Progress In Electromagnetics Research C, Vol. 128, 39–48, 2023

Deep Learning Algorithm for Automatic Breast Tumor Detection and Classification from Electromagnetic Scattering Data

Lalitha Kandasamy* **and Shreya Reddy K**

Abstract—Breast cancer is, by far, the most diagnosed disease for the death of women worldwide. Researchers are working with an alternative technology to detect tumors before they reach a terrible stage because of the numerous limitations in the current imaging approach. This article suggests a promising technique by utilising non-ionizing microwave signal and artificial intelligence especially deep learning algorithms for early detection of breast cancer. This contribution will present a method to detect and classify the tumor category using backscatter signals obtained from antenna simulation in CST microwave studio software. The post-processed scattering parameters are utilized to create image through MATLAB programming environment. The high intensity in the image represents the precise position of tumor. The automatic classification of tumor is achieved by YOLOv_5 deep learning model from the created microwave images. A training dataset with fifty image samples are formed by preprocessing, and then augmentation is applied to create final dataset with 1000 samples. This approach can identify the location and type of early-stage tumor with size of 5 mm.

1. INTRODUCTION

In 2025, breast cancer is anticipated to be the most common disease among women around the world [1, 2]. Adipose tissue and fibro glands are the two main healthy tissues of importance in the breast. The fibrous and glandular tissues are together referred to as "fibro glandular tissues." Adipose and fibro glands tissue proportions in the breast vary with age, general nutrition status, and hormonal changes brought on by menstruation or pregnancy. Breast tumors typically develop in the milk ducts connecting to the nipple or in the glandular tissue of the breast. Breast cancer is characterized by abnormal breast cell proliferation. The dielectric contrast between cancerous tumors and adipose breast tissue makes this technology particularly promising for the study of breast cancer. The substantial disparity in electrical characteristics (relative permittivity and conductivity) between malignant and normal tissues at microwave frequencies is the significant driver behind microwave imaging efficiency in tumor diagnostics. The discrepancy in dielectric constant values between different tissue types is due to the water content of each. High-water-content tissues (malignant tumors) have higher relative dielectric permittivity and conductivity, whereas fat, which is predominant in normal breast tissue, has a lower permittivity.

The gold standard procedure for breast abnormality detection is digital mammography. The breast is exposed to ionising x-rays during a mammogram, and the difference in X-ray attenuation of various tissues is used to visualize breast tissues [3]. Ionizing radiation exposure is a permanent health risk. Ultrasound is another technique for breast cancer screening. While ultrasound has a strong specificity for detecting the presence of cancerous lesions, accuracy depends on the operator and may have a higher percentage of false-positive results than mammography [4, 5]. So, an alternative technique

Received 6 November 2022, Accepted 20 December 2022, Scheduled 26 December 2022
* Corresponding author: Lalitha Kandasamy (lalithak@srmist.edu.in).
The authors are with the Department of Electronics and Communication Engineering, SRM Institute of Science and Technology, Ramapuram Campus, Chennai, India.

is required for early detection and continuous monitoring of the disease progression. For a clinical diagnosis of breast cancer, it is vital to distinguish between benign and malignant tumors. When a benign lump is misdiagnosed, unnecessary invasive procedures (such as a biopsy) may be performed. Breast microwave sensing systems work by exposing breast tissues with low power microwaves that typically operates between 900 MHz and 10 GHz with the help of an antenna. The printed microstrip antennas are extensively utilized in many noninvasive applications. Microstrip antennas with different types of patch structures and various shapes of ground plane are introduced to improve the bandwidth and reflection coefficient [6]. Here, parasitic patches in the form of cross shapes are placed in the substrate to improve the antenna characteristics. Breast tissues produce a scattered field because of their interactions with incident fields, and the image reconstruction from scattering parameters is done in MATLAB 2020 using time domain algorithms with dimension 640×640 pixels [7]. One of the intelligent classifications is based on fuzzy rules with self-configuring evolutionary genetic algorithm. The design of fuzzy classifier and its implementation process is complex and needs deep analysis of fuzzy set theory. A novel implementation fuzzy classifier is done by automated self-configuring fuzzy classifier. Fuzzy set theory and machine learning algorithms together proposed an approach to forecast the possibility of heart disease in an intelligent way [8]. Some researchers have suggested a deep learning-based approach for automatic tumor detection in recent years as deep learning has gained popularity [9]. However, the deep learning-based techniques for detecting breast cancer are essentially restricted to using image segmentation to find region of interest (ROI) regions or categorise a given ROI region as benign or malignant. They are unable to simultaneously find ROI regions and classify tumors, which leads to inefficient diagnosis. On the other hand, convolutional neural network (CNN) architectures with target detection-based algorithms are employed to classify electromagnetic images, and the generated images look noisy because of unavailability of large dataset. In this paper, we try to classify breast tumor and to find its location using YOLOv$_5$ algorithm. The simulation model creates fifty sample images, and image augmentation process artificially generates the final dataset with thousand images by rotation, scaling, cropping, zooming, shifting, etc. YOLOv$_5$ model is implemented in python using TensorFlow API. Six hundred images from the final dataset are utilized for training the YOLOv$_5$ model, and two hundred images are taken for validation purpose. The remaining two hundred of generated images are made available for testing purposes.

