Mammogram Image Classification Using Bi-LSTM and Random Forest

Dr. S. Satheesh[1], Sivanagireddy Kalli[2]

[1]*Professor, Department of ECE, RISE Krishna Sai Prakasam Group of Institutions, Ongole, A.P., India*

[2]*Associate Professor, Department of ECE, Sridevi Women's Engineering College, Hyderabad, Telangana*

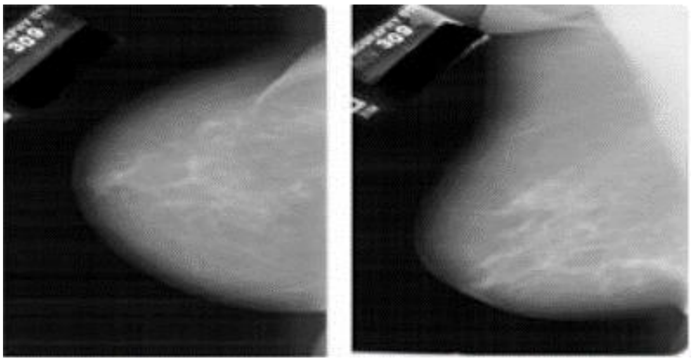
Abstract – Breast imaging diagnosis is one of the fastest-growing fields in radiology. A wide range of Mammograms have been marked during the past decade, and calcification of these images is prominent in diagnosis. Computer vision algorithms with appropriate feature vectors are used to perform this process. This work presents an accurate approach to extracting efficient features with the help of Attention layers and reducing error by the Bi-LSTM approach because LSTM has efficient memory to reduce error. We have worked on two benchmark datasets, MIAS and DDSM, which are scientifically proven datasets. An efficient approach is the CNN-Bi-LSTM Attention-based approach which Random forest learns. The proposed approach based on Attention and Bi-LSTM significantly improved all performance metrics. Compared to existing approaches described in the comparison section, its improved accuracy is 2-3%, precision 2%, recall 3% and ROC 3%.

Keywords: Mammogram Calcification, Convolutional Neural Networks, Bi- Long Short Term Memory (LSTM), machine learning.

I. INTRODUCTION

Breast disease is considered one of the most genuine clinical issues globally. Nowadays, breast tumor signifies 23% of every investigated sickness and 14% of development-associated deaths [1, 11]. Over the earlier decade, the assumption related to breast screening strategies similar to X-Ray mammography [3] acts as an early discovery of the tumor, thus reducing the passing/death rate. Women can get genuine investigation at the starting time of tumor acknowledgment. A mammogram is a broadly utilized and dependable screening innovation, i.e., helpful in identifying the growth of breast malignant [2]. The microcalcifications represent an early sign of malignant or cancer growth [12]. Distinguishing the microcalcifications and their confinement in breast tissue is typically done by radiologists in screening. According to the rule, a breast mass is believed to be destructive if its shape is theorized or unpredictable. The classification based on Microcalcifications relies upon their distribution, shape, size, and number. Bosom screening utilizing X-beam mammograms [9] is completed by taking images of a similar breast from two unique perspectives, i.e., mediolateral oblique (MLO) and craniocaudal (CC) view as appears in figure 1. With the help of breast cancer detection and lesion segmentation on mammograms, radiologists evaluate the hazard associated with the discovery of the masses and micro (small-scale) calcification bosom malignancy injuries that are either benign or malignant. Masses are typically grey to white in pixel-based intensity, and in the geometrical form, they can have ill-defined or obscured margins, lobulated or irregular shapes, and oval masses.

On the other hand, microcalcifications are generally bright round areas, as appears in figure 2. Generally, a bosom mass is viewed as dangerous if its shape represents an irregular lesion. Micro-calcifications depend on their distribution, shape, size, and number. The process of detecting, segmenting, and classifying the masses in mammograms, for the most part, is done on a manual basis i.e., tedious, and radiologist's skill and weakness level support in quick location or detecting of the lesion [2]. The sensitivity concerning manual classification and detection fluctuates between 80% to 90%, with 91% of specificity [3]. One method for expanding sensitivity and specificity is the twofold perusing of mammograms, which has been found to boost the sensitivity by 9% and lessen the number of women reviewed for additional tests by 45% [11]. The idea of a subsequent per user gives high specificity and sensitivity.



