A Comprehensive Study on the Phases and Techniques of Breast Cancer Classification

Adithya Narayanan[1], Mithilesh R[2], Swaminathan Chellappa[3], Anupama Y.K[*]

[1]UG Student, Department of Computer Science and Engineering, Dayananda Sagar College of Engineering Bangalore, Karnataka, India
[2]UG Student, Department of Computer Science and Engineering, Dayananda Sagar College of Engineering Bangalore, Karnataka, India
[3]UG Student, Department of Computer Science and Engineering, Dayananda Sagar College of Engineering Bangalore, Karnataka, India
[*]Assistant Professor, Department of Computer Science and Engineering, Dayananda Sagar College of Engineering Bangalore, Karnataka, India

---***---

Abstract - *Breast cancer, being one of the most prevalent cancers in the world, accounts for nearly one-third of all detected cancers in women. Researchers across the globe have put forth various techniques based on Machine Learning (ML) for the inspection and assessment of mammogram images, but these techniques have mainly resulted in errors with respect to observation and evaluation, which further resulted in erroneous diagnoses. This issue can however be overcome with the help of computer-assisted diagnosis which divides the diagnosis procedure into three phases: pre-processing phase, segmentation phase, and classification phase. Various techniques are utilised for each of the phases. Ef ective pre-processing deals with noise suppression, background removal, and edge detection. Segmentation deals with the separation of the area of concern from the breast tissue. Classification deals with the classification of the tumour present within the area of concern, into either malignant or benign. This paper sheds light on the comprehensive survey of the various algorithms and techniques available for breast cancer detection in detail.*

Key Words: Breast cancer, machine learning, pre-processing, segmentation, classification, area of concern.

1.INTRODUCTION

Breast cancer, being a prevalent form of cancer across the globe, accounts for about 27.7% of detected cancers in women in India [1]. The main causes of breast cancer in women are smoking, drinking, obesity, excessive stress, and other determinants. A sharp rise in cases has been recorded in Indian women, wherein a report from 2018 states that about 162,468 cases and 87,900 deaths were recorded, all due to breast cancer [1]. The year 2020 recorded 2.3 million cases of breast cancer and 685,000 fatalities across the globe. Its occurrence and fatality rates are also dependent on various factors, such as lack of access to medical facilities, annual income, living environment, etc., and vary from country to country [2]. Some countries oversaw an increase in the fatality rates, especially those that have large sections of the demographic classed as 'low-to-middle income'. This is probably a result of a lack of access to resources [2]. Studies conducted recently depict that on average, every tenth woman is in grave danger of contracting breast cancer, with it resulting in the fatalities of women who are aged between 35 and 54. Yearly, roughly 27% of new cases are that of breast cancer [3]. Hence, it is of foremost importance to educate women about the same and help them take part in screening examinations, which could possibly unearth the detection of some form of cancerous growth and could potentially save lives [4].

Mammography is widely recognized as the standard for screening. The data obtained post-screening should be analysed thoroughly, which would aid in providing a diagnosis in a quick and accurate manner. However, skilled radiologists are required for analysing the screening data. Unfortunately, the lack of skilled radiologists across the globe poses a major problem, mainly in most developing nations. India faces a severe shortage of trained radiologists, with only an estimated 10,000 available in the system [5]. This can lead to hindrances in the treatment and diagnosis. There still stands the task of classifying the lesions in the mammography images, which is a highly challenging task that requires at least an experienced radiologist or two to be able to accurately classify the lesions as benign or malignant.

Whilst experience plays an important role, human error plays a vital role as well, since there may exist error-prone mammogram interpretations due to fatigue, distractions, etc. This would result in harming the patients. Another statistic shows that 40-60% of biopsies of tumours diagnosed as malignant have been identified to be benign, which poses a rather tricky situation where if a tumour is left undetected and untreated it would cause further harm and if a tumour is misdiagnosed, then it would result in needless treatment. This has eventually led to the introduction and evolution of Computer-Aided Detection (CAD) systems. These help radiologists during a diagnosis. Lately, with the emergence of a number of CAD techniques, the relevant features pertaining to the mammogram images are learned and picked up by these systems, so as to help with the diagnosis. Therefore, using CAD systems aid in reducing the chances of misdiagnoses, whilst increasing mammography screening performance.

