

UWB Resonator-Based Supervised Learning for Breast Tumor Diagnosis

Sonal Patil* and Ashwini Naik

Department of Electronics and Telecommunication Engineering, Ramrao Adik Institute of Technology, Nerul, Navi-Mumbai, India

ABSTRACT: This paper proposes an application of ultra-wideband antenna in conjunction with supervised machine learning to detect the existence of breast tumor. The microstrip line fed octagonal shaped UWB antenna is designed by using Ansys high-frequency structure simulator 2022 R2. It is fabricated on double sided copper FR4 epoxy glass substrate of size 40 mm×40 mm and tested by using vector network analyzer N9916A. The antenna structure is optimized over the frequency spectrum of 3.1 GHz to 10.6 GHz to obtain minimum value of return loss. The optimized structure provides bandwidth spectrum of 8.38 GHz covering the frequency range of 2.76 to 11.15 GHz with maximum gain of 5.3 dB at 8 GHz. The homogenous artificial breast phantoms with and without tumor are fabricated using different chemical compositions. The dielectric traits of skin, fatty, glandular and tumor layers are analyzed. Microwave sensing for detecting the presence of breast cancer uses the disparity between tumor and breast tissues, requiring consideration of dispersiveness to accurately assess the dielectric characteristics of the breast model due to its lossy dispersive nature. The three sets of reflection characteristics of the entire system comprised of antenna with phantoms are recorded by using VNA with a gap of week to constitute the dataset. The ultrasonic gel serves as a medium for matching between the breast model and antenna. Further, the supervised machine learning approach is used to improve the detection accuracy. Supervised learning, a key category of machine learning, uses labeled data to predict unseen data. The Logistic Regression, Support Vector Machine, K-Nearest Neighbors, Random Forest and Multilayer Perceptron algorithms are applied on the measured data to classify the healthy and tumorous tissues. The random forest proven to be best fit on the data with auc score of 98.05%.

1. INTRODUCTION

Breast cancer is a global disease, causing a rise in mortality rates in females. The breast cancers in woman are complex structures made of abnormal cells that spread into tissues. It is caused by gene changes, physical genetic expression, and phenotypic features, breast cancer forms in breast cells, leading to uncontrolled growth and potential fatality if not diagnosed early. With declining survival rates as the disease progresses, breast cancer continues to most prevalent rank as the second root reason of deaths related to the cancer among women all over world. Being that the survival rate declines with increasing stages of cancer of the breast, prompt identification is essential for effective therapy, surgery, and extended postinterventional survival of patients. X-rays and mammograms are the standard method for clinical malignancy diagnosis; nevertheless, they have several shortcomings, notably the possibility of false indication and the procedure's difficulties for patients. The likelihood of a false alarm and callback rate is significant, and costly follow-ups for medical biopsies. It is difficult to image breasts with larger tissue densities, and not every lesion can be found. In addition, patients may experience uncomfortable physical and emotional side effects during mammography, and ionising radiation exposure may result in health problems. While MRI is good for screening women who have implants but is quite expensive, ultrasound is used to track the growth and identify different sorts of tumours [1–3]. Non-invasive can-

cer diagnostics, such as microwaves, are vital. Due to the fact that microwave sensing may avoid compressions and ionising radiation, it is a more economically viable means of detecting breast tumours than nuclear medicine and MRI. With the ability to detect tiny lesions and forecast the electrical characteristics of the tissues, this technology improves detection and minimises false positives, enabling harmless, pleasant diagnostic monitoring. The development of wireless communication and Ultra-Wideband standards has increased the demand for broadband antennas with high gain, particularly microstrip antennas, ideal for bio-medical applications due to their size, cost, and electrical parameters [4]. The Federal Communication Commission (FCC) has granted permission for emission measurement procedures in the 3.1–10.6 GHz broad frequency range, offering potential benefits for medical and communication applications [5]. Phantom plays an integral role in the era of bio-medical research. A phantom is a scientific, specialised object that can be employed to mimic a human body organ. Its design should have qualities similar to human tissue for accurate analysis and evaluation. To guarantee optimal performance, the measurable information gathered from the analysis should deliver consistent findings. It is a method for acquiring trustworthy and quantifiable data in order to compare systems utilised in real-world settings. The study on the fabrication of tissue-mimicking materials for phantoms operating at narrowband and ultrawideband have been reported. The investigation shown, with different variations in the combination of the materials used, a great variation in dielectric properties

