International Journal of Electrical and Computer Engineering (IJECE)
Vol. 12, No. 3, June 2022, pp. 2564~2570
ISSN: 2088-8708, DOI: 10.11591/ijece.v12i3.pp2564-2570

Intelligent computer aided diagnosis system to enhance mass lesions in digitized mammogram images

Ayman AbuBaker[1], Yazeed Yasin Ghadi[2], Nader Santarisi[3]
[1]Department of Electrical Engineering, Faculty of Engineering and Technology, Applied Science Private University, Amman, Jordan
[2]Department of Software Engineering, Faculty of Information Technology, Al Ain University, Al Ain, Uni Emirat Arab
[3]Department of Industrial Engineering, Faculty of Engineering and Technology, Applied Science Private University, Amman, Jordan

Article Info

Article history:

Received Mar 21, 2021
Revised Dec 24, 2021
Accepted Jan 5, 2022

Keywords:

Enhancing mass lesions
Fuzzy logic
Mammograms
Mass Lesions
Neuro-fuzzy

ABSTRACT

The paper presents an intelligent system to enhance mass lesions in digitized mammogram images. This system can assist radiologists in detecting mass lesions in mammogram images as an early diagnosis of breast cancer. In this paper, the early detection of mass lesion is visually detected by enhancing mass lesions in mammogram images using hybrid neuro-fuzzy technique. Fuzzified engine is proposed as a first step to convert all pixels in mammogram image to a fuzzy value using three linguistic labels. After that, artificial neural networks are used instead of the inference engine to accurately detect the mass lesions in the mammogram images in a short time. Finally, five linguistic labels are used as a defuzzifier engine to restore the mammogram image. Processed mammogram images are extensively evaluated using two different types of mammogram resources, mammographic image analysis society (MIAS) and University of South Florida (USF) databases. The results show that the proposed intelligent computer aided diagnosis system can successfully enhance the mass lesions in mammogram images with minimum number of false positive regions.

Corresponding Author:

Ayman AbuBaker
Department of Electrical Engineering, Applied Science Private University
Shafa Badran, Amman, Jordan
Email: a_abubaker@asu.edu.jo

1. INTRODUCTION

Breast cancer is the at most early death in women [1]. Now a days, the most significant tool in order to detect and monitor breast cancer is using conventional mammography [2]. Microcalcification and mass lesions are the most significant indications of malignancy using mammography. Microcalcifications are defined as a little fleck of calcium, normally smaller than 1 mm in size, this can be shown as bright white dots versus dark and black atmosphere on the mammogram. A cluster of microcalcifications is a sign of cancer. Otherwise, mass lesion cells are bigger than macrocalcifications, also it could consist of various kinds and signs [3], [4]. The size of mass lesion cells differs ranged from 1 mm to few cm, and the shape also varied from circumscribed to speculated [5].

Analyzing mammograms is not an easy task for radiologists. Their decision relays on their personal training, effort, and years of experience, and conditioned to selective criteria. Hence, good experts might have an interobserver differences ranged between 65 to 75% [6]. Computer aided diagnosis (CAD) systems might aid radiologists to interpret mammograms against overall detection and classification. Between 65% to 90% of the biopsies of tested damaged cells shown out to be benign, accordingly it is very significant to construct CADs which able to differentiate between benign and malignant damaged cells [7]. The add on CAD group and experts' experience could significantly enhance detection accuracy. The detection sensitivity without CAD is 80% and with CAD up to 90% [8]. Computer based evaluation of mammograms has been discussed and monitored in detection systems, abnormalities in the mammogram images. Vast number of author's methods are investigated to show the mass lesion in the mammogram pictures from the regarding to the complexity in detecting these regions [9], [10]. Vast investigation is a harder dilemma in comparison to microcalcifications cluster analysis system due to masses are joined to the ambient parenchymal tissue dense, especially those speculated lesions and they are normally around the by non-regular tissue image with similar specifications. Normally, there are three sizes of mass damaged cells: smaller size (3 to 15 mm), middle sizes (15 to 30 mm) and larger size (30 to 50 mm) that increment masses detection systems [11].

