Computer-aided diagnosis system for breast cancer based on the Gabor filter technique

Mohammed Y. Kamil
College of Sciences, Mustansiriyah University, Iraq

Article Info	ABSTRACT
	The most prominent reason for the death of women all over the world is breast cancer. Early detection of cancer helps to lower the death rate. Mammography scans determine breast tumors in the first stage. As the mammograms have slight contrast, thus, it is a blur to the radiologist to recognize micro growths. A computer-aided diagnostic system is a powerful tool for understanding mammograms. Also, the specialist helps determine the presence of the breast lesion and distinguish between the normal area and the mass. In this paper, the Gabor filter is presented as a key step in building a diagnostic system. It is considered a sufficient method to extract the features. That helps us to avoid tumor classification difficulties and false-positive reduction. The linear support vector machine technique is used in this system for results classification. To improve the results, adaptive histogram equalization pre-processing procedure is employed. Mini-MIAS database utilized to evaluate this method. The highest accuracy, sensitivity, and specificity achieved are 98.7%, 98%, 99%, respectively, at the region of interest (30×30). The results have demonstrated the efficacy and accuracy of the proposed method of helping the radiologist on diagnosing breast cancer.

Corresponding Author:

Mohammed Y. Kamil,
College of Sciences,
Mustansiriyah University,
Baghdad, Iraq.
Email: m80y98@uomustansiriyah.edu.iq

1. INTRODUCTION

Breast cancer is the chief cause of death ladies in their forties throughout the world [1]. Diagnosis of breast cancer earlier is necessary for improving active treatment approaches and decreasing the fatality rate [2]. Imaging systems become presented as an essential part of early diagnosis [3]. A common safety technique for recognizing breast tumors in early-stage is mammograms [4]. A mammogram is the most helpful screening technique for identifying breast tumors earlier. It utilizes a lower dosage of x-ray to recognize the tumor [5]. The specialist practice contributes to increasing the accuracy of using mammograms to determine breast cancer. To increases the patient's remainder probability including the idea of greater identification of breast, mammography improvements by the combination of the computer-aided diagnosis (CADs) system [6].

Preprocessing, feature extraction, and classification are the three important steps in a computer-aided detection technique [7]. Texture feature extraction techniques can be utilized to denote masses in mammograms more precisely [8]. As region of interests (ROIs) show texture in various directions and scales, so it can be described efficiently applying Gabor filters. There are several studies in the field of mammogram investigation for recognizing the tumors, including the Gabor filter technology, to extract features [9, 10].

Wei et al., [11] introduced to compute features to describe the textural pattern of the mammogram by the Gabor filter. Also, this study has shown that discriminated features can effectively to improve

performance in the image retrieval system. Lahmiri and Boukadoum [12] offered distinct techniques to extract features from mammograms depending on a combination of the DWT and the Gabor filter. First, the two-dimensional DWT is applied to prepare the image. A Gabor filter is employed on the mammograms at various wavelengths and directions. The features are computed from this image. Lastly, to distinguish normal and tumor images, certain features are provided to the support vector machine (SVM) classifier. Classification results prove the effectiveness of the method used. Hussain et al., [13] presented Gabor filter banks for deriving the most utmost characteristic and distinguish textural features of tumors in the image at different directions and frequencies to reduce FP and FN. The method is estimated on 512 ROIs selected from digital images of the DDSM database. Gabor filter banks applied on ROI at various directions and wavelengths. A robust feature selection system and SVM classifier used with 10-fold cross-validation. The obtained result achieves Az is 0.995. Khan et al., [14] obtained the orientation of textural characteristics by applying a sequence of Gabor filters at various directions and wavelengths. The structural characteristics of tumors and normal images in mammograms described by the orientation of textural features. ROIs selected from the MIAS database are utilized to estimate this method. (SELwSVM) is employed to classify mammograms. The mean accuracy achieved by this system varies from 68 to 100%. Zheng [15] suggested Gabor cancer detection (GCD) as a distinct breast tumor diagnosis method, using Gabor characteristics. GCD method included three important levels are preprocessing, segmentation, and classification (decreasing false signals). To decrease the false signals, fuzzy C-means clustering system and (KNN) classifier are employed. The best result of GCD algorithm which examined on the DDSM database is 90%.