2. LITERATURE SURVEY

Roshan Lal Chhokar et al. [6] put forth a hybrid methodology to improve the detection of edges in images. This methodology is centered around the improvement in the accuracy of detection of edges and noise removal. A median filter filters out salt and pepper noise, which is then followed by applying a hybrid of the Sobel and Canny edge detection filters on the denoised image. Results showed that the hybrid approach was an improvement over traditional edge detection filters, which resulted in the images being richer.

Amer et al. [7] presented and compared five different edge detection filters using performance as the primary metric. Results showed that Robert's and Laplacian edge detection filters are extremely sensitive to random noise and also result in weak responses for legitimate edges. The Prewitt kernel is too small to detect fine edges and requires further processing. Sobel and Canny edge detection filters are not receptive to random noise, though computationally slower and expensive.

Abhijith Reddy Beeravolu, Sami Azam et al. [8] proposed various techniques for removal of background, removal of the pectoral muscle, the addition of noise, and image enhancements. The methods put forth for the removal of background utilised the "Rolling Ball Algorithm" and "Huang's Fuzzy Thresholding", which result in the removal of the background from hundred percent of the images. The methods used for pectoral muscle removal include "Canny Edge Detection" and "Hough's Line Transform", and resulted in the removal of the pectoral muscle from 99.6% of the images.

Guangxing Guo and Navid Razmjooy [9] made use of histogram equalisation as a preprocessing step. This technique enhances the contrast of a digital image, which is done by normalising the histogram of that image.

Vishnukumar K. Patel et al. [10] proposed a frequency domain smoothing-sharpening technique to enhance mammogram images. This proposed methodology initially involves a sequence of smoothing, detection of edges, and modifications of the histogram of various images to improve image contrast. This is followed by combining the Gabor filter with a fast Fourier transform, along with an overlay mask segmentation, which proved to be effective in eliminating noise and enhancing edges, thus yielding an improvement of the signal-to-noise ratio. The results show the contrast between the proposed approach with various edge detection filters, wherein the former obtains higher Peak Signal-to-Noise Ratio (SNRPEAK) values than the latter.

Paul Bao et al. [11] put forth an approach that is based on scale multiplication, so as to enhance the results of most edge detection filters. Their methodology multiplies the responses of the Canny edge detection filter at two different scales and was tested on both synthetic and natural images. Results depict the enhancements in the significant edges of the image, whilst diluting noise within the image. This also results in an improvement in edge localization accuracy and hence, leads to improved edge detection results.

Raman Maini and J.S. Sohal [12] presented a study that evaluates the performance of the Prewitt edge detection filter. Their study is based on edge detection in digital images distorted with different noises. A sample image of a coin was used, into which four different kinds of noise were introduced, namely Poisson noise, Gaussian noise, Salt and Pepper noise, and Speckle noise. Results depicted the successful detection of edges in the image coupled with both Poisson noise and Gaussian noise, with the former yielding better results. However, in contrast, it doesn't work well for images containing Salt and Pepper noise and Speckle noise, as the pixel values in the images containing the aforementioned noises are different from their surrounding pixel values. Since the Prewitt edge detection filter takes the average of the neighbouring pixel values, it distorts the pixel average value for images containing Salt and Pepper noise and Speckle noise. The authors also state that it can be used in tandem with other edge detection filters or noise reduction filters, in order to harness its true potential and reduce the effect of distortion within the image.

Laurence Aroquiaraj and K. Thangavel [13] proposed three different algorithms for edge detection in mammograms. These algorithms make use of fuzzy rules. The first algorithm makes use of a hybrid of Fuzzy and Canny edge detection filters, which secludes noise from the image, and then finds and locates the edges without affecting the edges' features. Algorithm two is a Fuzzy Relative Pixel Edge Detector, which makes use of a sliding window of a predefined dimension (3x3), in which the pixels are shifted and assessed, and then are arranged in a horizontal manner, and then is incremented to retrieve and assess the locations across the following vertical section. Once the edges are highlighted, the image pixels are made to undergo fuzzy conditions, which help to remove the unwanted parts and retain the edges that are important. The third algorithm is based on Fuzzy Edge Detection, which attains the gradient and standard deviation of the pixel values and edges, after which a fuzzy logic is applied, and a decision is made as to whether or not each pixel represents an edge. Results show that the three algorithms were compared to different filters such as Canny, Sobel, Prewitt, Robert's, and Laplacian of Gradient, and lower RMSE values, and SNRRMS values and SNRPEAK values were obtained by the proposed algorithms in comparison to the different edge detection algorithms, which further is a testament to the superior quality of the extracted edges using these proposed algorithms.