The overall purpose of this paper is to create an intelligent CAD system platform that is able to accurately and reliably detect abnormalities, especially mass lesions, in a mammogram image. This CAD system will integrate different novel techniques which are designed and implemented to reliably detect the abnormalities in mammogram images from different sources. The proposed CAD system consists of two main processing stages: an image pre-processing stage and enhancement stage. The pre-processing stage will remove all the artifacts in mammogram images then a new novel algorithm will be implemented to enhance the mass lesion in the mammogram image. The enhancement algorithm use hybrid neuro-fuzzy technique. So, all the pixels in the mammogram image are converted to fuzzy values using three linguistic labels. Neural network is used next instead of the inference engine to speed up the processing time. Finally, the mass lesions are enhanced after defuzzification stage. The paper organization is as follow: related work is presented in section 2. Methodology is described in section 3. Finally, evaluation and conclusions are presented in sections 4 and 5, respectively.

2. RELATED WORK

Many authors applied different techniques to detect or enhance mass lesion in mammogram images. One of these techniques is using contour extraction as implement in Nakagawa *et al.* [12] article. The central point of the mass lesion is detected, then a contour technique is implemented. The sensitivity on this technique was 81%. Kom *et al.* [13] used an adaptive thresholding filter in detecting the mass lesion. The sensitivity was 95.91% TP detection for the mass lesion with 2 FP clusters/image. Also, Qian *et al.* [14] used using Ipsilateral multi-view CAD system. The average sensitivity of the algorithm is 89.6% TP with 1 FP cluster/image. Moreover, Sun *et al.* [15] also used Ipsilateral multi-view to detect mass lesions with a sensitivity 90% for TP and 3 FP/image. Cheng *et al.* [16] used fuzzy neural network on detection mass lesions. The proposed method used fuzzy neural network layer in detecting the tumors. As a result, the algorithm achieved 92% TP with 1.33 FP cluster/image. Convolution neural network is implemented by Chung *et al.* [17] to detect mass lesions in mammogram images. The results show 89% TP detection in mass lesion. On other way, wavelet transform is used to detect and enhance mass lesion by Zheng and Chan [18]. The DWT success in detecting the mass lesion with 97.3% TP with 3.92 FP/image. Speculated lesion, linear structures and a central mass are detected by Zwiggelaar *et al.* [19]. They proposed a new algorithm using recursive median filtering to detect different type of tumors in mammogram images. The results show 80% TP with 0.23 FP/image. The pattern recognition methodology is also used by Arod´z *et al.* [20] to detect mass lesion in mammogram images. They implemented support vector machine as a classifier in detecting mass lesion. The algorithm show sensitivity with 90% TP with 10 % FP. Finally, Eltonsy *et al.* [21] proposed Texture analysis in detecting mass lesion. The sensitivity for their algorithm was 92% TP with 5.4 FP cluster/image.

3. RESEARCH METHOD

Mass lesion enhancement technique is design based on using neuro-fuzzy technique. The proposed method (NFET) is divided to four stages as shown in Figure 1.

3.1. Pre-processing stage

Most of mammogram images have many artifacts which will reduce the efficiency of the enhancement process. Therefore, the first stage in our technique is removing the artifact from the mammogram images as in [22], [23]. Quality control (QC) sampling theory was investigated to remove the artifacts from mammogram images. In QC sampling theory two parameters should be verified: upper threshold (UT) limit and lower threshold (LT) limit [24]. These limits are declared in this technique since the intensity level for any mass lesion in all mammogram images was in the boundaries from [50 to 230] grey levels. So, UT is set to fixed value 230 grey level, whereas the LT is set be dynamic based on the image topology. So, the image mean ($\mu 0$) and standard deviation ($\sigma 0$) were the parameters to calculate the LT value as in (1).

$$LT=\mu_0-\sigma_0 \tag{1}$$

All artifacts and low level boundary regions will be eliminated from the mammogram images after implementing the QC threshold levels (LT and UT). As a results, the breast region and mass lesion become more focused.

