ROI Based Breast Cancer Detection Using Deep Learning

Kumari Neha[1], Prof. Prateek Gupta[2]
[1]Dept of CSE
[2]Professor, Dept of CSE
[1, 2] Shri Ram Institute of Science and Technology, Jabalpur, M.P.

Abstract- *The network attacks are increasing both in frequency and intensity with the rapid growth of internet of things (IoT) devices. Recently, denial of service (DoS) and distributed denial of service (DDoS) attacks are reported as the most frequent attacks in IoT networks. The traditional security solutions like firewalls, intrusion detection systems, etc., are unable to detect the complex DoS and DDoS attacks since most of them filter the normal and attack traffic based upon the static predefined rules. However, these solutions can become reliable and effective when integrated with artificial intelligence (AI) based techniques. During the last few years, deep learning models especially convolutional neural networks achieved high significance due to their outstanding performance in the text classification field. The potential of these convolutional neural network (CNN) models can be used to efficiently detect the complex DoS and DDoS and others attacks by converting the network traffic dataset into array of matrices. Therefore, in this work, we proposed a methodology to detect traffic data on IoT network and trained a state-of-the-art CNN model with different Activation Functions over the converted data. The proposed methodology accomplished very high accuracy for Binary Classification on large dataset and good results for multiclassification on the same dataset for detecting the DoS, DDoS and other attacks in case of binary classification. Furthermore, the proposed methodology achieved 87% average precision for recognizing various types of DoS and DDoS attack patterns which is higher as compared to the state-of-the-art.*

Keywords- Attacks on IoT Networks, DoS, DDoS, Normal, Machine Learning, Deep Learning, CNN, activation Functions.

I. INTRODUCTION

The Breast cancer is considered as one of the main health concerns amongst women around the world. Prevalence of breast cancer has increased lately. Recent statistics [1] show that breast cancer is one of the top causes of malignant-related deaths amongst women at the age of 40-70 years. Therefore, it is crucial to best utilize imaging systems and laboratory analysis in order to accurately identify breast cancer. When a breast cancer is detected early, there is a higher chance for proper treatment. Hence, it is highly important that new effective methods are introduced and developed to detect and investigate breast cancer in early stages. Accurate identification of breast cancer tissues in early stages leads to increase survival rate and save many people live. In the first stage of diagnosis, a very critical procedure in image analysis is sampling that is present in multiple implementations, like format identification, object discovery, and categorization in medical imaging. Image sampling has many methods in order to enhance the quality, and it has multiple uses categorized by an assortment of functions concerned by the intensity of pixels. In such method, distinct objects have distinct oscillation disseminations that they are distinguished according to the regular and average divergence of each dissemination.

Experts in modern medical areas are focusing more on technical approaches for a variety of chronic diseases. Even though many diseases are incurable, such as cancer, stroke, heart attack, chronic liver diseases, viral hepatitis, and coronary artery disease, the death rate from breast cancer is increasing every year. According to a statistical report on medical health, cancer is a genetic disease that leads to variations in genes involved in the functionality of human body cells. Variation of the gene in genetic diseases may affect the internal parts of human organs for future generations. It may also affect DNA structure, resulting in environmental exposure to substances such as UV radiation, smoking, and other variables that are significant in the development of breast cancer [2]. Despite this, 60% of women affected by breast cancer are diagnosed at the last stage, which leads to death in women.

Global cancer data show that breast cancer is the second most lethal form of cancer worldwide after lung cancer [3]. In 2018, 2 billion new cases of breast cancer were reported worldwide, where 627,000 deaths. A study in Australia [4] showed that breast cancer survival is strongly associated with the size of the tumor at the time of detection, with the size less than 10 mm, the probability of patient survival is 98%. A cohort study showed that 70% of breast cancer cases are detected when the tumor size was 30 mm [5]. Breast cancer usually becomes detectable during screening when the tumor is at least 20 mm in size [6]. Therefore, enabling early detection of breast cancer is crucial to facilitate early treatment. Early treatment may be beneficial following identification through screening examinations such as clinical-breast examination (CBE) and breast self-examination (BSE). CBE is a regular medical examination performed by healthcare professionals to detect breast lesions, whereas BSE is conducted by an individual to observe physical changes and appearance of breasts. The practice of BSE empowers women to take responsibility for their health. Consequently, BSE is recommended by the World Health Organization for raising awareness among women at risk [7].

