A Novel Deep Learning Based Approach for Breast Cancer Detection

Muhammad Aaqib, Muhammad Tufail and Shahzad Anwar*
Department of Mechatronics Engineering, University of Engineering and Technology Peshawar, Pakistan
*Corresponding author: shahzad.anwar@uetpeshawar.edu.pk

Abstract— **This study examine and investigate possibility of employing machine vision techniques for Breast cancer identification. Initially, data preparation was performed for labelling followed by preprocessing for removing pectoral muscle and inherent noise. Digital Database for Screening Mammography (DDSM) an internationally available well received database, which contain 2620 cases of mammograms was incorporated for experimentation. VGG-16 was employed for feature extraction subsequently, SSD was incorporated for tumor detection. Once method was established results were compared with other published methods for validation. The results exhibit an accuracy of 96.2% on DDSM. Therefore, a method is proposed and developed and its effectiveness is demonstrated in term of breast cancer detection.**

Keywords—Breast Cancer, Deep Convolutional Neural Network, Machine Learning, Image Processing.

I. INTRODUCTION

In the recent era cancer has been one of the most liable reason for high number of deaths and is anticipated to be one of the main causes for most deaths in the next decades. Cancer is challenges in terms of deaths and morbidity [1]. Globally, breast cancer (BC) has been diagnosed as a considerable universal health issue and the main reason of cancer associated deaths among women, especially in less resourced territory of the world [2]. In 2018, about 2.1 million women were estimated to have BC and 626,679 deaths related to breast cancer was stated. So, an increase of about 16% and 18% incidence and rate of mortality between the era of 2013 to 2018 [3]. According to World Health Organization, the expected number of cancer patients will be 19.3 million in 2025 [4]. In Asia, 1 in 9 women have developed breast cancer disease at some stages of life. The causes of breast cancer related deaths can be minimized by the subsequent actual diagnosis techniques at the earliest stages before considerable symptoms are developing in the body. Therefore, numerous approaches have been employed for breast cancer detection such as Support Vector Machine, Artificial Neural Networks and Convolutional Neural Networks, [5] [6][7]. Hence, mammography is an effective imaging approach for the early stage breast cancer identification [8] [9]. Indeed, mammography offers to detects minute level irregularities in human body, however, the precision and accuracy of detection is highly dependent on quality of image (QoI) and radiologist expertise. A study presented in [10] [11] suggest that error rate during examination among radiologist is recorded approximately 35%. In order to minimize errors and false negative detection, a patient with only 2% malignancy chances are prescribed for a biopsy, and 15% - 35% biopsy cases are observed to be malignant [12]. Therefore, suspicious biopsy cases are an expensive solution and cause anxiety and nervousness to the patients. Computer aided diagnosis (CAD) a framework was incorporated to aid radiologist in decision making process during clinical practices. Such a system decreases the amount of efforts require for the evaluation of a lesion and also minimize the number of false positive rates that guide for the biopsies which are unnecessary and cause discomfort for the patient. CAD frameworks offers two steps procedure; (i) suspicious lesion detection in mammogram and (ii) lesion diagnosis classification as malignant or benign respectively.

The remaining of this article is organized as following. Section 2 presents related work followed by methodology in section 3. Experimentation are presented in section 4. Section 5 provide the results of this paper and finally section 6 conclude this manuscript.

II. RELATED WORKS

Like the framework of Wang et al, [13] classification was evaluated as the forecasting technique to build a system model. In mass classification, Artificial neural network (ANN) and Linear discriminant analysis (LDA) are settled as fined and well classifier. For features selection, first the ANN was trained and then outputs for the known database is used to resolve the weights. To get the known output like that it is increased or decreased should be adjusted the weights properly and then for masses classification the ANN is integrated after training. Salem et al, [14] proposed a network for BC tumors classification whichever malignant or benign. The base of this classification technique is the combination of ontology and case-based reasoning (CBR). According to Jiang and his team [15] schemed a technique for detection using mammograms for classifying the suspicious regions as malignant or benign. A survey has been carried out exhibits how the frameworks of CAD aid in cancer detection. Xie et al, [16] propose CAD technique to