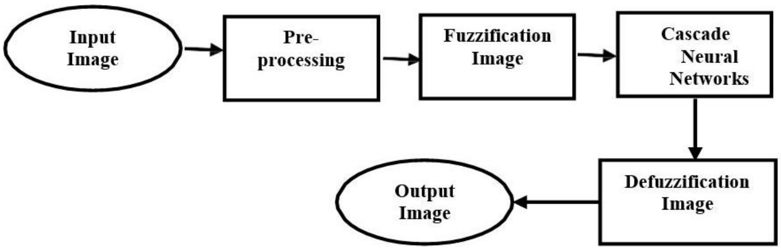


Figure 1. NFET flowchart

3.2. Fuzzification stage

Since the gray levels of boundaries of the mass lesion is very close to breast tissues grey level. The fuzzy logic technique is used to somehow distinguish between these values. In any fuzzy controller, the first step is fuzzification stage. So, intensity level of the pixels in mammogram images are processed to be converted to fuzzy value based on using three linguistics labels (low, medium, and high) as shown in Figure 2. Then the membership function for each intensity value is calculated based on (2) and (3).

$$M(x,a,b,c,d)=\begin{cases} 0, & x\le a \\ \frac{(x-a)}{b-a}, & a\le x\le b \\ 1, & b\le x\le c \\ \frac{(d-x)}{c-d}, & c\le x\le d \\ 0, & x\ge d \end{cases} \tag{2}$$

$$M(x,a,b,c)=\begin{cases} 0, & x\le a \\ \frac{(x-a)}{b-a}, & a\le x\le b \\ \frac{(c-x)}{c-b}, & b\le x\le c \\ 0, & x\ge c \end{cases} \tag{3}$$

The parameters a, b, c are changed dynamically based on UT and LT value for each mammogram image also steps (sampling) in fuzzification is calculated based on (4).

$$S=\frac{UT-LT}{4} \tag{4}$$

3.3. Neural network stage

To build fuzzy inference engine, a mask of size 5×5 is used to select number of inputs for the inference engine as shown in Figure 3. Mask center is a reference point, and the sixteen connected neighbor pixels are inputs to the inference engine. This process required 43,046,721 activation rules which make processing time very long. Therefore, artificial neural network (ANN) is used instead of the fuzzy inference engine to speed up the processing time. In any ANN, there should be three steps: collecting dataset, training process, and learning process which will presented in following sections in details.

3.3.1. Collecting dataset

In feedforward ANN, inputs and outputs must be declared as in [25]. The inputs for our ANN are the membership value for each 16 connected neighbor pixels. Since the intensity level boundaries for mass lesions is [50 to 230] grey level, number of samples for each vector is calculated based on (5). If we assume that LT=50 and UT=230 then number of samples will be 90 for each vector. As a result, the combination for all 16 connected neighbor pixels will be $(90)^{16}$ which equal to 1853020188851841000000000000000 rows with 48 columns.

$$Numbe\ of\ Samples\ for\ each\ input = \frac{UT-LT}{2} \tag{5}$$

Due to the huge dataset, the cascade neural network is used in training the dataset which will be explained in following section.

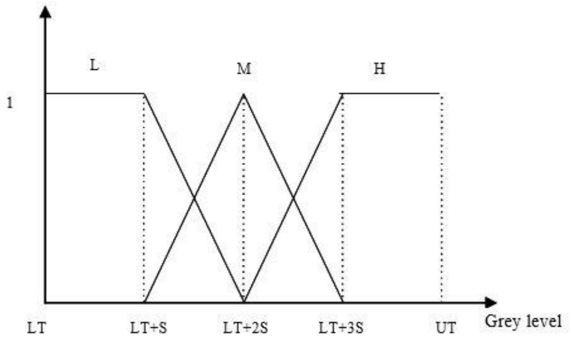


Figure 2. Fuzzification input image intensities

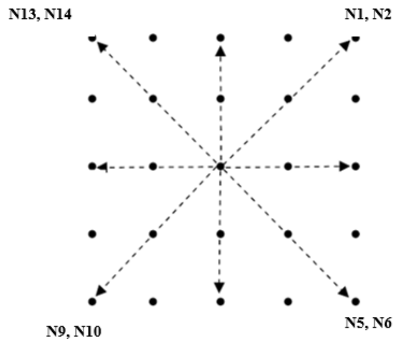


Figure 3. 16-connected neighbors of size 5×5

3.3.2. Training the neural network

Initially, neural network topology configuration is set based on input and output data. After many trials it was found that the best topology is in using cascade neural network. So, sixteen individual neural networks with structure (3 input and 5 outputs) are generated. Then the outputs of all previous neural networks are connected as an input to a decision neural network with structure of (48 inputs and 5 outputs) as shown in Figure 4.