(i) (ii)

Figure 1: Mammogram Images: (i) in breast Dataset: CC and MLO; (ii) DDSM: CC and MLO [11]

A few examinations conducted in [2, 11] carried out by radiologists explain that the sensitivity in detecting and classifying masses is practically improved by 10% with specificity at a similar level to the utilization of a CAD framework. A CAD framework may represent semi-automatic nature, needing conciliation by a specialist at some stage, or it may be fully or completely automatic, which needs no master help. Essentially, to the manual investigation of masses portrayed above, a CAD framework for the automatic examination of masses works when all is said in done in three stages: mass detection, mass segmentation, and mass classification. These stages are challenging as there is a drastic transformation of size, shape, location, and appearance of masses in mammogram technology, visual mass appearance based on low SNR, and the absence of openly accessible datasets. The datasets are precisely interpreted with the procedure FFDM, i.e., full-field digital mammograms, which is the fundamental imaging methodology utilized in bosom screening processes. Numerous strategies have been taken from mammograms utilized in automatic detection, segmentation, and classification of masses. For the most part, detecting mass from mammograms utilizing a candidate mass gives an FP, i.e., false positive.

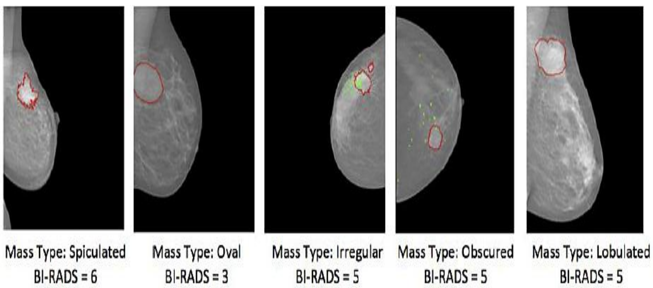


Figure 2: Lesions in mammograms from INbreast Dataset: Calcifications are symbolized using Green Contours and Masses are symbolized using Red Contours [11]

Decrease with various kinds of machine-learning (ML) classifiers [4]. Detection of mass candidate implemented with methods, for example, deforming models dependent on active (dynamic) contour models, thresholding edge-based detection utilizing various channels/filters, and statistical techniques, for example, area-based clustering, growing utilizing Markov random field (MRF) and k-means method.

Generally, the detection step of the mass candidate provides a very high rate for each image. It needs an additional step for decreasing false positives (FP) with classification and extraction of hand-made features [6]. These CAD frameworks generally deliver high FPRs [4] because the detection of the mass candidate and hand-made features are structured on a sub-optimal basis based on the geometry and appearance of masses. Classification of masses into malignant/ benign represents a two-step procedure where segmentation is regarded as the main step, and mass classification is considered the second step. Division or segmentation is typically done to remove the geometrically built features from the segmented contour of the mass. Customary graphical approaches and dynamic/active contour are the two best calculations based on algorithms for the segmentation of masses [1]. The primary issue with these mass division approaches is physically characterized appearance and shape terms. A cross-validation sub-optimal technique is used to learn the parameters of the segmentation model. The classification of mass finds the abstraction of hand-made features utilized by ML classifiers, for example, ANN, SVM, LDA, and so on [4, 8, 10].

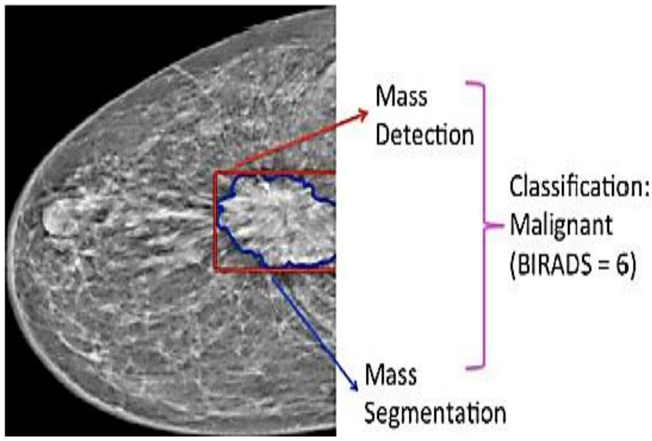


Figure 3: Mass Segmentation, Classification, and Detection [11]