* Corresponding author: Sonal Patil (sonalpatil606a@gmail.com).

can be obtained. The realistic 3D homogenous and heterogeneous breast phantoms were prepared using various ingredients, including oil-in-gelatin, p-toluic acid, n-propanol, deionized water, Bloom gelatin, Formaldehyde, oil, and Ultra Ivory detergent [6–8]. Innovations in wireless communication technology have driven demand for high-gain broadband antennas, necessitating the adoption of microstrip antennas for biomedical applications. These antennas are portable, lightweight, and low-profile, which renders them simpler to build and integrate into microwave integrated circuits [9–13]. The time-domain approach to breast cancer diagnosis employs a multi-static radar system to detect variations in dielectric characteristics between healthy and tumorous tissue. The system uses 240 signals from a 16-element antenna array to compensate for measurement inaccuracies [14–17]. A study on frequency-domain breast cancer detection found that benign, and malignant breast tissues differ in dielectric properties than healthy tissues [18–23]. The study also investigated the effect of electromagnetic wave interaction on breast cancer detection using a UWB antenna system [24–27]. The study examines various breast cancer classifiers, such as decision trees, Multi-Layer Perception (MLP), Naive Bayes (NB), Sequential Minimal Optimisation (SMO), and Instance Based for K-Nearest Neighbour (IBK) were applied on Wisconsin Breast Cancer (WBC), Wisconsin Diagnosis Breast Cancer (WDBC), and Wisconsin Prognosis Breast Cancer (WPBC) datasets. The WBC and WDBC dataset were found to be superior than the fusions of other classifiers. The study also looked into the usage of a UWB prototype microwave imaging technique for breast tumor categorization. Vivaldi antenna was employed in the system [28–32].

The proposed work aims to develop a system that improves the breast tumor detection accuracy by gathering dataset with the help of UWB antenna and breast phantoms. UWB antennas offer a wide frequency band and multiple resonance frequencies, providing better insights into the dense part of an object. In comparison to the literature, a single antenna sensor reduces the size, complexity and thus, the cost of a microwave sensing technique utilized for the detection of breast cancer. The proposed work is organized in five sections. The fabrication, design details and results of antenna geometry are described in second section, creation of breast model and analysis of dielectric properties in third section. The fourth section demonstrates theory, design details and optimal performance of various machine learning (ML) algorithms. The proposed system is compared with the literature in fifth section and concluded in sixth.

2. ANTENNA DESIGN DETAILS

An ultra-wideband resonator with octagonal patch configuration (OMSA) for analyzing the breast layers is proposed. The OMSA antenna structure is modeled and simulated by using ansys high-frequency structure simulator 2022 R2 simulator. The antenna structure of octagon shaped radiating element is designed on the dielectric substrate of the dimensions 40 mm × 40 mm. The structure is designed by referring rectangular microstrip antenna with following formulae.

For $\frac{W}{h} > 1$,

$$W = \frac{\lambda_o}{2\sqrt{\frac{\epsilon_r + 1}{2}}} \quad (1)$$

$$\epsilon_{re} = \frac{\epsilon_r + 1}{2} + \frac{\frac{\epsilon_r - 1}{2}}{\sqrt{1 + \frac{12h}{W}}} \quad (2)$$

$$\Delta L = 0.412h \frac{(\epsilon_{re} + 0.3) \left(\frac{W}{h} + 0.264\right)}{(\epsilon_{re} - 0.258) \left(\frac{W}{h} + 0.8\right)} \quad (3)$$

$$L = \frac{\lambda_o}{2\sqrt{\epsilon_{re}}} - 2\Delta L \quad (4)$$

where, W : width of the patch, ϵ_r : dielectric constant, ϵ_{re} : effective dielectric constant, λ_o : free space wavelength, ΔL : the effective length, h : the thickness of the dielectric substrate, L : length of the patch.