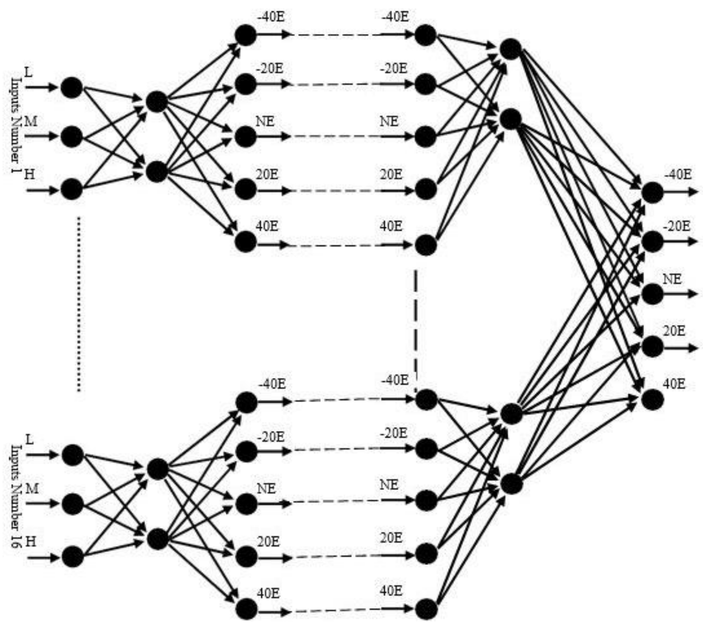


Figure 4. Cascade parallel neural network

3.3.3. Learning process

After generating the input and output dataset, the final step in ANN is the learning processing. Backpropagation feedforward learning procedure is used in our learning process for the cascade ANN. The input and output matrix are arranged based on Jakeknife method [26] where 70% of the data for training and 30% of the data for testing. Then different feedforward structures (changing number of hidden layer or number of hidden nodes or both) are carried out to find the minimal mean square error as shown in Figure 5.

As a result, it is found that the feedforward NN with one hidden layer and nine nodes for the parallel NN and feedforward NN with one hidden layer and seventeen nodes for the cascade NN produce the minimum errors.

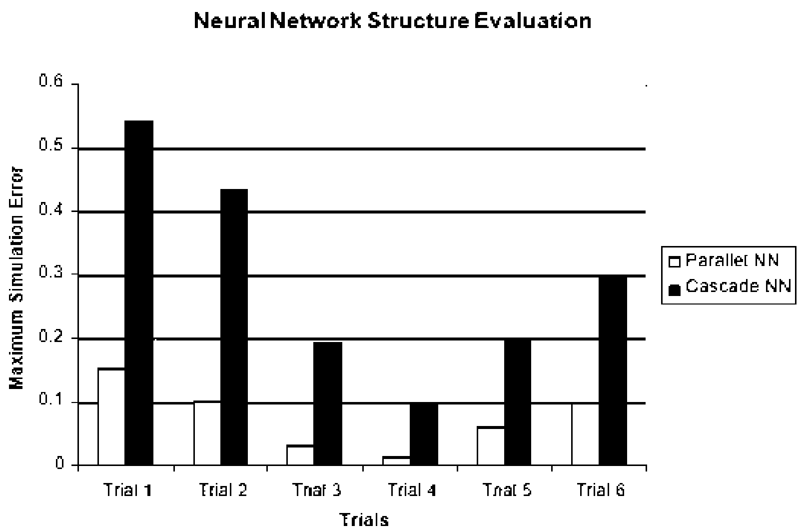


Figure 5. Evaluation of neural network structure

3.4. Defuzzification stage

The final stage in fuzzy controller is a defuzzification stage which will convert the fuzzy value to a crisp value. The outputs from cascade ANN are five logistics labels (-40E, -20E, NE, 20E, 40E) as presented in previous section. The membership value for these linguistics label will be processed in this defuzzification engine to restore the mammogram image. Center of gravity technique is used in defuzzification engine with universe of discourse (-50% to 50% of the intensity value of mask center) as shown in Figure 6.