The significant downsides of the techniques based on mass classification represent the sub-optimality in the features plan. For the most part, such frameworks are semi-automatic, needing the manual determination of mass candidates [26]. Deep learning (DL), with its progressive component's interpretation, has created better accuracy and better classification contrasted with other techniques of ML methods using hand-made features [5]. The models based on deep learning are prepared. The features at each level naturally learning the model hierarchy or order are enhanced or optimized via curtailing the loss function, which thus gives better segmentation, classification, and detection results. The features resulting from deep/profound learning techniques are optimal for detection, classification, or segmentation undertakings, which is a massive development contrasted with the previously mentioned hand-created features without limiting loss function. Though the training of DL models is troublesome due to their high limit or capacity-based issues (for example, a bosom mass examination from mammograms), it is very complex to find out marked datasets comprising many preparing tests that would take into account a robust preparation. DL models have been investigated for different clinical image examination issues, such as lymph node discovery, mitosis location, and significant level characterization/classification of multi-modular input [2]. A characteristic inquiry emerges if DL models can perform better than the cutting-edge strategies in CAD frameworks that examine mammograms. This inquiry has been our fundamental drive to test profound learning in mammograms' mass detection, mass segmentation, and mass classification.

II. RELATED WORK

Lee J.G et al. [1] offers points of view based on the history, improvement, and utilization of profound learning innovation applications, especially concerning its applications in clinical imaging. The interest in such a concept has resurfaced lately because of the accessibility of huge information, improved computing power with the present design processing units, and new algorithms to prepare the profound neural system. Late examinations of this innovation propose it possibly accomplished better than people in most visual and sound-related acknowledgment tasks that could predict its applications in healthcare and medicine, particularly in clinical imaging, within a reasonable time frame.

T. Kooi et al. [2] explore expanding temporal data to a profound CNN to identify dangerous, delicate tissue injuries in mammography. The analysts utilize a straightforward, direct mapping taking into account an area of a mass candidate and mapping it to either the prior or contralateral mammogram. ROIs are separated within every area. Two distinct designs are consequently investigated: (1) a combination model utilizing two information streams where the two ROIs are entered into the system during testing and training and (2) a stage-wise method where a solitary ROI-CNN is prepared on the essential image and in this manner utilized as a component extractor for both essential and earlier or contralateral ROIs. A "shallow" inclination supported tree classifier is finally prepared on the link of such features and is utilized for characterizing the joint portrayal. The pattern produced a 0.87 AUC with a certainty interval [0.893, 0.853].

Kenji Suzuki [3] reviewed the region of profound (deep) learning in clinical imaging together with (1) what was altered in ML before and after when the presentation of profound (deep) learning, (2) what represents the power source of the intensity-based profound learning, (3) two significant deep learning models: CNN and a massive-training ANN (TANN), (4) their applications to clinical imaging (5) differences and similarities amongst the two models. This audit shows that ML with include info (ML feature-based) was prevailing before the presentation of profound learning, which represents an essential and major distinction amongst ML after and before the deep learning represents the learning of image-based legitimately deprived of feature extraction or object segmentation; along these lines, it represents the power-based source of profound learning, even though the profundity of the model is a significant characteristic.

T. Kooi et al. [4] assessed the literature-based writing on segmenting BUS images as the systems embraced, particularly in recent years. By isolating into 7 classes (i.e., clustering-based, thresholding-based, graph-based, watershed-based, active contour model, neural network, and Markov random field), scientists have presented relating procedures and agent papers as needed. The scientists have outlined and thought about numerous BUS-based image segmentation systems and initiate that every one of these methods possesses its upsides and downsides. However, image BUS-based segmentation is an open and testing issue because of different ultrasound curios presented during the imaging process, including imaging involving blurry boundaries, high speckle noise, low intensity, low contrast inhomogeneity, and low speckle noise, low intensity, low contrast inhomogeneity SNR.

Gustavo Carneiro et al. [5] depict a robotized technique for the examination of craniocaudal (CC), mediolateral oblique (MLO), and mammography sees to evaluate the risk of emerging bosom malignant growth. This strategy's principle development uses deep learning models to mutually characterize unregistered mammographic audits and separate division maps of bosom injuries (i.e., masses and micro-calcifications). This represents an encompassing strategy that can order an entire mammographic test, containing the MLO and CC reviews and the division maps, instead of the characterization of individual sores, which is the predominant methodology in the field. The specialists also show that the proposed framework is fit for utilizing the segmentation-based maps created via micro-calcification detection systems and automated mass frameworks and delivering accurate outcomes. The semi-robotized method (utilizing physically characterized micro-calcification and mass segmentation maps) is tried on two openly accessible informational collections (DDSM and INbreast). Consequences illustrate that the volume for a 3-class issue (benign, malignant, and normal tissue) below ROC surface (VUS) is over 0.9, for a 2-class bosom screening issue (benign/ normal vs. malignancy) is likewise over 0.9, and the AUC for the 2-class screening issue is over 0.9. For the completely mechanized methodology, the VUS outcomes on INbreast is over 0.7, the AUC for 2-class breast screening is 0.86, and the AUC for the 2-class "malignant vs. benign" issue is over 0.78.