In this research, we offer an approach for breast mass diagnosis by examining the local textural characteristics of the tumors. For this, we utilize the Gabor filter to obtain the texture features. These features are important for accurately recognizing true tumors and decrease the false-positive diagnosis. We apply this method on a set of ROIs derived from the mini-MIAS database. The derived features are provided as information to the classifier to examine the input ROIs and classify them into exact masses and normal tissues.

2. GABOR FILTER

The texture is the essential feature for distinguishing the ROIs of different classes of images. Texture investigation is necessary for computerized analysis for distribution [16]. Tumors in a ROIs include micro-patterns in various frequencies and directions. These patterns are important in the identification of destructive regions in a Computer-aided diagnosis system. Gabor filters can be efficiently utilized to identify these patterns [17]. Gabor filters are linear filters utilized in many employments in the domains of computer vision problems, such as in texture analysis, face identification, and cancer diagnosis [18]. An impressive characteristic of Gabor filters has the best combined localization in frequency and spatial domains [19]. To obtain the value of mammogram, Gabor filters of distinguished direction used to transfer mammogram by determining the best magnitude of Gabor filter parameters, and the normalized mammogram is the output [20]. In this work, we offered the magnitude of the Gabor filter in figure. These filters represented in complex mathematics as:

$$g = \exp\left[-\frac{(x\cos\theta + y\sin\theta)^2 + \gamma^2(y\cos\theta - x\sin\theta)^2}{2\sigma^2}\right] . \exp\left[i\left[\frac{2\pi(x\cos\theta + y\sin\theta)}{\lambda} + \emptyset\right]\right] \quad (1)$$

where θ is the filter orientation, γ is the spatial aspect ratio, λ is the sinusoidal wavelength, σ is the standard deviation of Gaussian function, and $\emptyset$ is the phase offset.

3. PROPOSED METHOD

In this part, we handle every step of our approach to the diagnosis of a breast mass. The initial stage includes the mammogram acquisition, following we extract ROI from the mammogram, after that the enhancement process is employing to develop mammograms. The feature extraction step involves Gabor filters for the representative ROI at various wavelengths and directions. Finally, ROIs classification into normal and abnormal.

The mammograms are collected from the mini-MIAS database [21]. This database contains 322 mammograms from 161 women; mini-MIAS include normal and abnormal mammograms, the abnormal mammograms categorized to benign and cancerous. The dataset presents a report about estimated the position and radius (in pixels) of the mass indicated by the radiologist (ground truth). The mammogram dimension (image size) is 1024×1024 pixels. This database includes information about mammograms,

for instance, the kind of mass: ill-defined (ild.), asymmetry (asy.), architectural distortion (arch.), spiculated (spi.), circumscribed (cir.), and calcification (cal.).

The tumor zones selected based on the mini-MIAS tumor, indicating which represents the specialist analyzing and show the x-axis and y-axis for a specific tumor in the mammogram [22]. These indications are ground truth (GT) and determined as a rectangular region. To improve classification accuracy, ROI cropped manually inside GT [23]. ROI established as a square region. The normal mammogram, ROI selected manually from arbitrary locations. Extract ROI has shown in Figure 1.