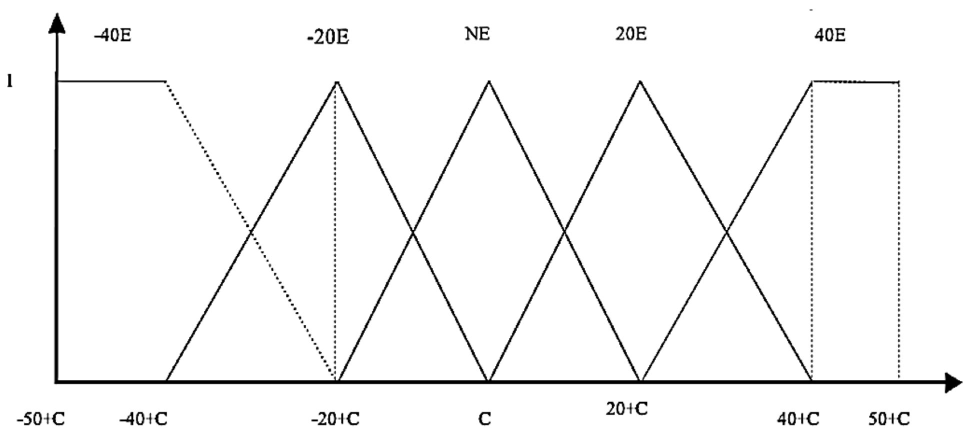
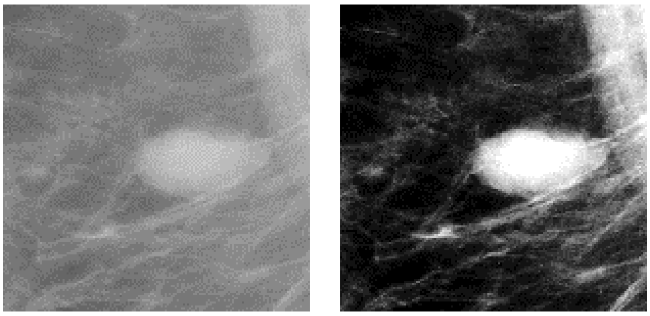


Figure 6. Defuzzification output sets

4. RESULTS AND DISCUSSION

Different resources of mammogram images are used to evaluate the performance of the proposed algorithm. Images are objectively compared with original image. As shown in Figures 7(a) and 7(b), intelligent CAD system success in improving brightness of mass lesion in mammogram images. Also, processing time of enhancement process is another challenge that is considered in this evaluation. In classical fuzzy inference engine with 43,046,721 activation rules, the processing time will be 3.4 sec. whereas, using neuro-fuzzy and specially relacing the ANN instead of inference engine enhance the processing time to be 102 μS comparing with 3.4 sec in traditional fuzzy controller.



(a) (b)

Figure 7. Intelligent CAD enhancement result (a) cropped image and (b) enhanced mass lesion

5. CONCLUSION

Intelligent technique to enhance mass lesions in digitized mammogram images is presented in this paper. This approach used the hybrid neuro-fuzzy technique to accurately enhanced the mass lesions with minimal number of false positive regions. This technique having many stages: in first stage, fuzzification engine is implemented to convert all the pixels in mammogram image to fuzzy values. Then cascade neural networks are used instead of the inference engine due to the huge number of activation fuzzy rules in the inference system. As a final stage, five linguistics labels which resulted from cascade ANN are processed in defuzzification engine to restore the mammogram image. Intelligent enhancement technique can accurately enhance only the mass lesion area with minimum number of FP regions. This will increase the sensitivity of the proposed algorithm in diagnosis like these cases.

ACKNOWLEDGEMENTS

The author is grateful to the Applied Science Private University, Amman, Jordan, for the full financial.

