lower grey levels of the mass in the mammogram images as shown in (4) to (8).

a. The modified mean feature (μ):

$$\mu = \sum_{i=40}^{240} iP(i) \tag{4}$$

b. The modified entropy feature (E)

$$E = -\sum_{i=40}^{240} P(i) \, log_2[P(i)] \tag{5}$$

c. The modified standard deviation feature (σ)

$$\sigma = \sqrt{\sum_{i=40}^{240}(i - \mu)^2 P(i)} \tag{6}$$

d. The modified third order of moment feature (M3)

$$M_3 = \sum_{i=40}^{240}(i - \mu)^3 P(i) \tag{7}$$

e. The modified kurtosis feature (K)

$$K = \sigma^{-4} \sum_{i=40}^{240}(i - \mu)^4 P(i) - 3 \tag{8}$$

4. RESULTS AND DISCUSSION

The proposed smart classifier was evaluated using 45 mammogram images from two databases (USF and MIAS database). At the beginning, the PMC algorithm is implemented to accurately detect mass lesion in mammogram images. Then a smart classifier that used five texture features (entropy, mean, standar deviation (STD), kurtosis, and skewness) was implemented to reduce the number of detected FP clusters. The results were compared with different author's results as shown in Table 1.

It is also worth mentioning that the TP and FP rates in these publications are reported for different mammogram images and using different benchmarks. In addition, most of these algorithms are tested on the MiniMIAS database, which contains smaller images, mainly 1,024×1,024, compared to the MIAS images we have used for evaluating our algorithm, which range from 623×1,774 to 1,269×2,132. Hence, we have conducted additional evaluations as described.

Table 1. The comparison results using other author's algorithms

Algorithm	TP	FP
Brijesh and Ping [27]	85	NA
Linguraru *et. al.* [28]	91	0.95 per image
Peng *et. al.* [29] KD-GA	98.9	40% FPF
Peng *et. al.* [29], GA	85	20% FPF
Zhang *et. al.* [30] discriminate	70	NA
Zhang *et. al.* [30], logistic regression	70	NA
Sentelle and Sutton [31]	94	17 per image
Malar *et. al.* [32]	94	NA
Mohanalin *et. al.* [33]	96.55	0.4 per image
Oliver *et. al.* [34]	80	1 per image
Rizzi *et. al.* [35]	98	1 per image
Ravi *et. al.* [36]	91	1.63 per image
The proposed algorithm	97.6	0.6 per image

From Table 1, it is clearly noticed that performance of the proposed algorithm is better than other algorithms in both TP and FP clusters. The algorithm achieved an accuracy rate around 97.6% with very low FP regions which is 0.6 per image. Figures 6(a)-(d) show some images with results from the mass detection algorithm superimposed.

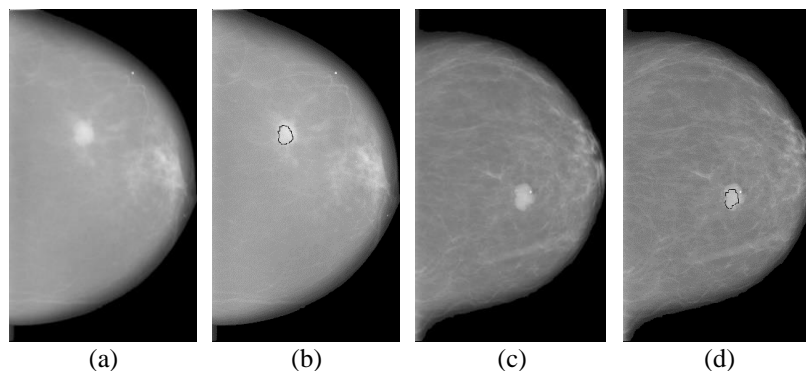<div align="center">(a)   (b)   (c)   (d)</div>

Figure 6. Accurate detection of mass lesions (a) original fatty breast, (b) processed image, (c) original dense breast, and (d) processed image

5. CONCLUSION

A new novel technique to accurately detect and classify the mass lesion in mammogram images is proposed in this paper. Initially, the contrast of mass lesion is enhanced using a customized Laplacian filter. The enhancement technique is slightly enhancing the contrast on mass lesion and it become more brighter than the surrounding regions. After that, two concentric circular masks were implemented to detect all the peaks in the mammogram images. These detected peaks are considered as PMC since all mass lesions are detected. Finally, five texture features were used to reduce the number of detected FP regions. The algorithm is evaluated by processing 45 mammogram cases from two databases (MIAS and USF database). In all cases the algorithm can successfully detect the mass lesion with an accuracy rate of 97.6% with minimum number of FP region 0.6.