1.2 Breast Cancer types

Breast cancer is a serious threat to women's life and health, and the morbidity and mortality of breast cancer are ranked first and second out of all female diseases [8]. Early detection of lumps can effectively reduce the mortality rate of breast cancer [9]. The mammogram is widely used in early screening of breast cancer due to its relatively low expense and high sensitivity to minor lesions [10]. In the actual diagnosis process, however, the accuracy can be negatively affected by many factors, such as radiologist fatigue and distraction, the complexity of the breast structure, and the subtle characteristics of the early-stage disease [11], [12]. The computer-aided diagnosis (CAD) for breast cancer can help address this issue.

The classical CAD for breast cancer contains three steps:

(a) Finding the Region of Interest (ROI) in the preprocessed mammogram, and hence locating the region of the tumor.
(b) Then, extracting features of the tumor based on expert knowledge, such as shape, texture, and density, to manually generate feature vectors.
(c) Finally, diagnosing benign and malignant tumors by classifying these feature vectors [13], [14]. Although the classical diagnosis method has been commonly used, its accuracy still needs to be improved [15].

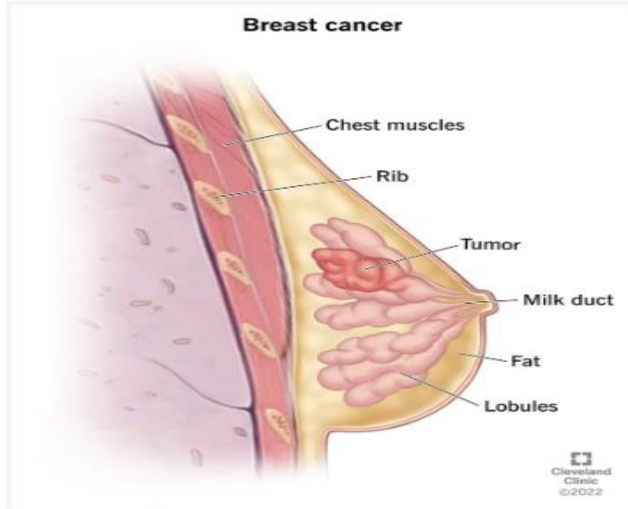

Figure 1.1: Structure of Breast.

Breast cancer is the most common cancers in women. It usually affects women over the age of 50. Although rare, men can also get cancer. About 2,600 men are diagnosed with prostate cancer each year in the United States, accounting for less than 1 percent of all cases. Tran's women are more likely to develop breast cancer than cisgender men. Additionally, transgender men are less likely to develop breast cancer than cisgender women.

There are several types of breast cancer, including:

- infiltrating (invasive) ductal carcinoma. These tumors begin in breast milk, extend beyond the walls of the ducts and spread to the breast tissue. This is the most common type of cancer, accounting for 80% of all cases.
- Ductal carcinoma in situ. Ductal carcinoma in situ, also known as stage 0 breast cancer, is considered by some to be the first because the cells do not spread into breast milk. These conditions are treatable. But urgent care is needed to prevent cancer cells from entering and spreading into other tissues.
- Lobular carcinoma in situ is a precancerous condition involving abnormal cells in the breast lobules. It is not actually cancer, but this marker may indicate the possibility of developing cancer later in life. Therefore, it is important for women with lcis to have regular breast exams and mammograms.
- Breast cancer. This type of cancer is rare and serious, similar to the disease. People with breast cancer often notice redness, swelling, pitting, and dimpling of the breast skin. It is caused by obstructive cancer cells in the lymphatic vessels.
- Paget disease of the breast. This cancer affects the skin of the nipple and areola (skin around the nipple).