Samir M. Badawy et al. [14] proposed a scheme involving a combination of Fuzzy Logic and Deep Learning (DL) for machine-driven segmentation systems. In the pre-processing phase, contrast enhancement is achieved using a Fuzzy Intensification Operator (FIO). The paper then uses eight pre-trained CNN based segmentation schemes. After the segmentation is performed, the results are then assessed using three evaluation metrics: global accuracy, mean IoU, and mean BF (Boundary F1) score.

W. Gomez and W. Pereira [15] presented a qualified disquisition for semantic segmentation of cancerous breast regions in ultrasound images (UIs) using eight pre-trained convolutional neural networks (CNN). The main objective of the authors was to present a CNN-based segmentation model that is conservative with resources so as to be further employed in CAD systems. In order to segment the images into two classes, benign and malignant, transfer learning (TL) is applied. The F1-score (F1s) and the Intersection over Union (IoU) are used as the evaluation metrics. For the implementation of machine-driven systems to successfully detect breast cancer, the study presented ResNet18 as a probable candidate.

Min Dong et al. [16] put forth a method to bring about automatic segmentation and categorization of cancerous breast regions. In order to extract the area of concern from the mammogram images, the authors use the chain codes offered by the datasets. The area of concern is then enhanced using the rough set (RS) method. They then perform segmentation of the cancerous regions from the concerned areas with an enhanced vector field convolution (VFC) snake followed by feature extraction to generate a features database with 32 dimensions. In this paper, the random forest technique is used for classification and the performance is further compared to other ML classification techniques. The authors also use Matthew's correlation coefficient (MCC) indicator as an evaluation metric. The proposed method achieves the best performance with 97.73 % accuracy.

Richa Agarwal et al. [17] presented a novel patch-based CNN method for the detection of cancerous regions present in mammogram images. The paper also presents a juxtaposition between the suggested method and the use of TL. The proposed patch-based CNN method uses a sliding window approach where multiple patches from the input image are extracted. The datasets used are INbreast and CBIS-DDSM. In order to limit the number of patches being extracted a stride value of 56 was used. Each patch is then labelled as a possible mass candidate or not depending on the information as presented by the dataset. Patch classification is then carried out and evaluated using three CNN architectures, namely VGG16, ResNet50, and InceptionV3. After evaluation of results based on accuracy, the InceptionV3 model is seen to perform the best in both data sets. The authors also investigate and evaluate the implementation of TL for a specific domain adaptation; the best performing model on the CBIS-DDSM data-set is then fine-tuned and trained on the INbreast dataset. The usage of TL between the two datasets results in a higher accuracy when in comparison to directly training on the INbreast data-set. For the final experiment, the authors select the finest performing model from the previous trial for the detection of masses in an automated process without the need for human interference. The results conclude that the usage of TL between similar datasets outperforms randomly initialised CNN.

Samir M. Badawy et al. [18] introduced an upgraded double thresholding-based method for segmentation of cancerous breast regions. Double Thresholding Segmentation is achieved by assigning two pixel values as a lower limit and an upper limit from the specified image. According to the proposed approach, if the grey level of a pixel lies between the two limits (lower and upper) specified it is considered to be white else it is defined to be black. The output binary image is then made to undergo Masking and Morphological Operations. Since the mammogram images consist of undesirable borders, a specific mask has been developed by the authors to remove them accordingly. After applying morphological operations the final enhanced segmented image is produced. The distinguishing features of the algorithm are the reduced memory and processing time.

Wessam M. Salama and Moustafa H. Aly [19], proposed their framework for the segmentation and categorization of cancerous breast regions. In this paper, the authors incorporate the use of TL and data augmentation (DA). Categorization of the cancerous regions into benign and malignant images is achieved using different pre-trained models on the ImageNet dataset. Region-based segmentation is achieved using a trained modified U-Net model. The segmented area of concern is then classified into benign or malignant by feeding them into the different deep learning models. Incorporating the use of DA with the enhanced U-Net model and InceptionV3 is concluded to present the best results.