REFERENCES

[1] P. Mathur *et al.*, "Cancer Statistics, 2020: Report From National Cancer Registry Programme, India," *JCO Global Oncology*, nol. 6, pp. 1063-1075, 2020, doi: 10.1200/GO.20.00122.

[2] R. Siegel, D. Naishadham, and A. Jemal, "Cancer statistics, 2013," *CA: A Cancer Journal for Clinicians*, vol. 63, no. 1, pp. 11–30, Jan. 2013, doi: 10.3322/caac.21166.

[3] C. DeSantis, J. Ma, L. Bryan, and A. Jemal, "*CA A Cancer Journal for Clinicians*," Breast cancer statistics, vol 64, pp. 52-62, 2013, doi: 10.3322/caac.21203.

[4] W. E. Barlow, *et al.*, "Performance of Diagnostic Mammography for Women With Signs or Symptoms of Breast Cancer," *JNCI: Journal of the National Cancer Institute*, vol. 94, no.15, pp. 1151–1159, 2002, doi: 10.1093/jnci/94.15.1151.

[5] M. J. Homer, "Breast imaging, standard of care, and the expert," *Radiologic Clinics of North America*, vol. 42, no. 5, pp. 963–974, Sep. 2004, doi: 10.1016/j.rcl.2004.03.012.

[6] R. W. Woods, G. S. Sisney, L. R. Salkowski, and K. Shinki, Y. Lin, E. S. Burnside, "The Mammographic Density of a Mass Is a Significant Predictor of Breast Cancer," *Radiology*, vol. 258, no. 2, pp. 417-425, 2011, doi. 10.1148/radiol.10100328.

[7] S. Banik, R. M. Rangayyan, and J. E. L. Desautels, "Measures of angular spread and entropy for the detection of architectural distortion in prior mammograms," *International Journal of Computer Assisted Radiology and Surgery*, vol. 8, no. 1, pp. 121–134, Jan. 2013, doi: 10.1007/s11548-012-0681-x.

[8] R. M. Rangayyan, R. J. Ferrari, and A. F. Frere, "Analysis of bilateral asymmetry in mammograms using directional, morphological, and density features," *Journal of Electronic Imaging*, vol. 16, no. 1, pp. 1–12, 2007, doi: 10.1117/1.2712461.

[9] L. Wang, "Early diagnosis of breast cancer," *Sensors*, vol. 17, no. 7, pp. 1–20, Jul. 2017, doi: 10.3390/s17071572.

[10] K. U. Sheba and S. G. Raj, "Objective quality assessment of image enhancement methods in digital mammography–a comparative study," *Signal & Image Processing : An International Journal*, vol. 7, no. 4, pp. 1–13, 2016, doi: 10.5121/sipij.2016.7401.

[11] Z. Wang, G. Yu, Y. Kang, Y. Zhao, and Q. Qu, "Breast tumor detection in digital mammography based on extreme learning machine," *Neurocomputing*, vol. 128, pp. 175–184, Mar. 2014, doi: 10.1016/j.neucom.2013.05.053.

[12] S. M. L. de Lima, A. G. da Silva-Filho, and W. P. dos Santos, "Detection and classification of masses in mammographic images in a multi-kernel approach," *Computer Methods and Programs in Biomedicine*, vol. 134, pp. 11–29, Oct. 2016, doi: 10.1016/j.cmpb.2016.04.029.

[13] P. Valarmathi and S. Robinson, "An improved neural network for mammogram classification using genetic optimization," *Journal of Medical Imaging and Health Informatics*, vol. 6, no. 7, pp. 1631–1635, Nov. 2016, doi: 10.1166/jmihi.2016.1862.