In order to improve the features and to sharpen the details in the image. Thus, we obtain high-performance results and increase the accuracy of the diagnosis. The enhancement step in CAD system is one of the important stages that determine the performance of the algorithm [24]. Here, we used the adaptive histogram equalization (AHE) method to enhance ROI. AHE assists in improving the contrast of each pixel. This technique calculates the various histograms, individually identical to the distinguished Part of the image identified as tiles. Every tile's contrast is improved to redistribute the grayscale of the image. The adjacent tiles later connected utilizing bilinear interpolation to reduce artificially produced edges. Figure 2 exposes the ROI before and after implementing the AHE system. The linear SVM (LSVM) is employed as a classifier. The linear function basis of the practice of LSVM in a high dimensional characteristic term that gets an optimal separating hyper plane [25]. The SVM produces great accuracy compared with other sets of systems.
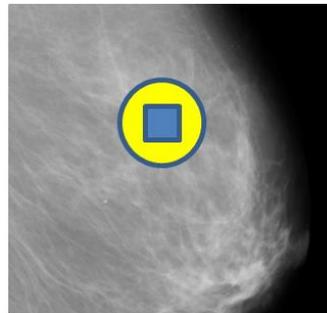
Figure 1. Ground truth ⬤ , ROI ⬛
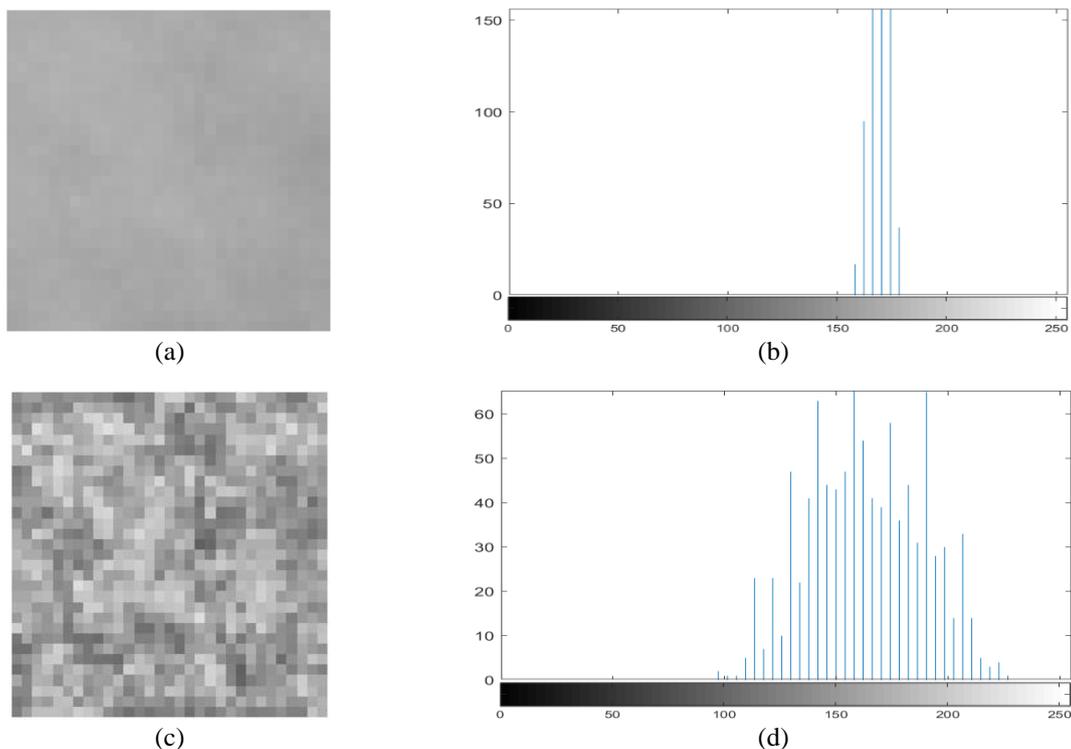
Figure 2. (a) Original ROI with (b) its histogram; and (c) ROI enhanced with (d) its histogram

4.   RESULTS AND DISCUSSIONS

The suggested method is tested by the mini-MIAS database. In mini-MIAS, tumor regions are selected (ground truth), these regions are manually cropped to generate ROIs within the ground truth in three dimensions (10×10, 20×20, 30×30) pixels base on the dimensions of tumor in mammograms. ROIs are defined manually from optional regions in normal cases with the same dimensions of abnormal cases. In the preprocessing step, the AHE method applied to enhance ROI and raise the accuracy of the analysis of breast tumors. After that, we employed the Gabor filter for all mammograms to extract texture features. In this research, 4 orientations (0, 45, 90, 135) and 3 wavelengths (30, 50, 70) used. The Gabor filter implementation represented in Figure 3.