978-1-7281-4956-1/19/$31.00 ©2019 IEEE

identify breast cancer as a consequence of extreme learning machine (ELM), that concern with the segmentation of the mammography through eliminate the inference and improved the quality of image and then extract region of interest based on collected features. By the combination of support vector machine and extreme learning machine selection of feature is obtained. So, this justify that the presented CAD method providing impressive results and also reduce the training time. Chan et al [17] described a breast mass density that employs a topographic map for significant rate. In their methodology, an associated component shows a frame which is assembled as a tree that represents topological framework. The salient features are used for detection of dense areas in the breast. The datasets used are MIAS and DDSM and achieve better accuracies i.e. 84% and 88% respectively. Yamazaki et al [18] proposed a system that mammogram is segmented in to three different regions employing linear discriminant analysis and histogram technique to classify the fibro glandular tissue density. The data used are 150 Japanese mammogram images which is prepared by a radiologist. The accuracy of this system is 90% which is grounded on four different classes. Since the development of machine and deep learning, much research has been presented in [19] employing deep learning architectures and the most familiar form of deep learning architecture is convolution neural network. According to Arevalo [20] various CNN algorithm was tested and distinguished them by labeling of two hand crafted for mass task detection. The dataset used in this experiment on the (BCDR-DM) and (BCDR-FM). Such technique presented the enhanced performance with the sequence of learned and hand-crafted illustration. But the author did not check the accomplishment of pre-trained system and utilized simple CNN models. Giger et al [21] used Alex-Net [22] pre trained for the problem of mass diagnosis with no fine tuning. Giger and his team examined the classification performance employ features from different layers of the network with the aid of support vector machine (SVM) and their results are compared with two different ways, a classifier employs on hand crafted features and an ensemble of soft voting classifier. Jiao and Gao [23] presented a framework on a sub sample of DDSM database which are pre trained calibration of CNN architecture. They used different layers of models to remove features of masses and then compare various scales to obtained features of high as well as middle level. Two classifiers of support vector machine are trained for the technique of decision i.e., one for feature class and other for their fused predictions. Rampun and Scotney [24] shows a collaborative of a small version of Alex net pre trained and fined tuned on CBIS DDSM. At the time of conclusion, they collect the best three performing model and fused their prognosis. Ting et al [25] design and prepared from scratch their system for the classification of breast masses. The database used for this experiment was mini MIAS. The framework utilizes 28 convolutional layers and fully connected layers and then through detection of ROI, it is fed by one shot detector. Kulsoom et al [26] proposed an effective technique for classification of malignant and benign breast cancer. This technique employs an optimal feature classification using artificial neural networks. Their system was trained, tested and validated on various datasets. According to receiver operating characteristics the accuracy rate for the system was 92%. On the basis of the aforementioned literature the proposed research aims to develop a novel framework using deep learning for classifying benign and malignant lesions from mammography. The proposed research overcomes the shortcomings explained in the above literature that includes higher accuracy rate and less computational power compared to the already used techniques.

III. MATERIAL AND METHOD

The main aim of this study is the identification of breast cancer at an initial stage and subsequently abnormalities classification in digital mammogram. Here, mammograms used as inputs for convolution neural networks for labeling the unknown. The following fig. 1 presents the overall process of proposed method.