The patch dimensions are calculated for the operating frequency of 6.85 GHz and wavelength of 43.796 mm/s covering UWB spectrum. The optimized size and side length of octagonal shaped radiating patch are 14 mm × 14 mm and 8.2 mm respectively. The structure is MS fed by a 50 Ω line of width 2 mm with partial ground plane of height 10 mm along the width of the substrate. The optimized structure offers minimum return loss below −10 dB and less VSWR below 2 over the operating frequency of 3.1–10.6 GHz. The Fig. 1(a) depicts the configuration details of the OMSA structure and Table 1 list details of the optimized dimensions of the designed structure. The parametric study has been conducted for different patch dimensions. With increase in patch dimensions return loss improves and gain increases. The optimized antenna structure provides 5.3 dB maximum peak gain at 8 GHz. The simulated results of return loss and gain variations are depicted in Fig. 2. The electric equivalent circuit of the optimized resonator is presented in Fig. 1(b). The values of the circuit components are ascertained by referring [25].

$$C = \frac{\epsilon_o \epsilon_{eff} LW}{2h[\cos^2 \left(\frac{\pi FP}{L}\right)]} \quad (5)$$

$$L = \frac{1}{\omega_r^2 C} \quad (6)$$

$$R = \frac{Q_r}{\omega_r C} \quad (7)$$

$$Q_r = \frac{c\sqrt{\epsilon_{eff}}}{4f_r h} \quad (8)$$

where, C : Capacitance, L : Inductance, R : Resistance of parallel RLC equivalent circuit, ϵ_{eff} : effective permittivity of the medium, ω_r : angular frequency, L : Patch Length, W : Patch Width, h : Substrate Thickness.

The feed line of length 10.4142 mm with width 2 mm is represented by series combination of inductor and capacitor of value 0.60 fH and 0.51 pF. A 50 Ω port is connected at the

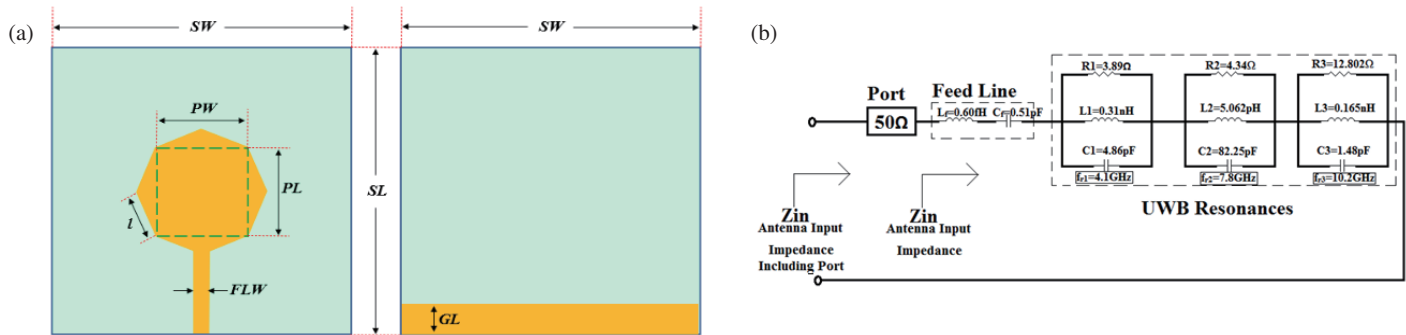


FIGURE 1. Octagonal shaped MSA (OMSA). (a) Geometry. (b) Equivalent circuit.

TABLE 1. Design dimension details of OMSA.

Parameters	Description	Dimension (mm)
SW	Width of the substrate	40
SL	Length of the substrate	40
GW	Width of the ground plane	40
GL	Length of the ground plane	12
PW	Width of the radiating patch	14
PL	Length of the radiating patch	14
FLW	Width of the feedline	2

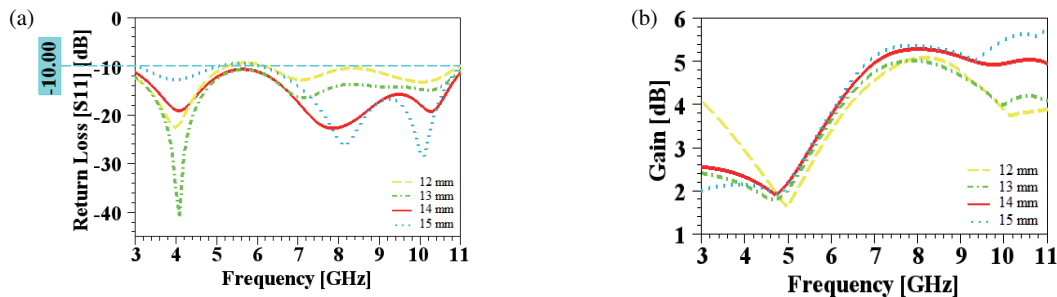


FIGURE 2. S_{11} and Gain variations of different patch dimensions of simple octagonal MSA. (a) Return loss. (b) Gain variations.