wavelength	Orientation			
	0˚	45˚	90˚	135˚
30				
50				
70				
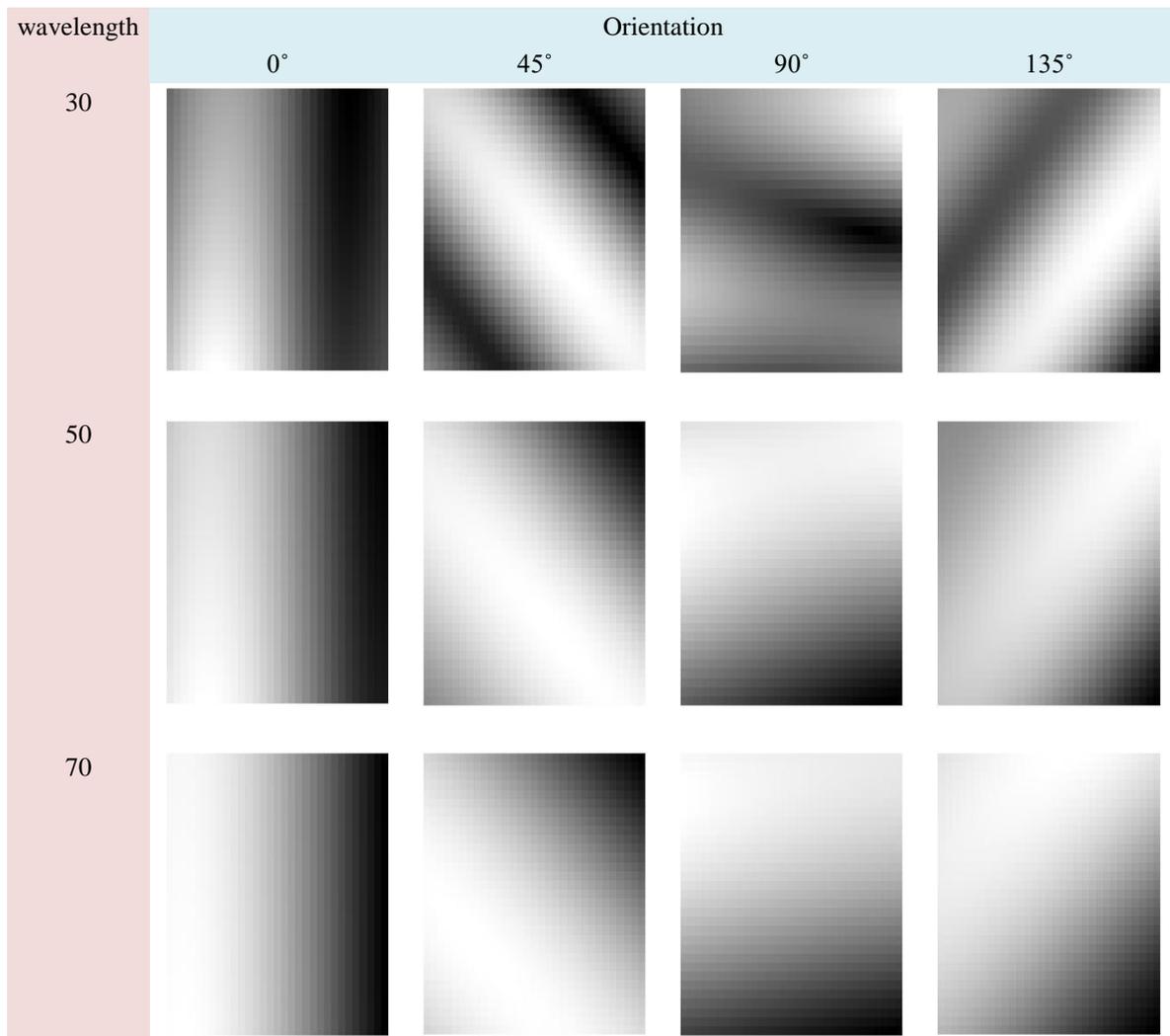
Figure 3. The magnitude of the Gabor filter for an image selected from database

Four features we derived when implementing the Gabor filters are contrast, correlation, energy, and homogeneity. The ROI is characterized as normal or mass by utilizing the LSVM classifier. The LSVM is trained with the derived features and using the practiced magnitudes. Testing features recognize the normal tissue and the tumor. Tables 1-3 show the representation of the Gabor design with LSVM and 5-fold cross-validation.