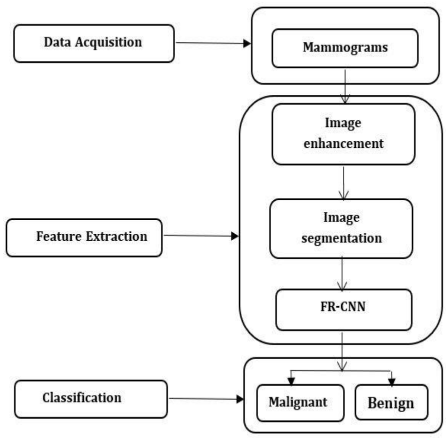

Fig.1 Block Diagram of the system architecture

In the study of mammogram classification, many researchers used their own databases. However, due to data protection act, most databases for mammograms are not public. In our study, a bench- mark database, Digital database for screening mammography (DDSM) and MIAS databases. The database comprises 2620 cases with labels marked by radiologists. Each image has the information of its categories (benign or malignant), the coordinates of the tumor center and the radius of the tumor. In the proposed method total of 1000 images are used in which half is malignant and half is benign for training and testing purposes. In pre-processing step details of image is taped by sensors having errors related to geometry and brightness values of the pixels. These errors are corrected by using appropriate mathematical models. Image enhancement is the improvement of image which is changing the pixel brightness values to modify its visual effect. In this a collection of frameworks that are used to enhanced the visual appearance of an image or to convert the image to a form which is better suited for human or machine interpretation. Deep neural network has made a tremendous success in image processing area including image classification, object detection etc. The performance of neural network surpasses traditional mathematical approaches in image processing. In this study, the techniques applied for lesion detection is SSD and for mammogram classification is VGG 16 respectively. In Network Architecture, a binary classifier was built with VGG16 and Convolutional Neural Network Enhancement for Breast Cancer Classification (CNNE-BCC) is develop as shown in the fig. 2.

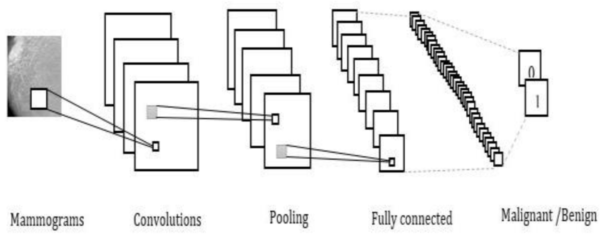

Fig. 2 CNNE-BCC Architecture

CNNE-BCC is 30-layer convolutional neural networks specialized at learning features in mammograms. It is modified based on VGG16. It has 28 convolutional layers, one pooling layer and one fully connected layer. Instead of using normal 2-dimensional (2D) convolutional layer, depth wise separable convolution layer is adopted to minimize the number of parameters required to train average pooling operation is used in the pooling layer and a dense layer is appended at the end. VGG16 as shown in figure 3 is a 16-layer convolutional neural network which is established its worth in image classification and been widely rewarded [27] having five blocks. First two blocks having two convolutional layers followed by a stack of convolutional layers and the subsequent three blocks consist of three convolutional layers followed by a max-pooling layer. A rectified linear unit (ReLU) is appended as activation function at the end of each convolution layer. Two fully connected layers are concatenated with the block of convolutional for classification.

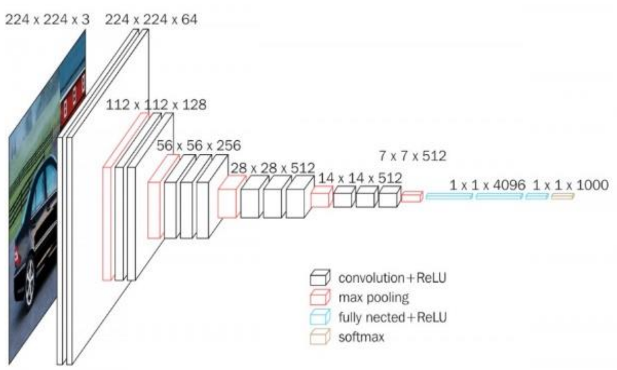

Fig.3 VGG 16 Architecture [27]