end for the excitation of antenna. As the OMSA is an UWB monopole antenna and resonates at three different frequencies 4.1, 7.8 and 10.2 GHz respectively below -10 dB reflection coefficients value. Hence the antenna has triple tuned ultra-wide band behavior. These three resonance circuits are represented by three separate RLC parallel resonance circuits. The evaluated values of all passive elements are indicated on the circuit parameters using fundamental parallel resonance and microstrip discontinuities concepts [33]. An Ultra-wideband (UWB) performance is achieved by series combination of three parallel resonance. The current distribution and radiation pattern at 4.1, 7.8, 10.2 GHz of OMSA are depicted in Figs. 3 and 4 respectively.

The designed UWB resonator is fabricated by using double sided copper flame retardant epoxy glass composite dielectric substrate. The substrate material used is of 1.6 mm thick, ϵ_r of 4.4 and $\tan \delta$ of 0.02. The developed antenna prototype is presented in Fig. 5. The vector network analyzer model N9916A

is used to test the structure and validate the results. The optimized and measured results of return loss and gain variations are shown in Fig. 5. The OMSA structure resonates at three different frequencies 4.1, 7.8 and 10.2 GHz and offers an impedance bandwidth of 8.38 GHz from 2.76–11.14 GHz. The simulated S_{11} values observed at resonating frequencies at 4.1, 7.8 and 10.2 GHz are -19.18 , -22.83 and -19.25 dB respectively. The S_{11} values measured at resonating frequencies at 3.96, 8.01 and 10.257 GHz are -19.62 , -20.39 and -17.32 dB respectively. This structure provides simulated peak gain of 5.3 dB at 8 GHz and measured of 5.18 dB at 7.43 GHz. An ultra-wideband performance is achieved as antenna resonates at different frequencies.

3. BREAST PHANTOM MODEL

The breast phantoms are artificial breast replicas purposive to mimic the properties of their matching aspects. High-fidelity

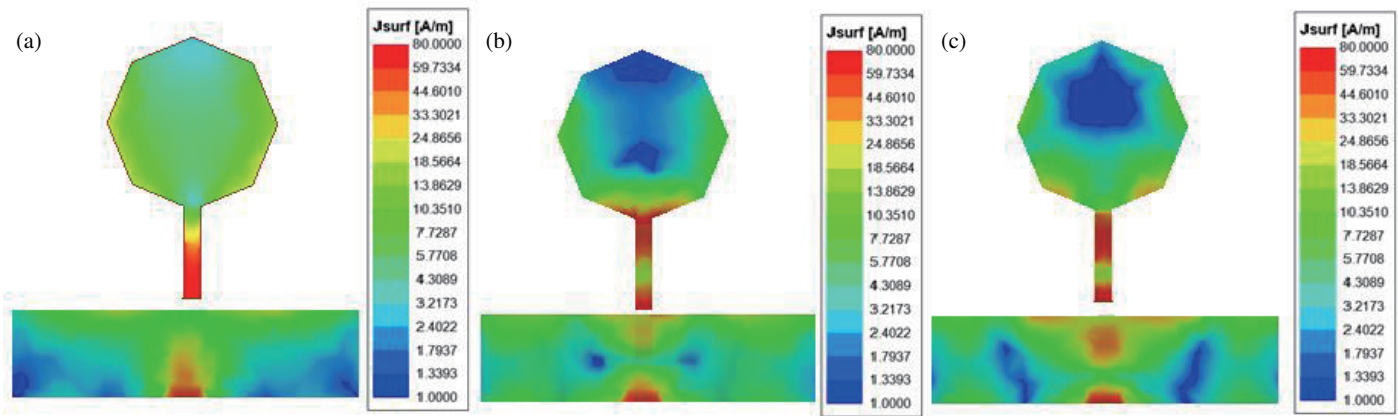


FIGURE 3. Current distribution of simple octagonal MSA. (a) 4.1 GHz. (b) 7.8 GHz. (c) 10.2 GHz.