In recent years, deep learning methods, such as the convolution neural network (CNN), that can extract hierarchical features from image data without the manual selection, which is also called objective features, have been successfully applied with a great improvement on accuracies in many applications, such as image recognition, speech recognition, and natural language processing [16], [17].

Breast cancer is a kind of malignant growth starting with breast tissue, usually in the interior lining of the Breast Lobules or Milk Ducts and metastasizing to other body parts [18]. The second most widespread disease worldwide is breast cancer. Therefore, it is crucial to check the number of deaths due to breast cancer prematurely in successful treatment and the drop. Research related to breast cancer has increased in the last decade [19]. Breast cancer medical imaging can be used to look inside the human body as a non-invasive method for helping doctors for diagnose and treat [20], [21]. An early breast cancer diagnosis can occur with any of the available imaging methods; it cannot be confirmed that these images are malignant alone [22]. There is a high risk of cancer cells being placed in the interstitial tissue veins or fluid until the microscopic exam of tissues from cancer to confirm their malignancy begins [23]. There is a possibility that cells drag along an operative incision or needle route that can increase the spread of cancer through biopsy [24]. The mammographic breast image is typically preprocessed to eliminate pectoral muscle in the diagnosis of breast cancer with a mammogram to encircle the detection process.

II. LITERATURE REVIEW

Developing a breast cancer screening method is very important to facilitate early breast cancer detection and treatment. Building a screening method using medical imaging modality that does not cause body tissue damage (non-invasive) and does not involve physical touch is challenging. Thermography, a non-invasive and non-contact cancer screening method, can detect tumors at an early stage even under precancerous conditions by observing temperature distribution in both breasts. The thermo grams obtained on Thermography can be interpreted using deep learning models such as convolutional neural networks (CNNs). CNNs can automatically classify breast thermo grams into categories such as normal and abnormal. Despite their demonstrated utility, CNNs have not been widely used in breast thermo gram classification. In this study, we aimed to summarize the current work and progress in breast cancer detection based on Thermography and CNNs. We first discuss of breast Thermography potential in early breast cancer detection, providing an overview of the availability of breast thermal datasets together with publicly accessible. We also discuss characteristics of breast thermo grams and the differences between healthy and cancerous Thermography patterns. Breast thermo gram classification using a CNN model is described step by step including a simulation example illustrating feature learning. We cover most research related to the implementation of deep neural networks for breast thermogram classification and propose future research directions for developing representative datasets, feeding the segmented image, assigning a good kernel, and building a lightweight CNN model to improve CNN performance [25].

An ultra wideband (UWB) radar-based breast cancer detection system, which is composed of complementary metal–oxide–semiconductor integrated circuits, is presented. This system includes Gaussian monocycle pulse (GMP) generation circuits, switching (SW) matrix circuits, equivalent-time sampling circuits, and a compact UWB antenna array. During the detection process, the GMP signal with the center frequency of 6 GHz is first generated and transmitted with a repetition frequency of 100 MHz. The GMP signal is sent to a selected transmitter antenna by the SW matrix module, and the reflected signal is captured by the receiver antennas. Next, the high-speed equivalent-time sampling circuits are employed to retrieve the reflected GMP signal. A confocal algorithm is used to reconstruct the breast image. The total size for the prototype module is 45 cm×30 cm×14.5 cm in length, width, and height, respectively, which is dramatically smaller than the conventional detection systems. Using our proposed system, we demonstrate a successful detection of 1-cm cancer target in the breast phantom [26].