2. METHODS

2.1. Design and Modelling of Microwave Wideband Antenna

CST Microwave Studio is an antenna modelling and simulation tool for various applications. In this paper, a CST microwave studio 2020 is used to build a microstrip patch antenna (96.45 mm × 82.98 mm) with a cross parasitic patch to perform wideband microwave imaging of breast tissue. The difference in dielectric properties of normal and tumor tissues is identified by radar-based microwave imaging setup. The parameters required to model the physical dimensions of the proposed antenna (substrate's height and width) are designed according microstrip antenna design equations from (1) to (6). The simulated antenna exhibits good gain at resonance frequency of 3.57 GHz. Figure 1 and Figure 2 depict the antenna's physical composition. FR-4 substrate material with a 1.6 mm width is utilised for antenna modelling. The substrate's relative permittivity is chosen to be 4.3, with parasitic cross-shaped elements for microwave applications. The antenna reflection coefficient characteristic curve is shown in Figure 3.

Width of the transmission line

$$w = \frac{c}{2f_c\sqrt{\frac{(\varepsilon_r+1)}{2}}} \tag{1}$$

where fc = operating frequency, C = velocity of light in free space.

The effective dielectric constant ε_{reff} is given by,

$$\varepsilon_{re_{ff}} = \frac{\varepsilon_r+1}{2} + \frac{\varepsilon_r-1}{2}\left(1+12\frac{h}{w}\right)^{-1/2} \tag{2}$$

$$\textit{Effective length} = L_{eff} = \frac{c}{2f\sqrt{\varepsilon_{reff}}} \tag{3}$$

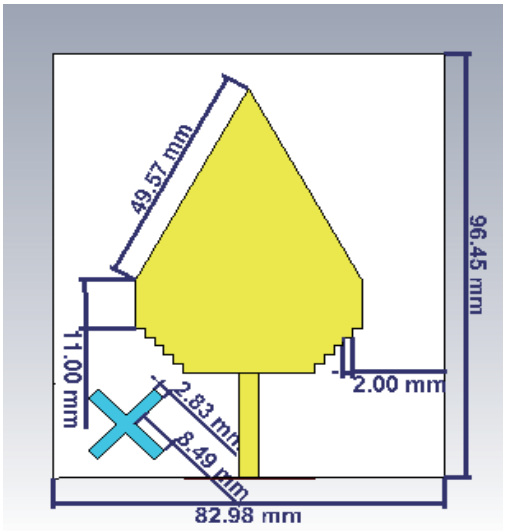

Figure 1. Proposed antenna structure front view.

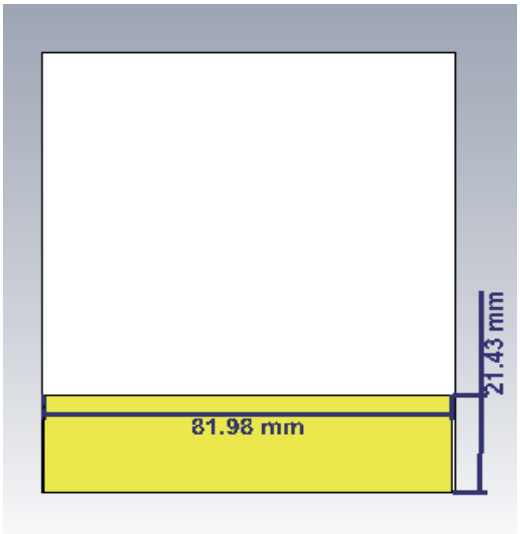

Figure 2. The antenna structure back view.