IV. EXPERIMENTATION

The main purpose of CNN is to determine where the object is and what the object is in an image. CNN works in layers in which the main function of feature extraction and classification occurs. The architectural layers are as following: (i) convolutional layers (ii) ReLU (iii) Pooling layers and (iv) fully connected layers. Convolution layer is the core block of CNN, take a multiple filter called kernel filters; they have the same depth as of the input. Kernel filters are applied on the input data generates the outputs, which would have the same spatial dimension as inputs and depth of the output depends on the number of kernel filters. Spatial dimensions of the output depend on the stride, if the stride value is one then the measurement of output is equal to its input and will process every pixel. If stride is two the dimensions changes and every second pixel is processed. The output dimension *'w /2 * h/2* D'* where '*w*' is width of input image, h is height of input image and '*D*' is the depth of an input image. In convolution layers, the basic process is to extract the features through different kernel filters, like one filter for edge detection, one for corner, others for color of an images etc. Further, Rectifier Linear Unit (ReLU) was introduced to bring non-linearity in network. Subsequent to ReLU, next is pooling layers (nonlinear layer) used for down sampling. Pooling take input from the previous layers i.e. the output of that previous layer having the same size. However, the output of the pooling layer depends on window size. Fully Connected Layer (FCC) was next to find the final output. In FCC each pixel is connected output class, thus making it heavy data layer. After processing, Softmax or Support Vector Machines (SVM) employed to generates outputs. Single Shot Multibox Detector (SSD) is one of the well-known object detection algorithms as shown in figure 4.

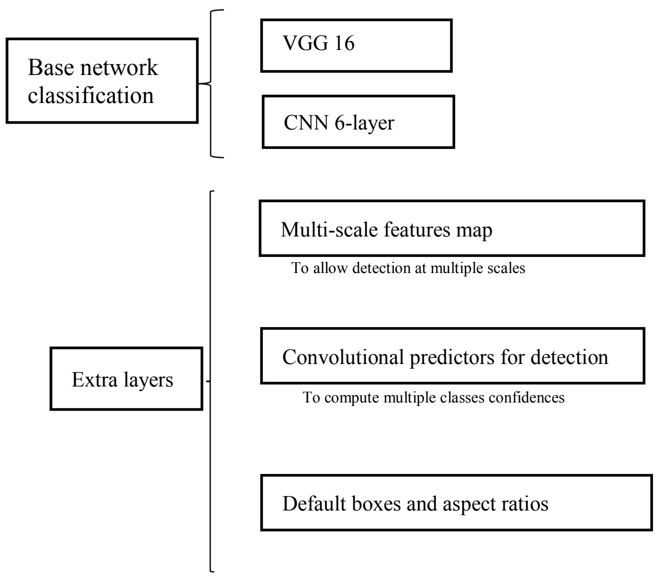

Fig. 4 SSD Architecture

Due to high accuracy; SSD was chosen for experimentation for BC detection. In SSD model, number of convolutional layers and max pooling layers are used to produce feature maps of different scales to detect objects of various sizes. In the source code, the SSD models suitable for 300x300 and 512x512 pixel size images. Once test was performed, layers were amended to fit the 256x256 size image. The next section of SSD model, as shown in figure 4, is for multiple classes confidence and learn the localization detection computation, which remain unchanged. Table 1 is SSD model architecture, illustrating the modified layers employed in the proposed experimentation.