[14] N. Gedik, "Breast cancer diagnosis system via contourlet transform with sharp frequency localization and least squares support vector machines," *Journal of Medical Imaging and Health Informatics*, vol. 5, no. 3, pp. 497–505, Jun. 2015, doi: 10.1166/jmihi.2015.1422.

[15] K. Hu, X. Gao, and F. Li, "Detection of suspicious lesions by adaptive thresholding based on multiresolution analysis in mammograms," *IEEE Transactions on Instrumentation and Measurement*, vol. 60, no. 2, pp. 462–472, Feb. 2011, doi: 10.1109/TIM.2010.2051060.

[16] B. Surendiran and A. Vadivel, "Mammogram mass classification using various geometric shape and margin features for early detection of breast cancer," *International Journal of Medical Engineering and Informatics*, vol. 4, no. 1, pp. 36–54, 2012, doi: 10.1504/IJMEI.2012.045302.

[17] M. Minavathi, S. Murali, and M. S. Dinesh, "Classification of mass in breast ultrasound images using image processing techniques," *International Journal of Computer Applications*, vol. 42, no. 10, pp. 29–36, Mar. 2012, doi: 10.5120/5730-7801.

[18] M. T. N. Uppal, "Classification of mammograms for breast cancer detection using fusion of discrete cosine transform and discrete wavelet transform features," *Biomedical Research (India)*, vol. 27, no. 2, pp. 322–327, 2016.

[19] M. M. Mehdy, P. Y. Ng, E. F. Shair, N. I. M. Saleh, and C. Gomes, "Artificial neural networks in image processing for early detection of breast cancer," *Computational and Mathematical Methods in Medicine*, pp. 1–15, 2017, doi: 10.1155/2017/2610628.

[20] J. Wang, R. M. Nishikawa, and Y. Yang, "Global detection approach for clustered microcalcifications in mammograms using a deep learning network," *Journal of Medical Imaging*, vol. 4, no. 2, pp. 1–10, Apr. 2017, doi: 10.1117/1.JMI.4.2.024501.

[21] A. M. Abdel-Zaher and A. M. Eldeib, "Breast cancer classification using deep belief networks," *Expert Systems with Applications*, vol. 46, pp. 139–144, Mar. 2016, doi: 10.1016/j.eswa.2015.10.015.

[22] Y. Wang, J. Li, and X. Gao, "Latent feature mining of spatial and marginal characteristics for mammographic mass classification," *Neurocomputing*, vol. 144, pp. 107–118, Nov. 2014, doi: 10.1016/j.neucom.2013.11.050.

[23] H. S. Sheshadri and A. Kandaswamy, "Computer–aided diagnosis of digital mammograms," *Information Technology Journal*, vol. 5, no. 2, pp. 342–346, Feb. 2006, doi: 10.3923/itj.2006.342.346.

[24] X. Zhang and H. Xie, "Mammograms enhancement and denoising using generalized gaussian mixture model in nonsubsampled contourlet transform domain," *Journal of Multimedia*, vol. 4, no. 6, pp. 389–396, Dec. 2009, doi: 10.4304/jmm.4.6.389-396.

[25] A. A. AbuBaker, R. S. Qahwaji, M. J. Aqel, H. Al-Osta, and M. H. Saleh, "Efficient pre-processing of USF and MIAS mammogram images," *Journal of Computer Science*, vol. 3, no. 2, pp. 67–75, Feb. 2007, doi: 10.3844/jcssp.2007.67.75.

[26] A. AbuBaker, "Automatic detection of breast cancer microcalcifications in digitized x-ray mammograms," Ph.D. dissertation, School of Inform., Univ. of Bradford, Bradford, United Kingdom, 2008.

[27] B. Verma and P. Zhang, "A novel neural-genetic algorithm to find the most significant combination of features in digital mammograms," *Applied Soft Computing*, vol. 7, no. 2, pp. 612–625, Mar. 2007, doi: 10.1016/j.asoc.2005.02.008.