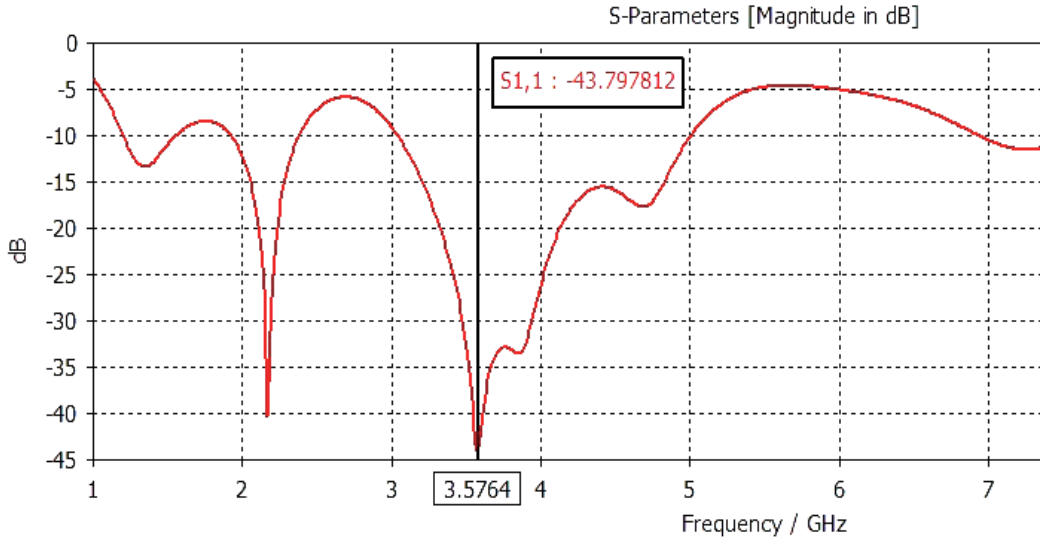

Figure 3. Reflection coefficient characteristics of proposed antenna.

$$Fringing\ length = \Delta L = 0.412h\frac{(\Sigma_{reff}+0.3)\left(\frac{w}{h}+0\cdot 264\right)}{(\varepsilon_{reff}-0\cdot 258)\left(\frac{w}{h}-0\cdot 8\right)} \tag{4}$$

$$Actual\ length = L = L_{eff}-2\Delta L \tag{5}$$

$$Length\ of\ the\ ground = L_g = 2*L \tag{6}$$

$$Width\ of\ the\ ground = W_g = 2*W$$

2.2. Dielectric Characteristics of Breast Tissue

Dielectric characteristics describe how electromagnetic waves interact at the cellular and molecular levels with biological tissues. An EM field predominantly affects parts of the materials found in biological tissues. These substances possess an electric net charge and/or an electric dipole moment. most

significant polar molecules serve as the tissue's primary source of electric dipole moments. Muscles, fat, and other body parts are some more sources. Since a tissue's electrical characteristics are influenced by a wide range of factors, as shown by various dielectric dispersions, these characteristics show large variances. The randomly oriented molecules in biological tissues will align in the direction of the applied electric field upon excitation. The molecules inside the tissues polarise because of this alignment caused by the external electric field. The molecules' polarisation produces an electric field that is opposite to the applied field's direction but larger in amplitude. This polarisation process takes place gradually over a period called as the relaxation time. The dielectric properties of different breast tissues are expressed in Table 1.

Table 1. Dielectric properties of breast tissue.

Sl. No.	Layer description	Dielectric constant	Electrical Conductivity (S/M) (σ)	Thickness in Diameter (mm)
1	Skin tissue	38	2.34	36
2	Fat tissue	4.8	0.26	32
3	Malignant Tumour tissue	**67**	**49**	**8**
4	Benign Tumour tissue	59	**34**	**5**

2.3. Simulation of Breast Phantom

The goal of the current study is to create an ultra-wideband (UWB) antenna and breast phantom model to diagnose breast cancer. The skin layer, breast tissue, and malignant or benign tumor tissue all make up the human breast. A simulation model is developed in CST microwave studio 2020 to mimic equivalent dielectric properties of breast tissue as shown in Figure 4. Two types of phantom models are