Table 1 shows the highest accuracy, 98.7% obtained in the ROI (30×30) at wavelength 50 and orientations 0˚ for ill-defined disease. While, the lowest accuracy is 72.7% in the ROI (30×30) at wavelength 50 and orientations 45˚for architectural distortion disease. The highest sensitivity and specificity achieved by applying the proposed classification is 100%.

Table 1. The accuracy provided by using the LSVM classifier at wavelengths (30, 50, 70)
and orientations (0, 45, 90,135)

	ROI type	arch.	asy.	cal.	cir.	ild.	spi.	wavelength
Orientation = 0°	10 × 10	84.6	82.7	81.4	76.7	85.5	79.5	30
	20 × 20	98.1	95.5	93.9	96.7	93	98.1	
	30 × 30	94.7	86.4	86.8	90	93	93.1	
	10 × 10	82.9	84.5	85	76.7	85.5	84.6	50
	20 × 20	96.4	90	90.4	98.3	96.8	96.4	
	30 × 30	86.3	84.5	81.4	86.7	98.7	86.3	
	10 × 10	86.3	86.4	79.6	76.7	85.5	81.2	70
	20 × 20	84.7	90	81.4	95	94.9	91.4	
	30 × 30	89.7	82.7	79.6	88.3	87.4	89.7	
Orientation = 45°	10 × 10	79.5	91.8	81.4	88.3	85.5	77.8	30
	20 × 20	77.8	95.5	88.6	88.3	87.4	89.7	
	30 × 30	77.8	82.7	85	83.3	85.5	77.8	
	10 × 10	79.5	90	85	86.7	87.4	77.8	50
	20 × 20	79.5	84.5	85	86.7	89.2	84.6	
	30 × 30	72.7	82.7	85	85	85.5	84.6	
	10 × 10	74.4	84.5	81.4	88.3	85.5	77.8	70
	20 × 20	79.5	84.5	88.6	90	93	77.8	
	30 × 30	81.2	82.7	81.4	80	85.5	82.9	
Orientation = 90°	10 × 10	79.5	84.5	81.4	78.3	91.1	79.5	30
	20 × 20	82.9	91.8	81.4	78.3	96.8	84.6	
	30 × 30	86.3	79.5	83.2	78.3	91.1	89.7	
	10 × 10	81.2	84.5	83.2	78.3	89.2	79.5	50
	20 × 20	86.3	86.4	85	85	94.9	86.3	
	30 × 30	88	93.6	81.4	85	89.2	82.9	
	10 × 10	77.8	86.4	81.4	76.7	87.4	77.8	70
	20 × 20	91.4	84.5	83.2	85	94.9	91.4	
	30 × 30	83.1	95.5	81.4	81.7	91.1	88	
Orientation = 135°	10 × 10	74.4	84.5	81.4	85	85.5	79.5	30
	20 × 20	75.1	88.2	88.6	85	85.5	77.8	
	30 × 30	76.1	88.2	83.2	80	85.5	81.2	
	10 × 10	74.5	80.9	81.4	76.7	85.5	79.5	50
	20 × 20	77.8	84.5	83.2	88.3	85.5	84.6	
	30 × 30	76.1	84.5	81.4	86.7	89.2	77.8	
	10 × 10	79.5	82.7	76.1	80	93	79.5	70
	20 × 20	80.9	84.5	83.2	86.7	85.5	81.2	
	30 × 30	77.8	86.4	86.8	85	85.5	77.8	