Table 1. SSD system based proposed model

S. No	Layer Types	Filter	Size	Stride	O/P
1	Conv2D	64	(3,3)	(1,1)	(128,128)
	Conv2D	64	(3,3	(1,1)	
	Maxpooling2D	--	(2,2)	(2,2)	
2	Conv2D	128	(3,3)	(1,1)	(64,64)
	Conv2D	128	(3,3)	(1,1)	
	Maxpooling2D	--	(2,2)	(2,2)	
3	Conv2D	256	(3,3)	(1,1)	(32,32)
	Conv2D	256	(3,3)	(1,1)	
	Conv2D	256	(3,3)	(1,1)	
	Maxpooling2D	--	(2,2)	(2,2)	
4	Conv2D	512	(3,3)	(1,1)	(16,16)
	Conv2D	512	(3,3)	(1,1)	
	Conv2D	512	(3,3)	(1,1)	
	Maxpooling2D	--	(2,2)	(2,2)	
5	Conv2D	512	(3,3)	(1,1)	--
	Conv2D	512	(3,3)	(1,1)	
	Conv2D	512	(3,3)	(1,1)	
	Maxpooling2D	--	(2,2)	(2,2)	
6	Conv2D	1024	(3,3)	(1,1)	--
	Conv2D	1024	(1,1)	(1,1)	
7	Conv2D	256	(1,1)	(1,1)	(8,8)
	Conv2D	512	(3,3)	(2,2)	
8	Conv2D	128	(1,1)	(1,1)	(4,4)
	Conv2D	256	(1,1)	(2,2)	
9	Conv2D	128	(3,3)	(1,1)	(2,2)
	Conv2D	256	(1,1)	(1,1)	
10	Conv2D	128	(2,2)	(1,1)	(1,1)
	Conv2D	256	--	(1,1)	

V. RESULTS

A series of experimentations were performed for mammogram classification with VGG16 and CNNE-BCC, respectively. In order to exhibits the generalization, this study was implemented employing python with anaconda based on TensorFlow backend. Data was randomly distributed for training and testing. At training stage, 70% of total data was allocated, whilst the remaining 30% were reserved for testing. The data reserved for testing (i.e. 30%) was kept hidden at training stage. As presented in fig. 5, the learning rate error is approaching to zero with increasing number of epochs and at $20x10^3$, 0.012 loss in accuracy was achieved.

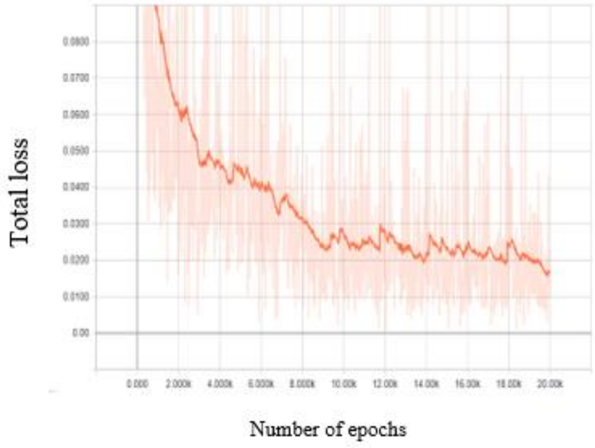

Fig. 5 Learning rate error

To test the capabilities of the develop method, malignant and benign lesion mages as shown in figure 6 (a) & (b). It is evident that 96% malignancy is detected as shown in figure 6 (c), while 95% benign is diagnosed as shown in figure 6 (d).

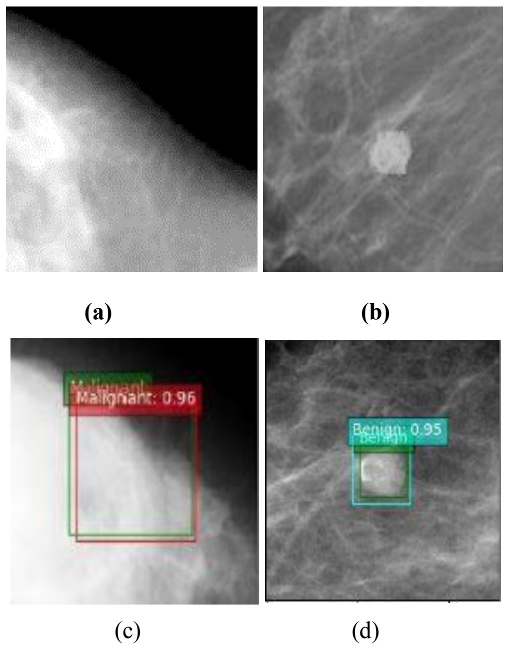

Fig. 6. (a) Original image of malignant (b) Original image of benign (c) Classification of malignant (d) Classification of benign

A) A comparison for evaluation

ROC study is used as a methodical form for measuring the cause of predictability between special thresholds resolutions. For classification purposes ROC curve is used as shown in the fig.7. The false positive rate (FPR) is shown on x-axis while true positive rate (TPR) is shown on y-axis respectively.