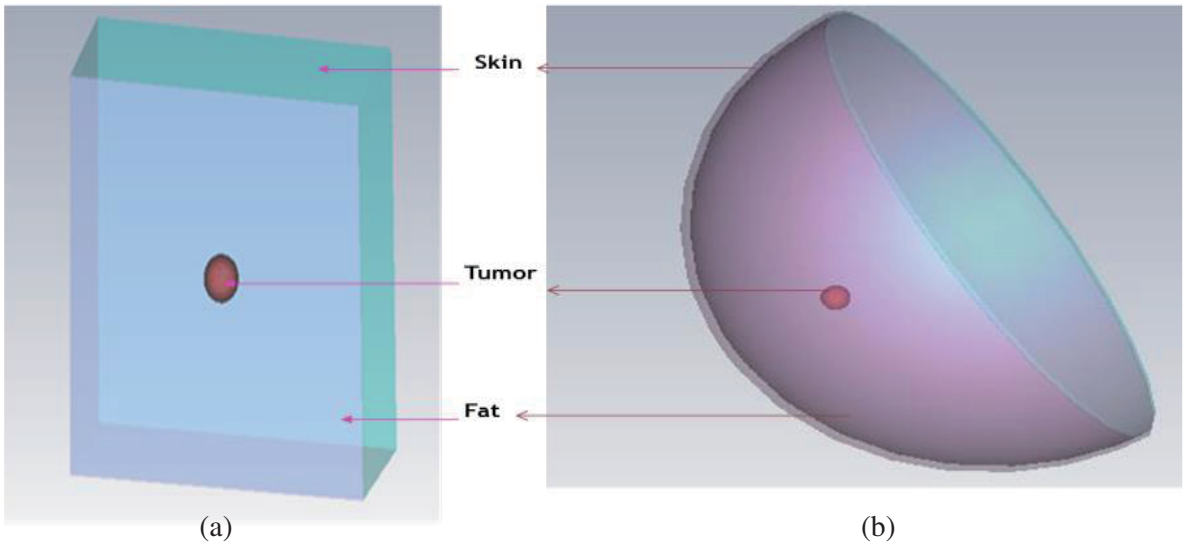

Figure 4. Rectangular and Hemispherical breast phantom. (a) Rectangular breast phantom, (b) hemispherical breast phantom.

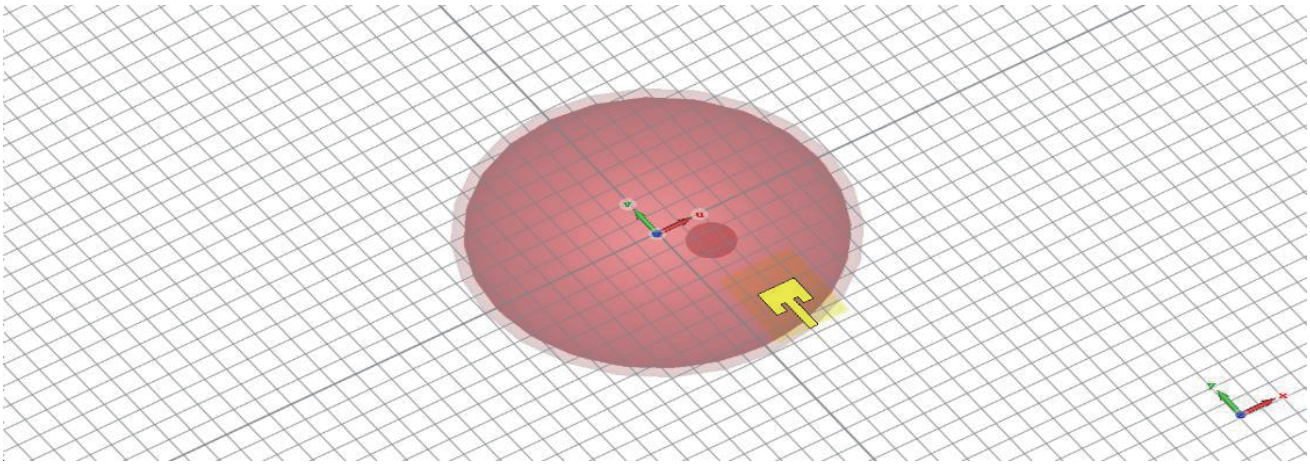

Figure 5. Simulation environment with antenna.

developed with rectangular and hemispherical shapes [10, 11]. The UWB antenna is directed to radiate low power microwave energy towards breast model, and the reflected signals are analysed to derive dielectric property distribution. Figure 5 denotes the arrangement of antenna and breast phantom in simulation environment. Different biological tissues with differences in water contents have unique electrical characteristics. This simulation setup makes it simple to distinguish between healthy tissue and cancerous tissue by revealing the presence of the tumor and its various stages of resolution. The antenna or sensing element is rotated around the breast phantom to measure reflection coefficient at different positions.

The interaction of electromagnetic waves with breast tissue is analysed by complex permittivity ε, and it represents the capability of material (tissue) to store energy. Equation (7) represents the complex permittivity of a human tissue and varies according to the frequency.

$$\varepsilon = \varepsilon_0 \left(\varepsilon_r' - j\varepsilon_r''\right) \tag{7}$$

where ε_r' and ε_r'' show the real and imaginary parts of complex permittivity, and $\varepsilon_0 = 8.854 \times 10^{-12}$. The relation between conductivity and complex permittivity is given by Equation (8)

$$\varepsilon_r'' = \frac{\sigma}{\omega} \tag{8}$$

where σ represents the conductivity, and ω indicates the angular velocity in radians.