[28] M. G. Linguraru, K. Marias, R. English, and M. Brady, "A biologically inspired algorithm for microcalcification cluster detection," *Medical Image Analysis*, vol. 10, no. 6, pp. 850–862, Dec. 2006, doi: 10.1016/j.media.2006.07.004.

[29] Y. Peng, B. Yao, and J. Jiang, "Knowledge-discovery incorporated evolutionary search for microcalcification detection in breast cancer diagnosis," *Artificial Intelligence in Medicine*, vol. 37, no. 1, pp. 43–53, May 2006, doi: 10.1016/j.artmed.2005.09.001.

[30] P. Zhang, B. Verma, and K. Kumar, "Neural vs. statistical classifier in conjunction with genetic algorithm based feature selection," *Pattern Recognition Letters*, vol. 26, no. 7, pp. 909–919, May 2005, doi: 10.1016/j.patrec.2004.09.053.

[31] S. Sentelle, C. Sentelle, and M. A. Sutton, "Multiresolution-based segmentation of calcifications for the early detection of breast cancer," *Real-Time Imaging*, vol. 8, no. 3, pp. 237–252, Jun. 2002, doi: 10.1006/rtim.2001.0285.

[32] E. Malar, A. Kandaswamy, D. Chakravarthy, and A. G. Dharan, "A novel approach for detection and classification of mammographic microcalcifications using wavelet analysis and extreme learning machine," *Computers in Biology and Medicine*, vol. 42, no. 9, pp. 898–905, Sep. 2012, doi: 10.1016/j.compbiomed.2012.07.001.

[33] Mohanalin, Beenamol, P. K. Kalra, and N. Kumar, "A novel automatic microcalcification detection technique using Tsallis entropy & a type II fuzzy index," *Computers & Mathematics with Applications*, vol. 60, no. 8, pp. 2426–2432, Oct. 2010, doi: 10.1016/j.camwa.2010.08.038.

[34] A. Oliver *et al.*, "Automatic microcalcification and cluster detection for digital and digitised mammograms," *Knowledge-Based Systems*, vol. 28, pp. 68–75, Apr. 2012, doi: 10.1016/j.knosys.2011.11.021.

[35] M. Rizzi, M. D'Aloia, and B. Castagnolo, "Computer aided detection of microcalcifications in digital mammograms adopting a wavelet decomposition," *Integrated Computer-Aided Engineering*, vol. 16, no. 2, pp. 91–103, Mar. 2009, doi: 10.3233/ICA-2009-0306.

[36] R. K. Samala *et al.*, "Detection of microcalcifications in breast tomosynthesis reconstructed with multiscale bilateral filtering regularization," in *Medical Imaging 2013: Computer-Aided Diagnosis*, Mar. 2013, pp. 1–8, doi: 10.1117/12.2008230.

BIOGRAPHIES OF AUTHORS

Ayman A. AbuBaker 🆔 has obtained his Ph.D. in the area of in Electronic Imaging and Media Communications (EIMC) from School of Informatics, University of Bradford-UK. Currently, he is a professor in image processing and machine learning at the Faculty of Engineering and Technology, Applied Science Private University, Jordan. He has published more than 70 research articles/papers in journals, and conferences in his main research interest such as medical images processing including mammogram images (detecting microcalcifications and mass lesions), ultrasound and MRI. Ayman A. AbuBaker has been actively working with several national professional communities and industries in the area of engineering education and accreditation such as ABET accreditation. Currently he is the dean of Faculty of Engineering and Technology at Applied Science Private University, Jordan. He can be contacted at email: a_abuabker@asu.edu.jo.

Yazeed Yasin Ghadi 🆔 received the Ph.D. degree from the Department of Biomedical Engineering, Central Queensland University, Australia. He was working as a Postdoctoral Research Fellow at CQU. He is currently an Assistant Professor with the Department of Computer Science and Software Engineering, Al Ain University, UAE. His research interests include multimedia contents and artificial intelligence, machine learning, computer vision, fuzzy logic, and renewable energy. He can be contacted at email: yazeed.ghadi@aau.ac.ae.