S.Punitha et al. [20] have presented an optimised region-growing(RG) technique for the detection of breast masses. To eliminate the noise from the images, the Gaussian filter is applied as a preprocessing step. A swarm optimization technique called Dragon Fly Optimization (DFO) is used to extract the area of concern where the starting seed points and thresholds are efficiently produced. Feature extraction is achieved using Gray level co-occurrence Matrix (GLCM) and Gray level Run-length Matrix (GLRLM). Once the features have been extracted they are fed into a Feed-Forward Neural Network, trained using a backpropagation algorithm called Levenberge Marquardt, to perform classification as normal, benign, or malignant. The system put forward by the authors performs at an accuracy of 98% and outperforms the other RG techniques when compared using ROC analysis.

Anuj Kumar Singh and Bhupendra Gupta [21] propose a novel technique to segment cancerous regions in breast cancer mammogram images along with detection. Since background information and other regions of a mammogram image have lower levels of intensity in comparison to that of malignant tissues, the authors try to detect the malignant tissues by considering regions with high-intensity values. In the detection phase, smoothing of the image is achieved using an averaging filter. The malignant region area is then extracted using a thresholding operation. A Max-Mean and Least-Variance technique is then applied on a rectangular window that is created on all sides of the output region to find the malignant tissues. Segmentation is achieved using morphological closing operation and image gradient technique, which outputs the area of concern with respect to the tumour regions. The resultant region boundary of the segmented images is then highlighted.

Meriem Amrane et al. [22] proposed the comparison of two machine learning techniques, namely the Naive Bayes(NB) classifier and the K-Nearest Neighbour classifier (KNN) for breast cancer classification. The Wisconsin breast cancer database is being used in this paper. The classification is binary in nature, with the two classes being benign and malignant. The database denoted benign classes with 2 and 4 respectively. The NB classifier and the KNN classifier (with k = 3) were trained using the above-mentioned dataset and they were evaluated using K-fold cross-validation. The results of their findings are that the KNN model performed better with a 97.51% accuracy whereas the NB classifier is slightly lower at 96.19% accuracy, in the case that the dataset is larger, the KNN classifier would require a larger amount of time to run.

Chelvian Aroef et al. [23] put forth a comparison between two machine learning algorithms, Random Forest and Support Vector Machine classifiers. The data is obtained from the UCI machine learning repository, which has a total of nine features,116 observations split between 64 patients and 54 healthy individuals; there exist two classes of classification. Feature selection is done using the Boruta Feature Selection which is built around the algorithm for random forest classification, the results obtained depict that Age, BMI, Glucose, HOMA, and Resistin are important features. The Random Forest (RF) classifier is an ensemble method classifier and the model had the best accuracy of 90.90% with 80% of training data and the worst accuracy of 74.75% with 10% of training data. The Support Vector Machine (SVM) classifier with the Radial Bias Function (RBF) kernel obtained an accuracy of 95.45% with 80% training data and the worst accuracy of 72.81% with 10% of training data. Both the models were evaluated using Hold-Out validation and inference is drawn that the SVM classifier with the RBF kernel outperforms the RF classifier.

T. Velmurugan [24] proposed evaluating the results of various decision tree machine learning algorithms in the classification of breast cancer based on the age of patients and categories of cancer type. The data set used contains data from 220 patients, the data is used considering the ages within 20 and 72 years. The classification algorithms evaluated are J48 algorithm, Classification and Regression Tree (CART), Alternating Decision Tree (ADTREE), and Best First Tree (BF TREE). The models are evaluated using the WKEA application, the performance is evaluated using 10-fold cross-validation and Percentage split where ⅔ is for training and ⅓ is for testing. The measures used are the rate of false positives, rate of true positive, Recall, Precision, Receiver Operating Characteristic Area, and F-measure. The accuracy of the J48 algorithm is 99%, BF Tree is 98%, AD Tree is 97% and CART is 96%. Hence, it is concluded that the J48 performs the best for this particular data set.