3. RESULTS AND DISCUSSIONS

The result of the proposed technique to locate and image of the breast phantom is discussed in this section.

3.1. Image Formation by Delay and Sum Algorithm

It is a confocal imaging algorithm that involves a beamforming technique to remove strong artifacts and noise. Delay and sum (DAS) is a time-domain approach for calculating the propagation model of an antenna. Signal to clutter ratio is a metric like the signal to noise ratio in signal processing. But it fails to differentiate the against artifacts and noise as represented in Figure 6. Microwave imaging in space-time is the most used algorithm to improve the signal to clutter ratio.

The simple delay and sum Equation (9) are given below.

$$I\left(\vec{r}\right) = \left[\sum_{i=1}^{M} b_{\dot{I}}\left(\tau_i\left(\vec{r}\right)\right)\right]^2 \tag{9}$$

where M — total number of channels, bi — backscattered signal recorded at Channel i.

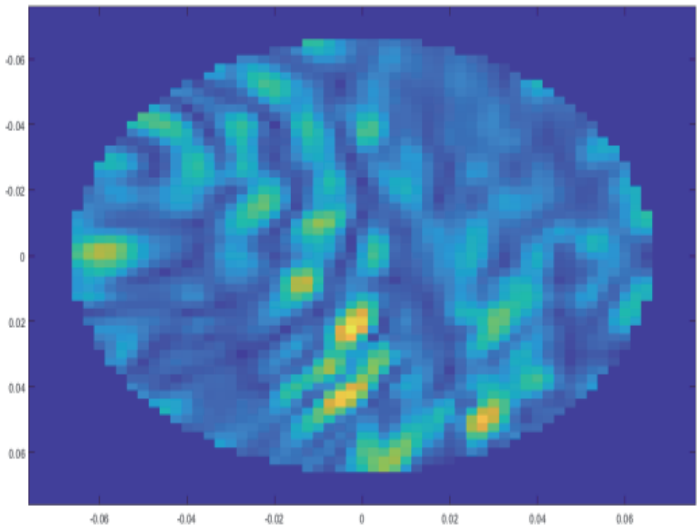

Figure 6. Reconstructed image in delay and sum algorithm.

3.2. Image Formation by Improved Delay and Sum Algorithm

An additional weighting factor is introduced at each focal point. The weighting factor is called Q factor, and it is introduced in the equation shown below. Figure 7 represents the reconstructed image.

$$I\left(\vec{r}\right) = QF\left(\vec{r}\right) \cdot \left[\sum_{i=1}^{M} b_{\dot{I}}\left(\tau_i\left(\vec{r}\right)\right)\right]^2 \tag{10}$$

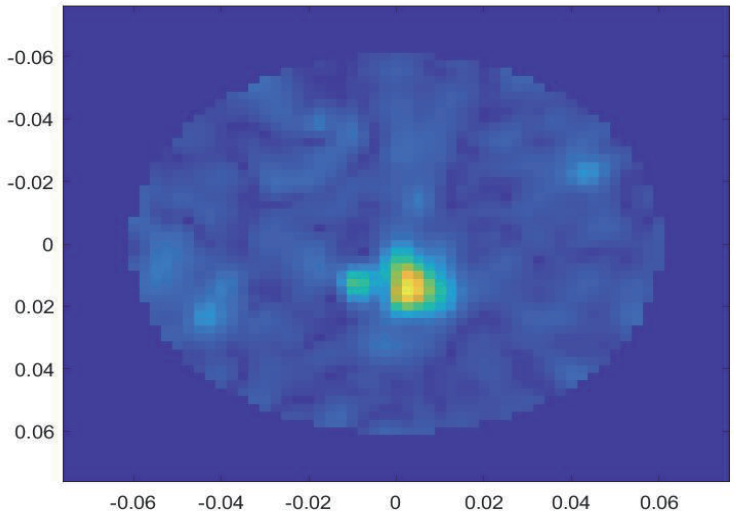

Figure 7. Reconstructed image in improved delay and sum algorithm.