Table 2. The sensitivity provided by using the LSVM classifier at wavelengths (30, 50, 70)
and orientations (0, 45, 90,135)

	ROI type	arch.	asy.	cal.	cir.	ild.	spi.	wavelength
Orientation = 0°	10 × 10	42	23	35	25	27	26	30
	20 × 20	84	77	73	80	41	84	
	30 × 30	95	23	29	60	64	78	
	10 × 10	63	30	29	18	79	47	50
	20 × 20	84	57	66	90	97	84	
	30 × 30	73	43	23	80	98	78	
	10 × 10	42	30	23	20	28	47	70
	20 × 20	84	83	85	95	72	68	
	30 × 30	84	43	29	65	33	68	
Orientation = 45°	10 × 10	73	67	45	65	28	25	30
	20 × 20	57	87	51	85	58	78	
	30 × 30	52	32	39	45	28	28	
	10 × 10	52	87	33	85	35	28	50
	20 × 20	52	73	64	70	66	68	
	30 × 30	31	40	39	80	30	57	
	10 × 10	36	67	51	85	28	25	70
	20 × 20	52	53	76	80	58	52	
	30 × 30	41	67	33	45	33	67	

Table 2. The sensitivity provided by using the LSVM classifier at wavelengths (30, 50, 70)
and orientations (0, 45, 90,135) (*continue*)

	type	arch.	asy.	cal.	cir.	ild.	spi.	wavelength
	ROI			Sensitivity %				
Orientation = 90°	10 × 10	57	60	45	50	66	46	30
	20 × 20	41	80	58	30	82	73	
	30 × 30	52	36	45	35	58	78	
	10 × 10	31	33	26	30	66	36	50
	20 × 20	83	47	70	70	74	73	
	30 × 30	67	87	26	60	51	73	
	10 × 10	31	80	28	25	66	25	70
	20 × 20	99	47	26	70	89	78	
	30 × 30	62	87	51	60	58	73	
Orientation = 135°	10 × 10	31	73	26	85	25	73	30
	20 × 20	25	67	58	85	28	28	
	30 × 30	41	73	39	65	32	57	
	10 × 10	25	67	58	65	30	36	50
	20 × 20	25	67	26	95	28	78	
	30 × 30	25	40	33	95	58	46	
	10 × 10	62	33	45	75	74	25	70
	20 × 20	53	60	45	75	28	62	
	30 × 30	67	60	51	100	35	32	

Table 3. The specificity provided by using the LSVM classifier at wavelengths (30, 50, 70)
and orientations (0, 45, 90,135)

	type	arch.	asy.	cal.	cir.	ild.	spi.	wavelength
	ROI			Specificity %				
Orientation = 0°	10 × 10	95	95	60	70	99	95	30
	20 × 20	95	93	93	95	99	95	
	30 × 30	88	100	100	95	93	90	
	10 × 10	83	95	98	100	98	93	50
	20 × 20	93	93	90	93	100	93	
	30 × 30	83	90	93	80	99	80	
	10 × 10	98	98	93	95	100	88	70
	20 × 20	90	83	70	85	93	93	
	30 × 30	83	80	90	73	95	90	
Orientation = 45°	10 × 10	78	95	80	95	100	98	30
	20 × 20	83	93	98	85	90	90	
	30 × 30	85	100	98	98	100	100	
	10 × 10	88	85	100	83	98	100	50
	20 × 20	88	83	88	90	90	90	
	30 × 30	88	83	98	83	100	93	
	10 × 10	88	85	88	85	98	98	70
	20 × 20	88	90	88	90	98	85	
	30 × 30	95	70	90	93	100	85	
Orientation = 90°	10 × 10	85	88	63	88	93	90	30
	20 × 20	98	90	93	98	95	85	
	30 × 30	98	95	93	95	95	90	
	10 × 10	100	98	100	98	90	95	50
	20 × 20	83	95	85	88	95	88	
	30 × 30	93	90	80	93	95	83	
	10 × 10	95	83	98	100	88	98	70
	20 × 20	83	93	100	88	90	93	
	30 × 30	85	93	88	88	95	90	
Orientation = 135°	10 × 10	90	83	88	80	100	78	30
	20 × 20	95	90	95	80	100	100	
	30 × 30	88	88	95	83	100	88	
	10 × 10	93	80	85	78	100	95	50
	20 × 20	98	85	100	80	100	83	
	30 × 30	95	95	93	78	93	88	
	10 × 10	83	88	83	78	93	100	70
	20 × 20	85	88	93	88	100	85	
	30 × 30	78	90	95	73	100	100	