Breast cancer plays a significant role in affecting female mortality. Researchers are actively seeking to develop early detection methods of breast cancer. Several technologies contributed to the reduction in mortality rate from this disease, but early detection contributes most to preventing disease spread, breast amputation and death. Thermography is a promising technology for early diagnosis where thermal cameras employed are of high resolution and sensitivity. The combination of Artificial Intelligence (AI) with thermal images is an effective tool to detect early stage breast cancer and is foreseen to provide impressive predictability levels. This paper reviews systematically the related works employing Thermography with AI highlighting their contributions and drawbacks and proposing open issues for research. Several different types of Artificial Neural Networks (ANNs) and deep learning models were used in the literature to process thermographic images of breast cancer, such as Radial Basis Function Network (RBFN), K-Nearest Neighbors (KNN), Probability Neural Network (PNN), Support Vector Machine (SVM), ResNet50, SeResNet50, V Net, Bayes Net, Convolutional Neural Networks (CNN), Convolutional and Deconvolutional Neural Networks (C-DCNN), VGG-16, Hybrid (ResNet-50 and V-Net), ResNet101, DenseNet and InceptionV3. Previous studies were found limited to varying the numbers of thermal images used mostly from DMR-IR database. In addition, analysis of the literature indicate that several factors do affect the performance of the Neural Network used, such as Database, optimization method, Network model and extracted features. However, due to small sample size used, most of the studies achieved a classification accuracy of 80% to 100% [27].

A computer-aided diagnosis (CAD) system based on mammograms enables early breast cancer detection, diagnosis, and treatment. However, the accuracy of the existing CAD systems remains unsatisfactory. This paper explores a breast CAD method based on feature fusion with convolutional neural network (CNN) deep features. First, we propose a mass detection method based on CNN deep features and unsupervised extreme learning machine (ELM) clustering. Second, we build a feature set fusing deep features, morphological features, texture features, and density features. Third, an ELM classifier is developed using the fused feature set to classify benign and malignant breast masses. Extensive experiments demonstrate the accuracy and efficiency of our proposed mass detection and breast cancer classification method [28].

Neural networks have recently become a popular tool in cancer data classification. In this paper, Deep Learning assisted Efficient Adaboost Algorithm (DLA-EABA) for breast cancer detection has been mathematically proposed with advanced computational techniques. In addition to traditional computer vision approaches, tumor classification methods using transfers are being actively developed through the use of deep convolutional neural networks (CNNs). This study starts with examining the CNN based transfer learning to characterize breast masses for different diagnostic, predictive tasks or prognostic or in several imaging modalities, such as Magnetic Resonance Imaging (MRI), Ultrasound (US), digital breast tomosynthesis and mammography. The deep learning framework contains several convolutional layers, LSTM, Max-pooling layers. The classification and error estimation that has been included in a fully connected layer and a softmax layer. This paper focuses on combining these machine learning approaches with the methods of selecting features and extracting them through evaluating their output using classification and segmentation techniques to find the most appropriate approach. The experimental results show that the high accuracy level of 97.2%, Sensitivity 98.3%, and Specificity 96.5% has been compared to other existing systems [29].

III. PROPOSED WORK

In this section, the proposed architecture phases to detect lesions in mammograms and classify them are discussed, with the evaluation method applied to check for cancerous tumors in new mammograms. As shown in Figure 4.1, the proposed architecture is composed of three phases. The first phase is the preprocessing phase that prepares the Digital Imaging and Communications in Medicine (DICOM) mammograms to be in the required format without any extra artifacts. The second phase is the detection phase responsible for configuring the model to localize masses in the abnormal mammograms through two detection paths. Finally, the third phase which is the classification phase applied by replacing with classification layers with other feature extractors to distinguish between benign or malignant detected lesions. Figure 4.1 illustrates the overall architecture of the detection model.

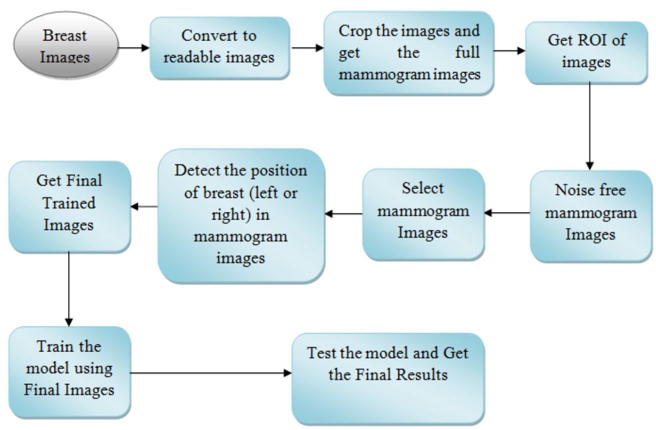

Figure 4.1: Proposed Model Working Architecture.