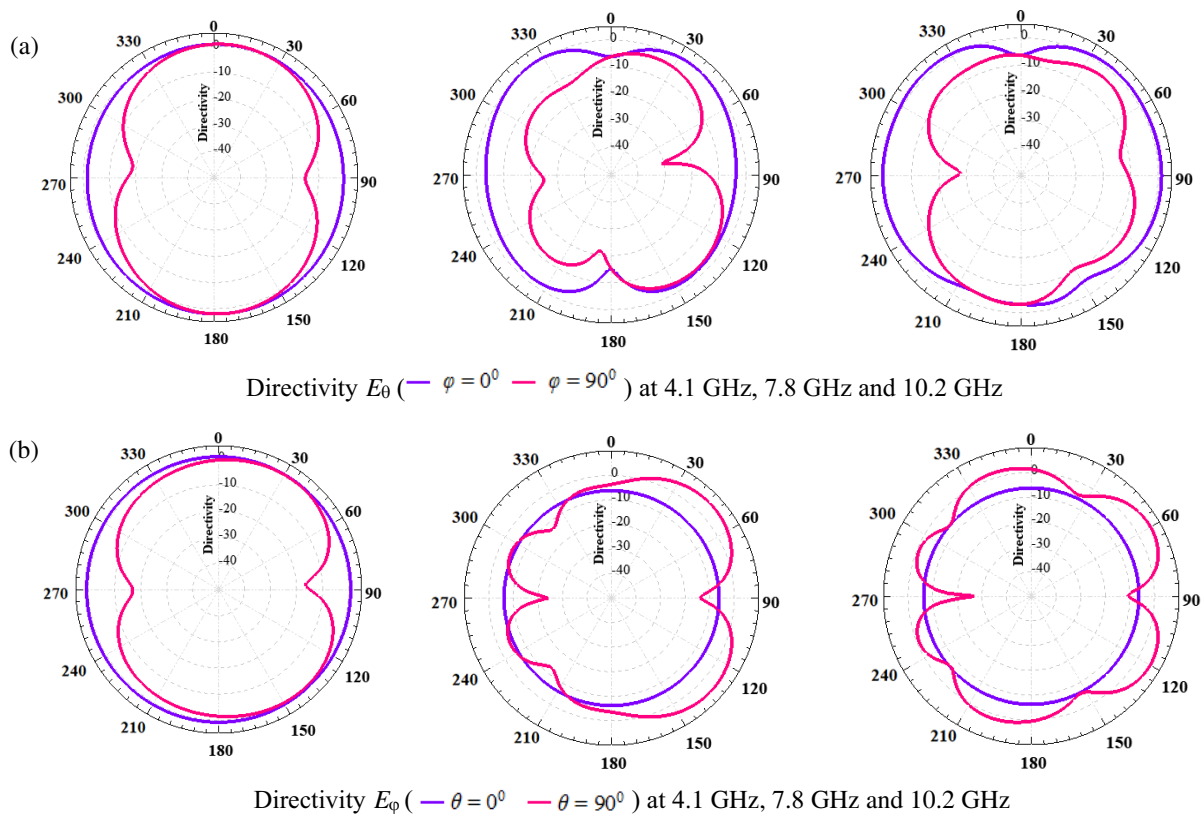


FIGURE 4. Radiation pattern of octagonal MSA.

breast phantoms are useful tools for both computational and experimental research. Homogeneous and heterogeneous breast phantoms are the two different types of breast phantoms. The homogeneous breast phantom has two layers composed of fatty layer on the top of the skin layer and a tumor within. The heterogeneous breast phantom is composed of four layers; epidermal, adipose, glandular and tumors embed inside. The heterogeneous is the most realistic type of the breast phantom and the layered structure is illustrated in Fig. 6. The thickness of skin layer is very thin whereas fatty and glandular layer takes 50% of thickness.