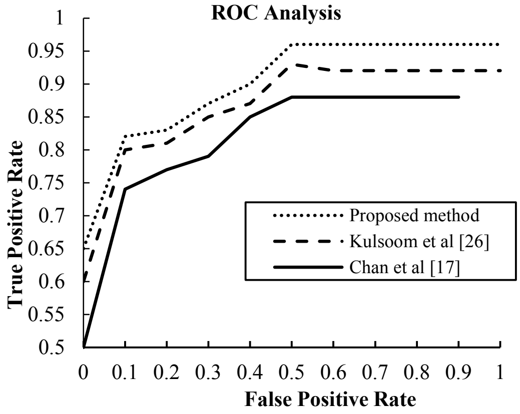

Fig. 7 ROC Analysis

In order to justify the effectiveness, a comparative analysis of the proposed deep learning framework (CNN, VGG 16) was performed with already established method such as Kulsoom et al. [26] and Chan et al [17]. It is evident from Fig. 8 that the total accuracy of the method is 96.28, and accuracy for others are 93 % and 88 % respectively. The results presented for developed method indicate better performance thus, the effectiveness of the proposed method more efficient and accurate.

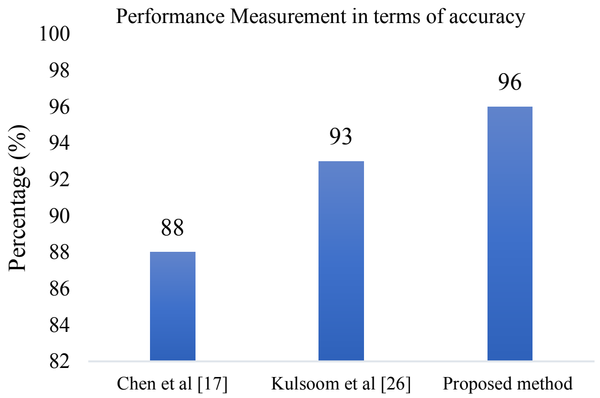

Fig. 8 Comparison of performance of the three techniques with previous techniques using DDSM database

Confusion matrix is measurement performance for machine learning classification where output can be two or more classes. It is the combination of predicted and actual values. This shows whether the system is classifying the classes correctly or it confuses the classes during classification. To check the validation results further experiments were carried out which is summarized in table 2

Table: 2 Confusion Matrix

	True positive	False negative
Predicted positive values	2143	76
Predicted negative values	16	1227

Further computations were performed in order to obtain crucial values from confusion matrix, such as, sensitivity, specificity, positive prediction values and negative prediction as shown in table.3.

Table: 3 Values obtained from confusion matrix

Sensitivity	TP/(TP+FN)	0.9881
Specificity	TN/(FP+TN)	0.9682
Positive prediction values	TP/(TP+FP)	0.9563
Negative prediction values	TN/(TN+FN)	0.9640
False positive rate	Sensitivity/(1-Specificity)	0.0418
False negative rate	(1-Sensitivity)/Specificity	0.0229