Mohamad Mahmoud Al Rahhal [25] proposes an approach using CNNs which are pre-trained on another domain with huge quantities of labelled images, this is integrated with another network made of fully connected layers. The data set used is large-scale as well as abundant in images, it contains 9109 images from 82 patients. The proposed method uses a pre-trained CNN to aid in feature extraction. The output CNN feature vectors are fed to an additional network that consists of 2 fully-connected layers, a hidden layer and a sigmoid layer for binary classification; a dropout technique is used to increase generalisation and reduce the possibility of overfitting. The results when using VGGm (8 layers) as the CNN which is trained on ILSVRC-12 challenge dataset, the average accuracy obtained for this model is 86.80%. Possible improvements to the model include customising and integrating several deep architectures to increase performance, domain adaptation can also produce improvements.

Royida A. Ibrahem Alhayali et al [26] proposed an ensemble machine learning technique involving the Naive Bayes Classifier (NBC) and the Hoffeding Tree. The dataset used is the Wisconsin Breast Cancer Database (WBC), it includes 10 patient features with 699 instances out of which 16 instances contain null values, hence a total of 683 instances are used. Their proposed model is an ensemble of the Hoffeding Naive Bayes Tree. In this technique, an NB prediction is applied to each training attribute, then has the performance of the prediction compared with the predominant class. The number of correct predictions of the true class is recorded when compared to the predominant class. While the test sample is being predicted, the leaf only outputs a prediction based on NB when the accuracy is more than the predominant class, if not, the output will be a predominant class prediction. When the incorrectly classified instance is forwarded to the NBC and is assessed in the next stage, the accuracy obtained was 95.9943%, without which it recorded an accuracy of 88.33%. This also concludes that the proposed model outperforms conventionally used algorithms or models. It can also be concluded that the accuracy of the Naive Bayes classifier can be improved by conditional independence being assumed.

Pachi Damodhar Shahare et al. [27] propose a comparison of performance with regards to accuracy between Support Vector Machines(SVM) with different kernels and Artificial Neural Networks. The dataset used is procured from the University of Wisconsin Hospitals, Madison by Dr. William H. Wolberg. The ANN proposed has the following architecture: 9 neurons in the input layer,121 neurons in the hidden layer, and 2 neurons in the output layer. The SVM was evaluated using 5 different kernels. The results obtained depicted that the ANN with a feed-forward network had an accuracy of 96.15%, the SVM had an accuracy of 99% with the linear kernel, 96% with the quadratic kernel,95% with the polynomial kernel, 98.5% with the Radial Bias Function kernel and 98.5% with the Multilayer Perceptron. Hence, it is concluded that the ANN does not perform as well as the SVM with its various kernels, it is also mentioned that the combination of kernels can be explored as it might improve performance.

Ines Domingues et al. [28] propose a couple of new pre-processing methods for the improvement in the performance of deep classifiers, that is data augmentation and pyramid of scales. These methods were tested on the InBreast dataset by classifying the dataset with AlexNet, a Convolutional Neural Network(CNN). The accuracy improved in excess of 33% with data augmentation and had seen an improvement in excess of 32% when data augmentation was performed on the training dataset by mirroring the images. The pyramid of scales approach has a slightly lighter improvement in accuracy of 3%. The pyramid of scales approach consists of using a three-channel matrix of Difference in Gaussians images (DoG) as the input to the CNN. It is concluded that in the future, non-linear transformations can be considered for data augmentation by co-registration and the combination of craniocaudal (CC) view and mediolateral oblique (MLO) views can be used to result in a well-informed decision.

Kundan Kumar et al. [29] proposed a CNNs for the categorization of breast cancer. The dataset used has 5,429 malignant and 2,480 benign samples. The proposed CNN has an input layer for RGB images with 32x32 or 64x64 resolution,6 convolutional layers,3 with 5x5 kernel and the remaining 3 with a 3x3 kernel, Rectified Linear Unit as the activation function, Pooling layers with a 3x3 kernel, the first 3 layers use max-pooling while the remaining 3 layers use average pooling and finally an output layer that uses the softmax function to classify an image as either benign or malignant. The model is further optimized using stochastic gradient descent. The model is validated using 10-fold cross-validation and the proposed model obtains 90% accuracy, it is concluded that while the obtained accuracy is acceptable, there exists room for improvement towards a model with 100% accuracy.