The heterogeneous complex structure found in breast tissue where a few amounts of contrast relay between normal and cancerous tissues. The principle of microwave sensing to detect lesion in the breast is based on dielectric properties of the tissue. The relative permittivity and conductivity of tumours region and different breast layers have large contrast. The radiations from antenna sensor will travel through these breast layers and interact with the breast tissues. In order to deal with the real electrical properties of the breast tissues, the dispersive effect needs to be taken into consideration because the breast tissues seem to be lossy dispersive material for microwave propagation. The resulting signal occurring from these biological ele-

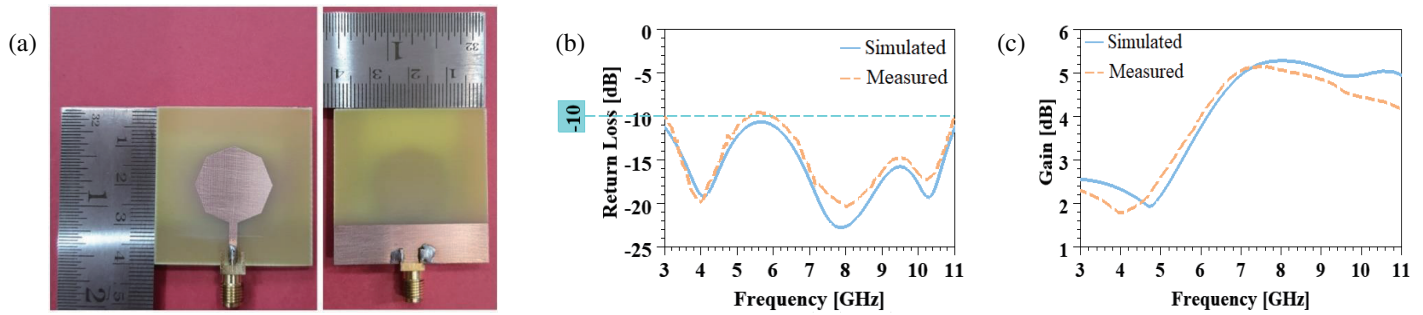


FIGURE 5. Fabricated octagonal microstrip antenna structure. (a) Prototype. (b) Return loss. (c) Gain variations.

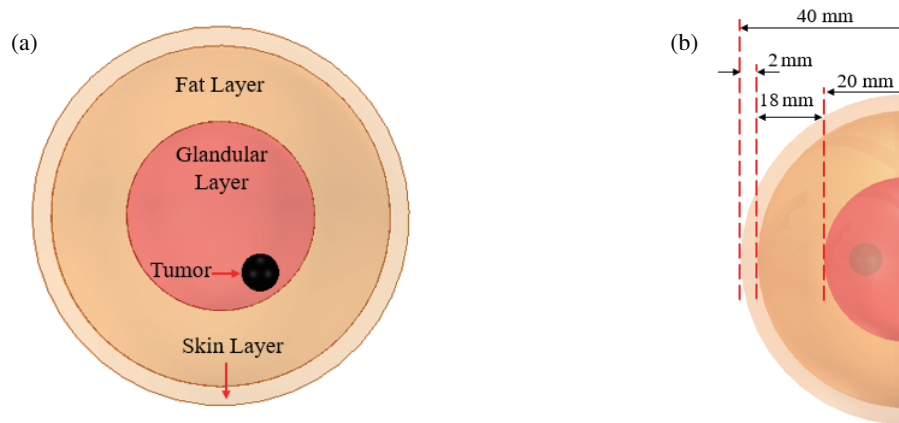


FIGURE 6. Different layers of breast model. (a) Different layers of homogeneous breast model. (b) Dimensions for the breast replica layout.

TABLE 2. Dielectric properties of different layers of breast tissues.

Sr. No.	Breast Tissue Type	ϵ_r	σ_s (S/m)	τ (ps)
1	Skin	40	1.1	7.37
2	Fatty	7	0.1	7.00
3	Glandular	47	2	7.00
4	Tumor	4	3	7.00

ments are owing to microwave characteristics which includes absorption, transmission and reflections. The single pole Debye model is used to describe how these tissues respond to variations in frequency. The dielectric characteristics of the homogeneous breast layers are listed in Table 2.

Debye dispersion model is expressed as,

$$\epsilon_r(\omega) = \epsilon_\infty + \frac{\epsilon_s - \epsilon_\infty}{1 + \omega^2 \tau^2} \quad (9)$$

$$\sigma_r(\omega) = \frac{(\epsilon_s - \epsilon_\infty)\omega^2 \tau \epsilon_0}{1 + \omega^2 \tau^2} + \sigma_s \quad (10)$$

where, ϵ_r — Complex permittivity, ϵ_0 — Free space permittivity, ϵ_s — Static dielectric constant, ϵ_∞ — Dielectric constant at infinite, σ_s — Static conductivity, τ — Pole relaxation constant, ω — Angular frequency.