4.2 Preprocessing Phase

The mammograms are usually stored in form of Digital Imaging and Communications in Medicine (DICOM) format in large dimensions greater than those that fit in any deep learning method. Besides that, their ground truth annotations are not stored in a direct format. So, through this phase, the preprocessing steps shown in Figure 4.1 are applied to prepare the mammographic datasets in the required format for checking. The preprocessing step is considered an essential phase for both types of mammograms, either the scanned mammograms or Full Field. The scanned mammograms are in a higher need of the preprocessing steps than Full Field.

CADs obtain better results when they have applied on the full-field mammograms than the scanned film ones but generally more accurate results are obtained when the preprocessing steps are applied on either one of them. Since even in the case of the Full field, sometimes labels exist at the top of some cases and even if the mammogram is free from any extra artifacts, some noise results in the obtained Full field due to some dust obtained on the breast screening tool itself or any useless motion. The first step as shown in Figure 4.1 comes here to prepare the mammograms in the required format with their ground truth.

In the beginning, all DICOM files are converted into readable images in TIFF (Tagged Image File Format) and (Portable Graphics Format) PNG formats. Since the mammograms' pixels' values are usually fit between 14-bits and 16-bits contrast resolution, so in our work, the PNG is used but by scaling the 14-bits or the 16-bits resolution. Then, any artifacts like the extra labels, mammograms view, etc. are removed by extracting the greatest component (breast) to get a noise-free mammogram.

Moreover, in order to reduce any useless space in the mammogram leaving the breast only, some mammograms contain large black space. So, in a trial to resize the mammogram naturally without losing any data, continuous black spaces in a mammogram are removed to reduce the original size to fit the breast only, and consequently, the coordinates of the lesion are updated related to the new mammogram.

It is concluded from the most recent work proposed for breast cancer detection that the full mammograms due to their large sizes, they are resized to small dimensions to fit the one of the deep learning voracious method [32]. The resizing issue is mainly summarized in a reduction in the image quality and the loss of some significant information that may feature the existent masses especially when the resizing occurred by a large ratio. The complete working of the proposed model is shown below in figure 4.2.

Deep Convolutional network performs the convolution of image data and specified Kernel in Euclidean space. Output after each of the convolutional operations depends on the considered stride, padding, image size and Kernel size. But TL is the most adapted and time efficient technique used to train the base model by transferring weights from the network trained on large size ImageNet dataset.

IV. RESULTS

Table 4.1 below shows the comparison of our model with some existing models on Histopathology dataset in terms of accuracy.

Algorithm Implemented	**Accuracy on Histopathology dataset in %**
Random Forest.	88.05
K-Means with GMM.	74.35
SVM	66.55
ResNet and Inception	86.45
CNN	89.82
Our Model	95.41

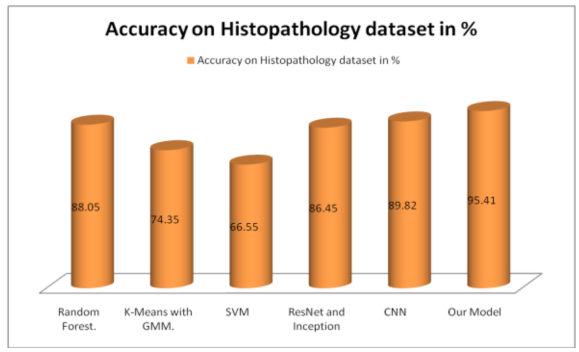

Figure 4.1: Comparison Chart for Accuracy.

V. CONCLUSION

In this paper, we have proposed a ROI-based Breast Cancer detection system to localize any suspected cancerous area in the breast and classify them into benign or malignant if existing with high accuracy. The proposed model overcomes the issue of resizing the breast mammograms into smaller sizes to fit CNNs for cancer localization by proposing a mammograms cropping idea that acts as a third part detector.