REFERENCES

[1] A. G. Haus and M. J. Yaffe, *Syllabus : a categorical course in physics : technical aspects of breast imaging*. Oak Brook, IL : RSNA Publications, 1994.

[2] K. Bovis, S. Singh, J. Fieldsend, and C. Pinder, "Identification of masses in digital mammograms with MLP and RBF nets," in *Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks. IJCNN 2000. Neural Computing: New Challenges and Perspectives for the New Millennium*, 2000, pp. 342--347 vol.1, doi: 10.1109/IJCNN.2000.857859.

[3] 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.

[4] A. M. Salih and M. Y. Kamil, "Mammography image segmentation based on fuzzy morphological operations," in *2018 1st Annual International Conference on Information and Sciences (AiCIS)*, Nov. 2018, pp. 40–44, doi: 10.1109/AiCIS.2018.00020.

[5] H. Li, Y. Wang, K. J. R. Liu, S.-C. B. Lo, and M. T. Freedman, "Computerized radiographic mass detection. I. Lesion site selection by morphological enhancement and contextual segmentation," *IEEE Transactions on Medical Imaging*, vol. 20, no. 4, pp. 289–301, Apr. 2001, doi: 10.1109/42.921478.

[6] P. Skaane, K. Engedal, and A. Skjennald, "Interobserver variation in the interpretation of breast imaging," *Acta Radiologica*, vol. 38, no. 4, pp. 497–502, Jul. 1997, doi: 10.1080/02841859709174375.

[7] A. Neekabadi, S. Samavi, N. Karimi, E. Nasr-Esfahani, S. A. Razavi, and S. Shirani, "Lossless compression of mammographic images by chronological sifting of prediction errors," in *2007 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing*, Aug. 2007, pp. 58–61, doi: 10.1109/PACRIM.2007.4313176.

[8] K. Doi, "Computer-aided diagnosis: potential usefulness in diagnostic radiology and telemedicine," in *Proceedings of the National Forum: Military Telemedicine On-Line Today Research, Practice, and Opportunities*, 1995, pp. 9–13, doi: 10.1109/MTOL.1995.504521.

[9] L. W. Bassett, D. H. Bunnell, R. Jahanshahi, R. H. Gold, R. D. Arndt, and J. Linsman, "Breast cancer detection: one versus two views.," *Radiology*, vol. 165, no. 1, pp. 95–97, Oct. 1987, doi: 10.1148/radiology.165.1.3628795.

[10] J. Wei *et al.*, "Computer-aided detection of breast masses on full field digital mammograms," *Medical Physics*, vol. 32, no. 9, pp. 2827–2838, Aug. 2005, doi: 10.1118/1.1997327.

[11] H. Li, Y. Wang, K. J. R. Liu, S.-C. B. Lo, and M. T. Freedman, "Computerized radiographic mass detection. II. Decision support by featured database visualization and modular neural networks," *IEEE Transactions on Medical Imaging*, vol. 20, no. 4, pp. 302–313, Apr. 2001, doi: 10.1109/42.921479.

[12] T. Nakagawa, T. Hara, H. Fujita, T. Iwase, T. Endo, and K. Horita, "Automated contour extraction of mammographic mass shadow using an improved active contour model," *International Congress Series*, vol. 1268, pp. 882–885, Jun. 2004, doi: 10.1016/j.ics.2004.03.172.

[13] G. Kom, A. Tiedeu, and M. Kom, "Automated detection of masses in mammograms by local adaptive thresholding," *Computers in Biology and Medicine*, vol. 37, no. 1, pp. 37–48, Jan. 2007, doi: 10.1016/j.compbiomed.2005.12.004.