5.   CONCLUSION

In this study, we have examined the Gabor filter as a technique for avoiding tumor classification difficulties and FP reduction. This method has employed to excerpt the features from textural ROIs at various wavelengths and orientations. The features obtained based on the Gabor filter shown to strongly distinguish between the normal and abnormal tissues utilized in the tests and improves the identification rate of the breast tumor diagnosis system. This filter estimated upon ROI mammograms derived from the mini-MIAS database. AHE has used as a pre-processing procedure to enhance local gray levels values of ROIs. To recognize ROI as normal or mass, LSVM is employed. This classifier produces the most approving results at ROI (30×30), orientation 0˚, and wavelength 50, where the accuracy, sensitivity, specificity is 98.7%, 98%, 99%, respectively. In the CAD system, the analysis of breast mass can be developed by the combination Gabor filter method with other feature extraction techniques.

REFERENCES

[1]   A. Unni, N. Eg, S. Vinod, and L. S. Nair, "Tumour detection in double threshold segmented mammograms using optimized GLCM features fed SVM," in *2018 international conference on advances in computing, communications and informatics (ICACCI)*, pp. 554-559, 2018.

[2]   N. F. Lattoofi, et al., "Melanoma Skin Cancer Detection Based on ABCD Rule," *2019 First International Conference of Computer and Applied Sciences (CAS)*, pp. 154-157, 2019.

[3]   M. Y. Kamil, "Morphological gradient in brain magnetic resonance imaging based on intuitionistic fuzzy approach," in *2016 Al-Sadeq International Conference on Multidisciplinary in IT and Communication Science and Applications (AIC-MITCSA)*, pp. 1-3, 2016.

[4]   S. B. Y. Tasdemir, K. Tasdemir, and Z. Aydin, "ROI Detection in Mammogram Images Using Wavelet-Based Haralick and HOG Features," in *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pp. 105-109, 2018.

[5]   C. Abirami, R. Harikumar, and S. S. Chakravarthy, "Performance analysis and detection of micro calcification in digital mammograms using wavelet features," in *2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, pp. 2327-2331, 2016.

[6]   S. Pashoutan, S. B. Shokouhi, and M. Pashoutan, "Automatic Breast Tumor Classification Using a Level Set Method and Feature Extraction in Mammography," in *2017 24th National and 2nd International Iranian Conference on Biomedical Engineering (ICBME)*, pp. 1-6, 2017.

[7]   F. Saki, A. Tahmasbi, H. Soltanian-Zadeh, and S. B. Shokouhi, "Fast opposite weight learning rules with application in breast cancer diagnosis," *Computers in biology and medicine,* vol. 43, no. 1, pp. 32-41, 2013.

[8]   X. Lladó, A. Oliver, J. Freixenet, R. Martí, and J. Martí, "A textural approach for mass false positive reduction in mammography," *Computerized Medical Imaging and Graphics,* vol. 33, no. 6, pp. 415-422, 2009.

[9]   S. Bagchi, K. G. Tay, A. Huong, and S. K. Debnath, "Image processing and machine learning techniques used in computer-aided detection system for mammogram screening - a review," *International Journal of Electrical and Computer Engineering (IJECE),* vol. 10, no. 3, pp. 2336-2348, 2020.

[10]   A. Pradeep and X. F. Joseph, "Binary operation based hard exudate detection and fuzzy based classification in diabetic retinal fundus images for real time diagnosis applications," *International Journal of Electrical and Computer Engineering (IJECE),* vol. 10, no. 3, pp. 2305-2312, 2020.