Nowadays, due to the huge number of mammograms taken daily and with the necessity to early discover breast cancer to reduce the death rate, the model play an essential role to be a second reader for the screened mammograms to decide with the radiologists one. The proposed model is a real-time detector that can detect and classify any existing masses. Moreover using different experiments, it is proved that it is better to use ROI-Based Model for detecting only cancerous areas whatever their types. The Proposed model outperforms the existing model in terms of Accuracy.

REFERENCES

[1] C. E. DeSantis, J. Ma, A. G. Sauer, L. A. Newman, and A. Jemal, ''Breast cancer statistics, 2017, racial disparity

in mortality by state,'' CA, A Cancer J. Clinicians, vol. 67, no. 6, pp. 439–448, 2017.

[2] S. Aminikhanghahi, S. Shin, W. Wang, S. I. Jeon, and S. H. Son, ''A new fuzzy Gaussian mixture model (FGMM) based algorithm for mammography tumor image classification,'' Multimedia Tools Appl., vol. 76, no. 7, pp. 10191–10205, Apr. 2017.

[3] IAFR Cancer. Global Cancer Observatory. Accessed: Jun. 30, 2019. [Online]. Available: http://gco.iarc.fr/.

[4] C. Nickson and A. M. Kavanagh, ''Tumor size at detection according to different measures of mammographic breast density,'' J. Med. Screening, vol. 16, no. 3, pp. 140–146, Sep. 2009.

[5] S. A. Narod, ''Tumors size predicts long-term survival among women with lymph node-positive breast cancer,'' Current Oncol., vol. 19, no. 5, pp. 249–253, Sep. 2012.

[6] OncoLink Team. (Jan. 2020). All about Breast Cancer. [Online]. Available: https://www.oncolink.org/cancers/breast/all-about-breast-cancer.

[7] Breast Cancer: Prevention and Control, World Health Org., Geneva, Switzerland, 2019.

[8] R. L. Siegel et al., ''Colorectal cancer statistics,'' CA, Cancer J. Clinicians, vol. 64, no. 2, pp. 104–117, Mar. 2014

[9] J. B. Harford, ''Breast-cancer early detection in low-income and middle-income countries: Do what you can versus one size fits all,'' Lancet Oncol., vol. 12, no. 3, pp. 306–312, Mar. 2011.

[10] C. Lerman et al., ''Mammography adherence and psychological distress among women at risk for breast cancer,'' J. Nat. Cancer Inst., vol. 85, no. 13, pp. 1074–1080, Jul. 1993.

[11] P. T. Huynh, A. M. Jarolimek, and S. Daye, "The false-negative mammogram,'' Radiographic, vol. 18, no. 5, pp. 1137–1154, Sep. 1998.

[12] M. G. Ertosun and D. L. Rubin, ''Probabilistic visual search for masses within mammography images using deep learning,'' in Proc. IEEE Int. Conf. BioInform. Biomed. (BIBM), Nov. 2015, pp. 1310–1315.

[13] S. D. Tzikopoulos, M. E. Mavroforakis, H. V. Georgiou, N. Dimitropoulos, and S. Theodoridis, ''A fully automated scheme for mammographic segmentation and classification based on breast density and asymmetry,'' Comput. Methods Programs Biomed. vol. 102, no. 1, pp. 47–63, 2011.

[14] D. C. Pereira, R. P. Ramos, and M. Z. do Nascimento, "Segmentation and detection of breast cancer in mammograms combining wavelet analysis and genetic algorithm," Comput. Methods Programs Biomed. vol. 114, no. 1, pp. 88–101, Apr. 2014.

[15] S. A. Taghanaki, J. Kawahara, B. Miles, and G. Hamarneh, ''Paretooptimal multi-objective dimensionality reduction deep auto-encoder for mammography classification,'' Comput. Methods Programs Biomed. vol. 145, pp. 85–93, Jul. 2017.