Dinggang Shen et al. [6] cover PC helped examine images in clinical imaging. Ongoing advances in ML, particularly in the mechanism of profound learning, assist with distinguishing, characterizing, and measuring designs in clinical images. At the center of these advances is the capacity for exploiting various levels, including portrayals gained exclusively from information, rather than features planned by hand as indicated by domain-specific information. Deep (or profound) learning quickly becomes the cutting edge, prompting upgraded execution in different clinical applications. The specialist presents the basics of profound learning strategies and surveys their achievements in image enlistment, the discovery of cellular and anatomical structures, and tissue-based division. CAD helps disease prognosis and diagnosis, etc. They close by talking about research issues and recommending future headings for additional improvement.

Neeraj Dhungal et al. [7] present a new CAD framework (i.e., automated in nature with negligible client mediation that can identify, section, and further classify the bosom masses from mammographic images. The scientists investigate profound (or deep) learning and organized yield models for the structure and advancement of the proposed CAD framework. All the more explicitly for the discovery, the analysts propose a course of deep learning strategies to choose refined theories dependent on the Bayesian-based optimization method. For the division or segmentation, the analysts propose the utilization of deeply organized yield discovery, i.e., refined by a level set strategy. At long last, for the arrangement, the analysts propose a mechanism of deep learning classifier that is pre-prepared with regression to hand-made feature (or elemental) values and adjusted depending on the explanations of the bosom mass grouping dataset. The proposed CAD framework delivers the present best in class detection, segmentation, and outcomes classification for the INbreast dataset.

Xiaofei Zhang et al. [8] got more than 3000 excellent unique mammograms with endorsement from an institution-based audit board at the Kentucky University. Various classifiers dependent on CNNs were assembled. Every classifier was assessed depending on its exhibition comparative with truth-based values created by 2-year negative mammogram follow-up and histology consequences from biopsy acknowledged by master radiologists. The outcomes demonstrated that the CNN model was created and optimized using data augmentation and transfer learning. They have an extraordinary potential for programmed bosom malignant growth recognition utilizing mammograms.

Thijs Kooi et al. [9] give a face-to-face examination between a cutting-edge art in mammography-based CAD framework, depending on a physically structured list of capabilities, and a CNN, focusing on a framework that can look at last read mammograms freely. The two frameworks are prepared on a huge informational collection of around 45,000 images. Results show that the CNN outflanks the conventional CAD framework at low sensitivity and performs equivalent at high sensitivity. The scientists in this manner examine to what degree features, for example, area and patient data and regularly utilized manual features, can supplement the system and see enhancements at high-specificity over the CNN, particularly with the area and contextual features that comprise data not accessible to CNN.

Estefanía D. Avalos-Rivera et al. [10] center to group the process into 2, 3, and 4 BIRADS classifications in which the development of malignant growth can be forestalled. Our strategy comprises choosing an ROI; at that point, as a pre-processing system, a High Pass Gaussian channel is applied to the image to upgrade the sores and afterward develop a double mask to get its morphological descriptors. This information is standardized and driven into ANN, prepared with a dataset procured from UNEME DEDICAM in Mexico. Such a dataset incorporates its determination into BIRADS. This research utilizes a two-layered feed-forward network system, with 10 shrouded systems and a sigmoid covered up. It grouped BIRADS 2 and 3 into an individual classification presenting the best outcome as 80% of the expectations are right.

Neeraj Dhungal et al. [11] diagrammed this manual procedure of clarifying mammograms that includes the location of bosom sores (e.g., masses), the division-based segmentation of lesion-based boundaries, and the grouping of sores dependent on their appearance, shape, and textural features. This manual examination of bosom sores from mammograms presents huge variable interpretations among radiologists. This inconstancy can be diminished with the guidance of CAD-supported frameworks that can go about as a second per-user in examining bosom injuries.