3.3. Image Formation by Delay Multiply and Sum Algorithm

The delay multiply and sum (DMAS) beamformer multiplies the signals in pair after time alignment, and it provides a better signal to clutter ratio. The microwave image acquired using DMAS with tumor

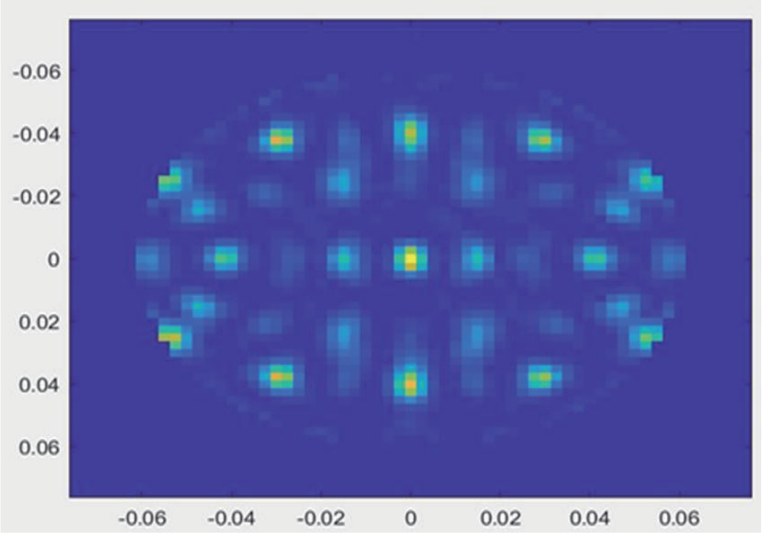

Figure 8. Reconstructed image in delay multiply and sum algorithm.

is shown in Figure 8. The highest yellow colour pixel value in the centre of the image shows the sharp image of the tumor. The bright yellow intensity at the centre of the image is due to the highest reflection from the breast. The remaining blue colour and low yellow colour intensity show the lowest scattering region from the breast.

Image clutter is a detection quality metric that is defined as the strongest tumor response to the strongest clutter energy in the breast tissue. Signal-to-clutter ratio (SCR) is given by $20\log_{10}(S\max/C\max)$, where $S\max$ is the maximum response in the known tumor region, and C max is the maximum response in the clutter region of the image as represented by Table 2.

Table 2. Quality metrics of image reconstruction.

Tumour size	Signal-to-clutter ratio (SCR)		
	Delay and Sum Algorithm (dB)	Improved Delay and Sum Algorithm (dB)	Delay Multiply and Sum Algorithm (dB)
5 mm	1.54	3	2
10 mm	1	4	2

One of the most cutting-edge real-time object identification algorithms was created by Joseph Redmon et al. [14] and is called YOLO (You only look once). The overall detection approach, dataset justification, image preprocessing, augmentation strategies, YOLOv$_5$ architecture with classification analysis, and detection mechanism are covered in this section [12, 13]. Figure 9 illustrates the full approach of the classification and detection flowchart. The simulated microwave breast imaging pictures were used in this work. The YOLOv$_5$ models are trained using the processed images and their accompanying tagging of tumor objects in YOLO format. The input image is divided into a grid of cells. Each cell in YOLO v5's deep convolutional neural network (DCNN) directly predicts a bounding box (BB) and object classification. The architecture of YOLOv$_5$ model is illustrated in Figure 10. There are three subsections namely backbone, neck, and prediction. Input images are preprocessed, labelled, and annotated to classify them into different categories. The dataset contains 1000 images including 50 image samples and augmented images. Initially, 75% of images present in the dataset are utilized for training, and remaining 20% are utilized for validation and testing. It is apparent that for the YOLOv$_5$ model to be trained effectively to detect and categorize the target tumor with locations in the microwave