[14] W. Qian, D. Song, M. Lei, R. Sankar, and E. Eikman, "Computer-aided mass detection based on ipsilateral multiview

mammograms," *Academic Radiology*, vol. 14, no. 5, pp. 530–538, May 2007, doi: 10.1016/j.acra.2007.01.012.
[15] X. Sun, W. Qian, and D. Song, "Ipsilateral-mammogram computer-aided detection of breast cancer," *Computerized Medical Imaging and Graphics*, vol. 28, no. 3, pp. 151–158, Apr. 2004, doi: 10.1016/j.compmedimag.2003.11.004.
[16] H. D. Cheng and M. Cui, "Mass lesion detection with a fuzzy neural network," *Pattern Recognition*, vol. 37, no. 6, pp. 1189–1200, Jun. 2004, doi: 10.1016/j.patcog.2003.11.002.
[17] S.-C. B Lo, H. Li, Y. Wang, L. Kinnard, and M. T. Freedman, "A multiple circular path convolution neural network system for detection of mammographic masses," *IEEE Transactions on Medical Imaging*, vol. 21, no. 2, pp. 150–158, 2002, doi: 10.1109/42.993133.
[18] L. Zhen and A. K. Chan, "An artificial intelligent algorithm for tumor detection in screening mammogram," *IEEE Transactions on Medical Imaging*, vol. 20, no. 7, pp. 559–567, Jul. 2001, doi: 10.1109/42.932741.
[19] R. Zwiggelaar *et al.*, "Model-based detection of spiculated lesions in mammograms," *Medical Image Analysis*, vol. 3, no. 1, pp. 39–62, Mar. 1999, doi: 10.1016/S1361-8415(99)80016-4.
[20] T. Arodź, M. Kurdziel, E. O. D. Sevre, and D. A. Yuen, "Pattern recognition techniques for automatic detection of suspicious-looking anomalies in mammograms," *Computer Methods and Programs in Biomedicine*, vol. 79, no. 2, pp. 135–149, Aug. 2005, doi: 10.1016/j.cmpb.2005.03.009.
[21] N. H. Eltonsy, G. D. Tourassi, and A. S. Elmaghraby, "A concentric morphology model for the detection of masses in mammography," *IEEE Transactions on Medical Imaging*, vol. 26, no. 6, pp. 880–889, Jun. 2007, doi: 10.1109/TMI.2007.895460.
[22] L. Bocchi, G. Coppini, J. Nori, and G. Valli, "Detection of single and clustered microcalcifications in mammograms using fractals models and neural networks," *Medical Engineering & Physics*, vol. 26, no. 4, pp. 303–312, May 2004, doi: 10.1016/j.medengphy.2003.11.009.
[23] M. Azzeh, D. Neagu, and P. I. Cowling, "Fuzzy grey relational analysis for software effort estimation," *Empirical Software Engineering*, vol. 15, no. 1, pp. 60–90, Feb. 2010, doi: 10.1007/s10664-009-9113-0.
[24] A. B. Nassif, M. Azzeh, A. Idri, and A. Abran, "Software development effort estimation using regression fuzzy models," *Computational Intelligence and Neuroscience*, vol. 2019, pp. 1–17, Feb. 2019, doi: 10.1155/2019/8367214.
[25] A. B. Nassif, M. Azzeh, L. F. Capretz, and D. Ho, "Neural network models for software development effort estimation: a comparative study," *Neural Computing and Applications*, vol. 27, no. 8, pp. 2369–2381, Nov. 2016, doi: 10.1007/s00521-015-2127-1.
[26] K. Fukanga, *Introduction to Statistical Pattern Recognition*. Elsevier, 1990.

BIOGRAPHIES OF AUTHORS

Ayman AbuBaker has obtained his PhD 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 has obtained his PhD in software Engineering, Qeens, Queensland University, Australia Artificial intelligence, Robotics, Computer Vision, System and Software Modeling and Analysis, System and Software Verification, Validation and Testing, Big Data, Indoor and Outdoor Smart Environments, Computational Intelligence, Machine learning. He can be contacted at email: Yazeed.ghadi@aau.ac.ae.

Nader Santarisi had his PhD in Industrial Engineering from the University of Iowa, USA, in 1993. He had started his academic career at the University of Jordan in 1993. His research interests include operations management, quality management, neural networks, machine learning, artificial intelligence, and big data analytics. In 2006, he was appointed as the Head of the Board of Examiners of King Abdullah II Award for Excellence. During the period from July 2009-July 2020, he was working as an Expert (CEO Advisor) with Dubai Government. Nowadays, he is an Associate Professor in Industrial Engineering Department at the Applied Science Private University, Amman, JORDAN. He can be contacted at email: n_santarisi@asu.edu.jo.