Globally, cancer is a major public health disease causing millions of deaths and in women breast cancer (BC) is a second most detected malignancy. One in eight women develop BC in their life. Early detection and accurate diagnoses of the disease could improve cancer survival and minimize cost of medications. In recent years Computer-aided diagnosis (CAD) systems have been employed to aid radiologist for digital mammographic inspection. Mammograms Classification is a complicated and challenging task as the tumor is usually located in a small region of interest of the entire mammography. Hence, the ternary label of mammograms may not be an obvious feature for the neural network to learn. Furthermore, the mammograms usually have low contrast and sometimes it's even hard for radiologists to interpret. Applying appropriate techniques for enhancing the image contrast before training could be a plausible approach to improve the performance of neural network. In addition, VGG16 in its nature is designed for classifying images into 1000 categories, which leads to the great number of parameters in the last two dense layers. However, in our study the mammograms only need to be classified into 2 categories, which implies a possibility of applying some hyperparameter tuning algorithm to find a better hyperparameter set that could reduce the computing time without sacrificing the performance. Besides, it is also possible to employ pre-trained neural network with large dataset such as ImageNet to extract rudimentary features of the mammograms. The proposed system is capable and having the potential to be incorporated in an unknown circumstance in which sufficient data is not provided for examination.

VI. CONCLUSION AND FUTURE WORK

In this study, system is trained and evaluated on convolutional neural networks for mammograms classification and tumor detection. Digital Database for Screening Mammography (DDSM) which contain 2620 cases of mammograms was incorporated for experimentation. Initially, pre-processing was applied to adjust mammographic image size in order to have efficient calculation. Deep Convolutional Neural Networks (CNN-s) were employed on mammographic images. Various features are extracted for training data at Rectified- Linear-Unit (RELU), Pooling layers and Fully connected layers and subsequently data fed to CNN for decision. The results exhibit an accuracy of 96.2% on DDSM a publicly available international dataset.

In future it is interesting to observe how the proposed method performs in noisy mammographic images and to observe the effect upon the accuracy of performance by increasing the number of cases.

REFERENCES

[1] DeSantis, C., Ma, J., Bryan, L., & Jemal, A. Breast cancer statistics. A Cancer Journal for Clinicians, 64(1), pp.52–62,2013.

[2] T. Matsuda and M. Niino, “Bladder cancer incidence rates in the world from the Cancer Incidence in Five Continents XI,” Jpn. J. Clin. Oncol., vol. 48, no. 10, pp. 953–954, 2018.

[3] J. Greenhalgh *et al.*, “Eribulin for the Treatment of Advanced or Metastatic Breast Cancer: A NICE Single Technology Appraisal,” Pharmacoeconomics, vol. 33, no. 2, pp. 137–148, 2014.

[4] D. A. Ragab, M. Sharkas, S. Marshall, and J. Ren, “Breast cancer detection using deep convolutional neural networks and support vector machines,” Peer J, vol. 7, pp. 1-23, 2019.

[5] Jalalian, A, Mashohor, S., Mahmud, R., Karasfi, B., Saripan, B. and Ramli, A., Foundation and methodologies in computer-aided diagnosis systems for breast cancer detection. *EXCLI journal*, *16*, pp.113-137,2017.

[6] Gerasimova-Chechkina, E., Toner, B., Marin, Z., Audit, B., Roux, S.G., Argoul, F., Khalil, A., Gileva, O., Naimark, O. and Arneodo A. Combining multifractal analyses of digital mammograms and infrared thermograms to assist in early breast cancer diagnosis. In *AIP conference proceedings* Vol. 1760, No. 1, pp.1-11, 2016.

[7] V. Vijayasarveswari, S. Khatun, M. M. Fakir, M. Jusoh, and S. Ali, March “UWB Based Low-Cost and Non-Invasive Practical Breast Cancer Early Detection,” Vol. 1808, No. 1, p. 0200601-0200605, 2017.

[8] C. Muramatsu, T. Hara, T. Endo, and H. Fujita, “Breast mass classIfication on mammograms using radial local ternary patterns,” *Comput. Biol. Med.*, vol. 72, pp. 43–53, 2016.

[9] J. Dheeba, N. A. Singh, and S. T. Selvi, “Computer-aided detection of breast cancer on mammograms : A swarm intelligence optimized wavelet neural network approach,” Journal of Biomed. Inform., *49*, pp.45-52 2014.

[10] Leonard Berlin, “Radiologic errors , past , present and future,” vol. 1, no. 1, pp. 79–84, 2014.