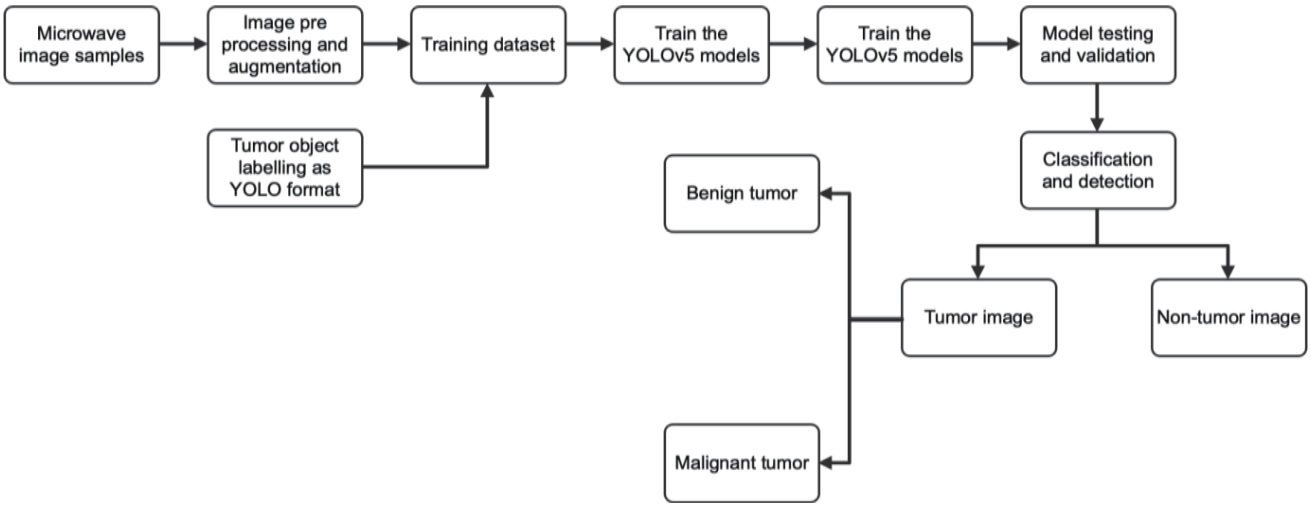

Figure 9. Process flow diagram of classification model.

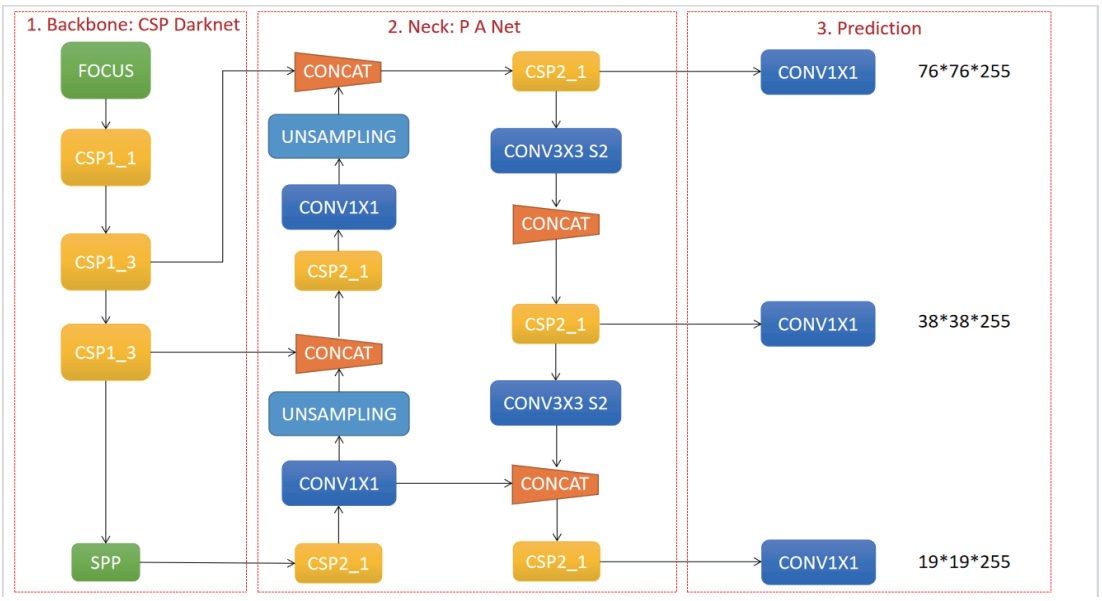

Figure 10. Architecture of YOLOV$_5$ model.

head images, a substantial dataset is needed. Tumor detection using YOLOv$_5$ model is illustrated in Figure 11.

Image labelling and annotations are done by computer vision technique, and the classification performance is evaluated by the metrics like Precision (P), Recall (R), and $F1$ score (Fs) shown by the following Equations (11), (12), and (13). Table 3 shows various metrics associated with classification procedure.

$$P = \frac{N_{TP}}{(N_{TP} + N_{FP})} \tag{11}$$

$$R = \frac{N_{TP}}{(N_{TP} + N_{FN})} \tag{12}$$

$$F_s = \frac{(2XN_{TP})}{(2 \times N_{TP} + N_{FN} + N_{Fp})} \tag{13}$$

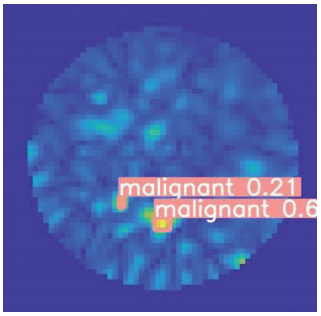

Figure 11. Output image after classification.

Table 3. Classification performance parameters.