3. CONCLUSION

A survey of Breast Cancer classification is explained in this paper which includes three phases: image pre-processing, segmentation, and classification of the image into the various classes. This aims to aid in the improvement of existing or future classification or detection of Breast Cancer and help people across the globe, especially those who lack the finances or necessary infrastructure to get themselves screened and diagnosed for the same.

REFERENCES

[1] Gouri Shankar Bhattacharyya, Dinesh C. Doval, Chirag J. Desai, Harit Chaturvedi, Sanjay Sharma, and S.P. Somashekhar JCO Global Oncology 2020 :6, 789-798

[2] F. Z. Francies, R. Hull, R. Khanyile, and Z. Dlamini, ``Breast cancer in low-middle income countries: Abnormality in splicing and lack of targeted treatment options,'' Amer. J. Cancer Res., vol. 10, no. 5, p. 1568, 2020.

[3] A. Jemal, R. Siegel, E. Ward, Y. Hao, J. Xu, and MJ. Thun. Cancer statistics, 2009, CA Cancer J Clin., Vol.59, pp. 225-249, 2009

[4] D. Saslow, C. Boetes, W. Burke, S. Harms, M. O. Leach, C. D. Lehman, E. Morris, E. Pisano, M. Schnall, S. Sener, R. A. Smith, E. Warner, M. Yaffe, K. S. Andrews, and C. A. Russell, ''American cancer society guidelines for breast screening with MRI as an adjunct to mammography,'' CA A, Cancer J. Clinicians, vol. 57, no. 2, pp. 7589, Mar. 2007.

[5] Apollo Hospitals partners with the Royal College of Radiologists to address the shortage of skilled workforce in India and the UK. Available: https://apolloradiologyintl.com/apollo-hospitals-partners -with-royal-college-of-radiologists-to-address-shortage-of skilled-workforce-in-india-and-the-uk/#:~:text=Currently %2C%2083%20countries%20fall%20below,10%2C000% 20available%20in%20the%20system.

[6] Kalra, Anchal & Lal, Roshan. (2016). A Hybrid Approach Using Sobel and Canny Operator for Digital Image Edge Detection. 305-310. 10.1109/ICMETE.2016.49.

[7] Amer, Ghassan & Ahmed, Mohamed & Abushaala, Ahmed. (2020). Edge Detection Methods.

[8] A. R. Beeravolu, S. Azam, M. Jonkman, B. Shanmugam, K. Kannoorpatti, and A. Anwar, "Preprocessing of Breast Cancer Images to Create Datasets for Deep-CNN," in IEEE Access, vol. 9, pp. 33438-33463, 2021, DOI: 10.1109/ACCESS.2021.3058773.

[9] Guangxing Guo & Navid Razmjooy (2019) A new interval differential equation for edge detection and determining breast cancer regions in mammography images, Systems Science & Control Engineering, 7:1, 346-356, DOI: 10.1080/21642583.2019.1681033.

[10] Patel, Vishnukumar & Uvaid, Syed & Suthar, Anilkumar. (2012). Mammogram of Breast Cancer detection Based using Image Enhancement Algorithm. 10.13140/RG.2.2.13042.86723.

[11] P. Bao, Lei Zhang and Xiaolin Wu, "Canny edge detection enhancement by scale multiplication," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, no. 9, pp. 1485-1490, Sept. 2005, DOI: 10.1109/TPAMI.2005.173.

[12] Maini, Raman & Sohal, J.. (2006). Performance Evaluation of Prewitt Edge Detector for Noisy Images. GVIP Journal. 6.

[13] Aroquiaraj, I. Laurence and K. Thangavel. “Mammogram Edge Detection Using Hybrid Soft Computing Methods.” ArXiv abs/1307.4516 (2013): n. pag.

[14] Badawy SM, Mohamed AEA, Hefnawy AA, Zidan HE, GadAllah MT, El-Banby GM. Automatic semantic segmentation of breast tumors in ultrasound images based on combining fuzzy logic and deep learning-A feasibility study. PLoS One. 2021 May 20;16(5):e0251899. DOI: 10.1371/journal.pone.0251899. PMID: 34014987; PMCID: PMC8136850.