[11]   C.-H. Wei, Y. Li, and C.-T. Li, "Effective extraction of Gabor features for adaptive mammogram retrieval," in *2007 IEEE International Conference on Multimedia and Expo*, pp. 1503-1506, 2007.

[12]   S. Lahmiri and M. Boukadoum, "Hybrid discrete wavelet transform and Gabor filter banks processing for mammogram features extraction," in *2011 IEEE 9th International New Circuits and systems conference*, pp. 53-56, 2011.

[13]   M. Hussain, S. Khan, G. Muhammad, M. Berbar, and G. Bebis, "Mass detection in digital mammograms using Gabor filter bank," *IET Conference on Image Processing (IPR 2012),* 2012.

[14]   S. Khan, M. Hussain, H. Aboalsamh, and G. Bebis, "A comparison of different Gabor feature extraction approaches for mass classification in mammography," *Multimedia Tools and Applications,* vol. 76, no. 1, pp. 33-57, 2017.

[15]   Y. Zheng, "Breast cancer detection with gabor features from digital mammograms," *Algorithms,* vol. 3, no. 1, pp. 44-62, 2010.

[16]   N. Ponraj and M. Mercy, "Texture analysis of mammogram for the detection of breast cancer using LBP and LGP: A Comparison," in *2016 Eighth International Conference on Advanced Computing (ICoAC)*, pp. 182-185, 2017.

[17]   S. Khan, M. Hussain, H. Aboalsamh, H. Mathkour, G. Bebis, and M. Zakariah, "Optimized Gabor features for mass classification in mammography," *Applied Soft Computing,* vol. 44, pp. 267-280, 2016.

[18]   Z. Lei, S. Z. Li, R. Chu, and X. Zhu, "Face recognition with local gabor textons," in *International Conference on Biometrics*, pp. 49-57, 2007.

[19]   M. Hussain, S. Khan, G. Muhammad, I. Ahmad, and G. Bebis, "Effective extraction of Gabor features for false positive reduction and mass classification in mammography," *Appl. Math,* vol. 8, no. 1L, pp. 397-412, 2014.

[20]   A. Kamra, V. Jain, and S. Singh, "Extraction of orientation field using Gabor Filter and Gradient based approach for the detection of subtle signs in mammograms," *Journal of Medical Imaging and Health Informatics,* vol. 4, no. 3, pp. 374-381, 2014.

[21]   M. Y. Kamil and A. M. Salih, "Mammography Images Segmentation via Fuzzy C-mean and K-mean," *International Journal of Intelligent Engineering and Systems,* vol. 12, no. 1, pp. 22-29, 2019.

[22] K. C. Tatikonda, C. M. Bhuma, and S. K. Samayamantula, "The Analysis of Digital Mammograms Using HOG and GLCM Features," in *2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, pp. 1-7, 2018.

[23] F. S. S. de Oliveira, A. O. de Carvalho Filho, A. C. Silva, A. C. de Paiva, and M. Gattass, "Classification of breast regions as mass and non-mass based on digital mammograms using taxonomic indexes and SVM," *Computers in biology and medicine,* vol. 57, pp. 42-53, 2015.

[24] A. AlQoud and M. A. Jaffar, "Hybrid gabor based local binary patterns texture features for classification of breast mammograms," *International Journal of Computer Science and Network Security,* vol. 16, no. 4, pp. 16, 2016.

[25] F. Shirazi and E. Rashedi, "Detection of cancer tumors in mammography images using support vector machine and mixed gravitational search algorithm," in *2016 1st conference on swarm intelligence and evolutionary computation (CSIEC)*, pp. 98-101, 2016.

BIOGRAPHY OF AUTHOR

Mohammed Y. Kamil obtained his M. Sc. in Optics from Mustansiriya University, College of Science, Physics department, Iraq, in 2005. He received his Ph.D. in digital image processing from Mustansiriya University in 2011. He is interested in medical image processing, computer vision, and Artificial Intelligence. He is a member of the IEEE Iraq section.