No. of Training image	Precision (P) (%)	Recall (R) (%)	F1 score (%)	Train Classification Loss	Validation Classification Loss
500	81.7	86	86.9	0.00665	0.0225
700	82.6	87.8	88	0.00623	0.0192
1000	84.8	88.4	89	0.00620	0.0189

where N_{TP} = Number of True Positive, N_{FP} = Number of True Negative, N_{Fp} = Number of False Positive, N_{FN} = Number of False Negative.

4. CONCLUSION

Microwave imaging is an efficient technique to differentiate healthy and malignant tissue in the breast. Antenna plays a major role to identify tumors in the breast in the early stage. Hence, a high-performance Ultra-Wideband Dielectric Resonator Antenna (DRA-UWB) is used to identify the tumor in the breast. An antenna is sketched in different locations of the breast phantom. Because of the hemispherical structure, a mean value of the reflected signal is high at the centre compared to that at the edge. Hence, the difference in mean value is calculated with and without breast phantom for identifying the tumor location. The overall efficiency of this technique can be improved by using high-performance UWB antenna. The image of the breast is reformed by the DMAS beamforming algorithm. This study suggests using the YOLOV$_5$ algorithm to concurrently identify ROI zones and categorise benign and malignant tumors.

REFERENCES

1. Brenner, D. R., H. K. Weir, A. A. Demers, L. F. Ellison, C. Louzado, A. Shaw, D. Turner, R. R. Woods, and L. M. Smith, "Projected estimates of cancer in Canada in 2020," *Canadian Medical Association Journal*, Vol. 192, No. 9, E199–E205, 2020.
2. Canadian Cancer Statistics Advisory Committee, "Canadian cancer statistics 2019," Toronto, ON, Canadian Cancer Society, 2019, Accessed: 2020-04-09.
3. Gøtzsche, P. C. and K. J. Jørgensen, "Screening for breast cancer with mammography," *Cochrane Database of Systematic Reviews*, Vol. 22, No. 1469–493X (Electronic), CD001877, 2013.
4. Bourqui, J., E. C. Fear, and S. Member, "System for bulk dielectric permittivity estimation of breast tissues at microwave frequencies," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 64, No. 9, 3001–3009, 2016.

5. Kelly, K. M., J. Dean, W. S. Comulada, and S. J. Lee, "Breast cancer detection using automated whole breast ultrasound and mammography in radiographically dense breasts," *European Radiology*, Vol. 20, No. 3, 734–742, 2010.
6. Lalitha, K. and J. Manjula, "Design of UWB antipodal vivaldi antenna with parasitic patch for microwave head imaging system," *2022 International Conference on Computing, Communication, Security and Intelligent Systems (IC3SIS)*, 1–6, 2022, doi: 10.1109/IC3SIS54991.2022.9885397.
7. Fouda, A. and F. L. Teixeira, "Ultra-wideband microwave imaging of breast cancer tumours via Bayesian inverse scattering," *Journal of Applied Physics*, Vol. 115, No. 6, 1–8, 2014.
8. Srinivas, K., G. R. Rao, and A. Govardhan, "Rough-fuzzy classifier: A system to predict the heart disease by blending two different set theories," *Arab J. Sci. Eng.*, Vol. 39, 2857–2868, 2014, https://doi.org/10.1007/s13369-013-0934-1.
9. Edwards, K., J. LoVetri, C. Gilmore, and I. Jeffrey, "Machine-learning-enabled recovery of prior information from experimental breast microwave imaging data," *Progress In Electromagnetics Research*, Vol. 175, 1–11, 2022, doi:10.2528/PIER22051601.
10. Lai, J. C. Y., C. B. Soh, E. Gunawan, and K. S. Low, "Homogeneous and heterogeneous breast phantoms for ultra-wideband microwave imaging applications," *Progress In Electromagnetics Research*, Vol. 100, 397–415, 2010.
11. Selvaraj, V., J. B. J. J. Sheela, R. Krishnan, L. Kandasamy, and S. Devarajulu, "Detection of depth of the tumor in microwave imaging using ground penetrating radar algorithm," *Progress In Electromagnetics Research M*, Vol. 96, 191–202, 2020, doi:10.2528/PIERM20062201.
12. Lalitha, K. and J. Manjula. "Novel method of characterization of dispersive properties of heterogeneous head tissue using microwave sensing and machine learning algorithms," *Advanced Electromagnetics*, Vol. 11, No. 3, 84–92, Oct. 2022, doi:10.7716/aem. v11i3.1821.
13. Khoshdel, V., M. Asefi, A. Ashraf, and J. LoVetri, "Full 3D microwave breast imaging using a deep-learning technique," *J. Imag.*, Vol. 6, No. 8, 80, Aug. 2020.
14. Redmon, J., S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 779–788, Jun. 2016.