[15] Gómez-Flores W, Coelho de Albuquerque Pereira W. A comparative study of pre-trained convolutional neural networks for semantic segmentation of breast tumors in ultrasound. Comput Biol Med. 2020 Nov;126:104036. DOI: 10.1016/j.compbiomed.2020.104036. Epub 2020 Oct 8. PMID: 33059238.

[16] Dong, Min & Lu, Xiangyu & Ma, Yide & Guo, Ya’nan & ma, Yurun & Wang, Keju. (2015). An Efficient Approach for Automated Mass Segmentation and Classification in Mammograms. Journal of digital imaging. 28. 10.1007/s10278-015-9778-4.

[17] R. Agarwal et al., “Mass detection in mammograms using pre-trained deep learning models,” Proc. SPIE 10718, 107181F (2018).

[18] Samir M. Badawy, Alaa A. Hefnawy, Hassan E. Zidan and Mohammed T. GadAllah, “Breast Cancer Detection with Mammogram Segmentation: A Qualitative Study” International Journal of Advanced Computer Science and Applications(IJACSA), 8(10), 2017. http://dx.doi.org/10.14569/IJACSA.2017.081016

[19] Salama, Wessam & Aly, Moustafa. (2021). Deep learning in mammography images segmentation and classification: Automated CNN approach. Alexandria Engineering Journal. 60. 4701-4709. 10.1016/j.aej.2021.03.048.

[20] Punitha, S. & Amuthan, A. & Joseph K, Suresh. (2018). Benign and Malignant Breast Cancer Segmentation Using Optimized Region Growing Technique. Future Computing and Informatics Journal. 3. 10.1016/j.fcij.2018.10.005.

[21] Singh, Anuj & Bhupendra, Gupta. (2015). A Novel Approach for Breast Cancer Detection and Segmentation in a

Mammogram. Procedia Computer Science. 54. 10.1016/j.procs.2015.06.079.

[22] M. Amrane, S. Oukid, I. Gagaoua and T. Ensarİ, "Breast cancer classification using machine learning," 2018 Electric Electronics, Computer Science, Biomedical Engineerings' Meeting (EBBT), 2018, pp. 1-4, DOI: 10.1109/EBBT.2018.8391453.

[23] Aroef, Chelvian & Yuda, Rivan & Rustam, Zuherman. (2020). Comparing random forest and support vector machines for breast cancer classification. TELKOMNIKA (Telecommunication Computing Electronics and Control). 18. 815. 10.12928/telkomnika.v18i2.14785.

[24] Venkatesan, E Venkatesan & Thambusamy, Velmurugan. (2015). Performance Analysis of Decision Tree Algorithms for Breast Cancer Classification. Indian Journal of Science and Technology. 8. 10.17485/ijst/2015/v8i1/84646.

[25] Mohamad Mahmoud Al Rahhal, "Breast Cancer Classification in Histopathological Images using Convolutional Neural Network" International Journal of Advanced Computer Science and Applications(IJACSA), 9(3), 2018. http://dx.doi.org/10.14569/IJACSA.2018.090310

[26] Alhayali, Royida & Ahmed, Munef & Makki, Yasmin & Ali, Ahmed. (2020). Efficient method for breast cancer classification based on ensemble hoffeding tree and naïve Bayes. Journal of Electrical Engineering. 18. 1074-1080. 10.11591/ijeecs.v18.i2.pp1074-1080.

[27] Shahare, Prachi Damodhar, and Ram Nivas Giri. "Comparative analysis of artificial neural network and support vector machine classification for breast cancer detection." Int. Res. J. Eng. Technol 2 (2015): 2114-2119.

[28] I. Domingues, P. H. Abreu and J. Santos, "Bi-Rads Classification of Breast Cancer: A New Pre-Processing Pipeline for Deep Models Training," 2018 25th IEEE International Conference on Image Processing (ICIP), 2018, pp. 1378-1382, DOI: 10.1109/ICIP.2018.8451510.

[29] K. Kumar and A. C. S. Rao, "Breast cancer classification of image using convolutional neural network," 2018 4th International Conference on Recent Advances in Information Technology (RAIT), 2018, pp. 1-6, DOI: 10.1109/RAIT.2018.8389034.