[16] X.-W. Chen and X. Lin, ''Big data deep learning: Challenges and perspectives,'' IEEE Access, vol. 2, pp. 514–525, 2014.

[17] K. Ganesan, U. R. Acharya, C. K. Chua, L. C. Min, K. T. Abraham, and K. H. Ng, ''Computer-aided breast cancer detection using mammograms: A review,'' IEEE Rev. Biomed. Eng., vol. 6, pp. 77–98, 2012.

[18] P. Gupta and A. Kaur Malhi, ''Using deep learning to enhance head and neck cancer diagnosis and classification,'' in Proc. IEEE Int. Conf. Syst., Compute. Autom. Netw. New York, NY, USA, Jul. 2018, pp. 1–6.

[19] A. M. Abdel-Zaher and A. M. Eldeib, ''Breast cancer classification using deep belief networks,'' Expert Syst. Appl., vol. 46, pp. 139–144, Mar. 2016.

[20] A. Cruz-Roa, H. Gilmore, A. Basavanhally, M. Feldman, S. Ganesan, N. N. C. Shih, J. Tomaszewski, F. A. González, and A. Madabhushi, ''Accurate and reproducible invasive breast cancer detection in whole-slide images: A deep learning approach for quantifying tumor extent,'' Sci. Rep., vol. 7, no. 1, Jun. 2017, Art. no. 46450.

[21] S. Mambou, P. Maresova, O. Krejcar, A. Selamat, and K. Kuca, ''Breast cancer detection using infrared thermal imaging and a deep learning model,'' Sensors, vol. 18, no. 9, p. 2799.

[22] M. H. Yap, G. Pons, J. Marti, S. Ganau, M. Sentis, R. Zwiggelaar, A. K. Davison, and R. Marti, ''Automated breast ultrasound lesions detection using convolutional neural networks,'' IEEE J. Biomed. Health Informant. vol. 22, no. 4, pp. 1218–1226, Jul. 2018.

[23] D. Levy and A. Jain, ''Breast mass classification from mammograms using deep convolutional neural networks,'' (2016), arXiv: 1612.00542. [Online]. Available: https://arxiv.org/abs/1612.00542.

[24] S. Suzuki, X. Zhang, N. Homma, K. Ichiji, N. Sugita, Y. Kawasumi, T. Ishibashi, and M. Yoshizawa, ''Mass detection using deep convolutional neural network for mammographic computer-aided diagnosis,'' in Proc. 55th Annu. Conf. Soc. Instrum. Control Eng. Jpn. (SICE), Sep. 2016, pp. 1382–1386.

[25] R. Roslidar *et al.*, "A Review on Recent Progress in Thermal Imaging and Deep Learning Approaches for Breast Cancer Detection," in *IEEE Access*, vol. 8, pp. 116176-116194, 2020, doi: 10.1109/ACCESS.2020.3004056.

[26] H. Song *et al*., "A radar-based breast cancer detection system using CMOS integrated circuits," in IEEE Access,

vol. 3, pp. 2111-2121, 2015, doi: 10.1109/ACCESS.2015.2496101

[27] M. A. S. A. Husaini, M. H. Habaebi, S. A. Hameed, M. R. Islam and T. S. Gunawan, "A Systematic Review of Breast Cancer Detection Using Thermography and Neural Networks," in IEEE Access, vol. 8, pp. 208922-208937, 2020, doi: 10.1109/ACCESS.2020.3038817.

[28] Z. Wang *et al*., "Breast Cancer Detection Using Extreme Learning Machine Based on Feature Fusion With CNN Deep Features," in IEEE Access, vol. 7, pp. 105146-105158, 2019, doi: 10.1109/ACCESS.2019.2892795.

[29] J. Zheng, D. Lin, Z. GAO, S. Wang, M. He and J. Fan, "Deep Learning Assisted Efficient AdaBoost Algorithm for Breast Cancer Detection and Early Diagnosis," in IEEE *Access*, vol. 8, pp. 96946-96954, 2020, doi: 10.1109/Access.2020.2993536.