[11] Whang, J.S., Baker, S.R., Patel, R., Luk, L. and Castro III, A. The causes of medical malpractice suits against radiologists in the United States. *Radiology*, *266*(2), pp.548-554, 2013.

[12] Sickles, Edward A. Periodic mammographic follow-up of probably benign lesions*:* results in 3,184 consecutive cases*.* Radiology, 179(2*),* pp.*463–468*,1991.

[13] T. Wang, C. Cheng, and H. Chiu, “Predicting Post-Treatment Survivability of Patients with Breast Cancer Using Artificial Neural Network Methods,” pp. 1290–1293, 2013.

[14] Essam Amin. M. Lotfy Abdrabou and A. M. Salem, “A breast cancer classifier based on a combination of case-based reasoning and ontology approach,” In Proceedings of th*e* International Multiconference on Computer Science and Information Technology pp. 3–10, 2014.

[15] Y. Jiang, R. M. Nishikawa, R. A. Schmidt, C. E. Metz, M. L. Giger, and K. Doi, “Improving breast cancer diagnosis with computer-aided diagnosis,” Acad. Radiol., vol. 6, no. 1, pp. 22–33, 1999.

[16] W. Xie, Y. Li, and Y. Ma, “Breast mass classification in digital mammography based on extreme learning machine,” Neurocomputing, vol. 173, pp. 930–941, 2016.

[17] Z. Chen, A. Oliver, E. Denton, and R. Zwiggelaar, “Automated mammographic risk classification based on breast density estimation,” Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 7887 LNCS, pp. 237–244, 2013.

[18] Matsubara, T., Yamazaki, D., Kato, M., Hara, T., Fujita, H., Iwase, T. and Endo, T. An automated classification scheme for mammograms based on amount and distribution of fibro glandular breast tissue density. In International congress series (Vol. 1230, pp. 545-552,2001.

[19] L. Tsochatzidis, L. Costaridou, and I. Pratikakis, “Deep Learning for Breast Cancer Diagnosis from Mammograms”, A Comparative Study." Journal of Imaging 5, no. 3: pp.37-46,2019.

[20] Arevalo, J., González, F. A., Ramos-Pollán, R., Oliveira, J. L., & Guevara Lopez, M. A. Representation learning for mammography mass lesion classification with convolutional neural networks. Computer Methods and Programs in Biomedicine, 127, pp.248–257,2016.

[21] N. Antropova, B. Q. Huynh, and M. L. Giger, “A deep feature fusion methodology for breast cancer diagnosis demonstrated on three imaging modality datasets,” Med. Phys., vol. 44, no. 10, pp. 5162–5171, 2017.

[22] A. Krizhevsky and G. E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks,” Proceedings of the IEEE Computer Society Conference on, vol.2 pp.1-9,2015.

[23] Z. Jiao, X. Gao, Y. Wang, and J. Li, “Neurocomputing A deep feature based framework for breast masses classification,” Neurocomputing, pp. 1–11, 2016.

[24] A. Rampun, B. W. Scotney, P. J. Morrow, and H. Wang, “Breast Mass Classification in Mammograms using Ensemble Convolutional Neural Networks,” 2018 IEEE 20th Int. Conf. e-Health Networking, Appl. Serv., pp. 1–6, 2018.

[25] Ting Fung, F. Yun. J. Tan, and Kok. S. Sim, “Convolution neural network improvement for breast cancer classification” Expert Syst.em with Applications,120 pp.103-115, 2019.

[26] K. Iftikhar, S. Anwar, I. U. Haq, M. T. Khan, and S. R. Akbar, “An Optimal Neural Network Based Classification Technique,” Journal of Engineering and Applied Sciences vol. 35, no. 1, pp. 39–46, 2016.

[27] X. Zhang, J. Zou, K. He, and J. Sun, “Accelerating Very Deep Convolutional Networks for Classification and Detection,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 38, no. 10, pp. 1943–1955, 2016.