III. PROPOSED APPROACH

Though for a CAD framework to be helpful in a clinical setting, it should viably group injuries as considerate or threatening. The process of detecting, segmenting, and classifying the bosom sores are the fundamental three stages engaged with completely computerized CAD frameworks that can work in examining mammograms. Constructing a CAD framework is troublesome because a low sign defaces mammograms to noise proportion for the perception of bosom sores. Furthermore, bosom injuries present a huge variety in size, appearance, and shape. Many strategies have been applied to build robotized CAD frameworks for the two kinds of injuries, particularly mass and small-scale calcification. Yet, right now, scientists focus just on investigating the masses.

Figure 4 depicts the flow chart of the proposed approach. The input images are gathered from the MIAS dataset, which is of grayscale intensity level images whose pixel values range [0 255]. Later, during the pre-processing of the images, features are extracted from the images and marked with their labels for classification. For extraction of features, convolutional layers are employed, which are forwarded to Bi-LSTM layers, which provide a long-term dependency for a series of image sequences. These features are optimized and trained with a Random forest classifier for prediction.

In the second phase of the system, when an unknown sample is provided as input, the trained classifier predicts the class of mammogram images. The Bi-LSTM layer employs a variety of algorithms to calculate the hidden information. An LSTM is trustworthy when used with the standard RNN structure. It overcomes the limitation of Recurrent ANNs in managing long-distance dependency. An LSTM approach entails many memory units, each of which has three gates with various functions. The next section provides settings for the criteria of the LSTM unit of s^{th} patches using the feature vector S as input and the s^{th} patches as an example. A specific calculation equation is used, which denotes the sigmoid function and denotes dot multiplication.

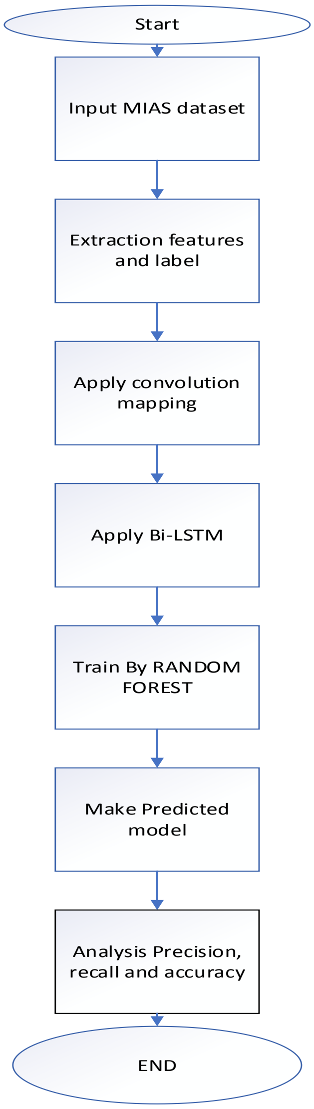


Figure 4: Proposed Approach

In the Layer Bi-LSTM the X_t shows the forget gate:

$$X_t = \propto (W_x w_s + U_x h_{s-1} + bias_x) \quad (1)$$

The j_s is the input gate:

$$j_s = \propto \left(W_j w_j + U_j h_{s-1} + bias_j\right) \quad (2)$$

The $\tilde{u}_s$variable represents the status of the candidate memory cell at the most recent time step, where tanh signifies the tangent hyperbolic function.

$$\tilde{u}_s = tanh\left(W_u w_s + U_j h_{s-1} + bias_u\right) \quad (3)$$

IV. EXPERIMENTAL RESULTS

In order to validate the proposed approach, the method is tested with standard MIAS dataset with 322 samples. Apart from these DDSM dataset images are used which comprises of 10,480 samples. Figure 5, depicts the performance comparison of proposed method with earlier SVM based classifier and traditional CNN network methods. In the proposed method Bi-LSTM with random forest classifier is employed to classify the defective or abnormal mammogram images.

Table 1: Comparison of proposed and existing

Proposed Approaches	Accuracy	Precision	Recall	ROC
CNN	98.56	98.23	98.45	98
BilSTM-Random forest	99.23	98.99	99.123	99.12
SVM	98.12	98	97.12	97

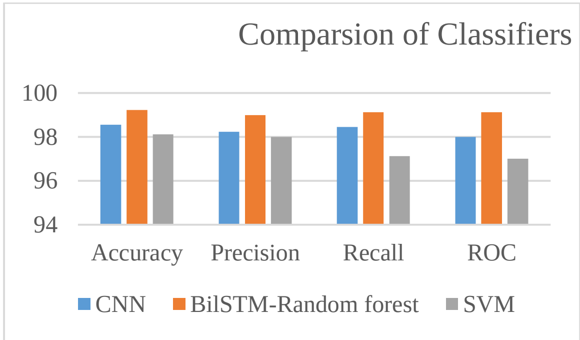


Figure 5: Comparisons of proposed classifier performance with earlier classifiers

The experimental results and analysis can be observed from table 1 and figure 5 that the proposed Bi-LSTM-based Random forest classifier could classify the mammogram with an average accuracy of 99.23%, which is a 6.7% improvement when compared with the traditional CNN based method. However, it can be observed that from all perspectives, the proposed approach is yielding considerable improvements against traditional classifiers.

V. CONCLUSIONS

Breast cancer detection and classification is an essential medical diagnosis by biomedical images, and it becomes challenging for radiologists every time they look into a new type of sample. Computer vision with machine learning algorithms paved the way for an easy and accurate way of diagnosing the mammogram accurately. In the area of biomedical images, Convolution neural networks get significant results. On the same path, there is considerable improvement in mammography image classification. In this work, Bi-LSTM based feature engineering approach is proposed integrating computer vision methods with machine learning base Random forest classifier used in the proposed method by this motivation.

VI. REFERENCES

[1]. Lee, J. G., Jun, S., Cho, Y. W., Lee, H., Kim, G. B., Seo, J. B., & Kim, N. (2017). Deep learning in medical imaging: general overview. *Korean journal of radiology*, *18*(4), 570-584.

[2]. Kooi, T., & Karssemeijer, N. (2017). Classifying symmetrical differences and temporal change for the detection of malignant masses in mammography using deep neural networks. *Journal of Medical Imaging*, *4*(4), 044501.

[3]. Suzuki, K. (2017). Overview of deep learning in medical imaging. *Radiological physics and technology*, *10*(3), 257-273.

[4]. Huang, Q., Luo, Y., & Zhang, Q. (2017). Breast ultrasound image segmentation: a survey. *International journal of computer assisted radiology and surgery*, *12*(3), 493-507.

[5]. Carneiro, G., Nascimento, J., & Bradley, A. P. (2017). Automated analysis of unregistered multi-view mammograms with deep learning. *IEEE transactions on medical imaging*, *36*(11), 2355-2365.

[6]. Shen, D., Wu, G., & Suk, H. I. (2017). Deep learning in medical image analysis. *Annual review of biomedical engineering*, *19*, 221-248.

[7]. Dhungel, N., Carneiro, G., & Bradley, A. P. (2017). A deep learning approach for the analysis of masses in mammograms with minimal user intervention. *Medical image analysis*, *37*, 114-128.

[8]. Zhang, X., Zhang, Y., Han, E. Y., Jacobs, N., Han, Q., Wang, X., & Liu, J. (2017, November). Whole mammogram image classification with convolutional neural networks. In *2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)* (pp. 700-704). IEEE.

[9]. Lu, L., Zheng, Y., Carneiro, G., & Yang, L. (2017). Deep learning and convolutional neural networks for medical image computing. *Advances in Computer Vision and Pattern Recognition*.

[10]. Kooi, T., Litjens, G., Van Ginneken, B., Gubern-Mérida, A., Sánchez, C. I., Mann, R., & Karssemeijer, N. (2017). Large scale deep learning for computer aided detection of mammographic lesions. *Medical image analysis*, *35*, 303-312.

[11]. Avalos-Rivera, E. D., & Pastrana-Palma, A. D. J. (2016). Classifying microcalcifications on digital mammography using morphological descriptors and artificial neural network. In *IEEE CACIDI 2016-IEEE Conference on Computer Sciences* (pp. 1-4). IEEE.

[12]. Dhungel, N. (2016). *Automated detection, segmentation and classification of masses from mammograms using deep learning* (Doctoral dissertation).

[13]. Vikhe, P. S., & Thool, V. R. (2016). Mass detection in mammographic images using wavelet processing and adaptive threshold technique. *Journal of medical systems*, *40*(4), 82.

[14]. Gastounioti, A., Conant, E. F., & Kontos, D. (2016). Beyond breast density: a review on the advancing role of parenchymal texture analysis in breast cancer risk assessment. *Breast cancer research*, *18*(1), 91

[15]. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., & Rabinovich, A. (2015). Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1-9).

[16]. He, W., Hogg, P., Juette, A., Denton, E. R., & Zwiggelaar, R. (2015). Breast image pre-processing for mammographic tissue segmentation. *Computers in biology and medicine*, *67*, 61-73.