Here an attempt is made to analyze the dielectric properties of heterogeneous breast phantom using proposed UWB antenna.

The entire system geometry of different breast layers with antenna structure is designed using ansys HFSS software. The different modeling stages are depicted in Fig. 7. The breast phantom of diameter 40 mm is modeled by referring dielectric properties mentioned in Table 2. The 2 mm thick skin layer, 18 mm thick fatty layer and 20 mm thick glandular layer is designed and tumor of 5 mm is embedded inside the gland layer. The proposed antenna is used to analyze the dielectric traits of the different layers of breast model.

The different methods are reported to fabricate breast phantom [6–8]. The heterogeneous breast phantom is developed by using the materials specified in Table 3.

Fabrication steps:

- Mix Propylene glycol and distilled water.
- Place mixture in double boiler with temperature raised to 80°C.

TABLE 3. Material and composition for breast phantom fabrication.

Sr. No.	Material	Skin	Fat	Gland	Tumor
1	Distilled Water (ml)	100	50	100	120
2	Propylene Glycol (ml)	7	2	7	7
3	Agar-Agar Gelatin Powder (g)	6	7	5	10
4	Safflower Oil (ml)	14	40	21	9
5	Liquid Detergent (ml)	0.5	0.5	0.5	0.5
6	Formalin (ml)	0.3	0.3	0.3	0.3
7	Xanthan Gum (g)	2.5	2	2	2

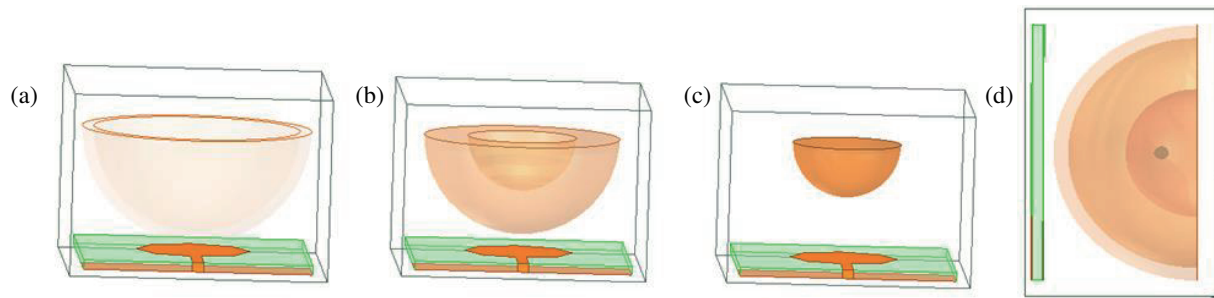


FIGURE 7. Modeling stages of breast phantom. (a) Skin layer. (b) Fat layer. (c) Gland layer. (d) Breast phantom with tumor.

- iii. Add agar-agar gelatin powder and mix until it dissolves completely until mixture changes the color.
- iv. Add liquid detergent and formalin to safflower oil. Mix it with heated solution.
- v. Remove solution from the double boiler. Add Xanthan gum and stir until it cools.
- vi. Place solution in ice bath and stir carefully.
- vii. When mixture reaches 25°C, place it in mold and refrigerate.

The chemical concentration of distilled water, propylene glycol and safflower oil are found superior to modify the permittivity and conductivity of different layers. Distilled water has a high permittivity, which is perfect for increasing the phantom's permittivity; propylene glycol and safflower oil possess low conductivity, which is ideal to creating a low conductivity phantom. Phantom shape is retained by adding powder of agar-agar gelatin. Formalin (37% formaldehyde solution), liquid detergent and xanthan gum are acting as surfactant and thickener. Formalin (37% formaldehyde solution) is used to determine the liquefy temperature of agar-agar gelatin and it also stabilizes the phantom. The mixtures of agar-agar gelatin powder exhibit dielectric characteristics that are strikingly similar to the properties of realistic breast tissue. These materials also offer excellent mechanical strengths and are easy to synthesize. Although varying concentrations of these materials are utilized for creating phantom layers, the chosen materials make the fabrication of breast phantoms simple. The constructed breast tissues, phantoms without and with tumor are presented in Fig. 8.

The system model comprising antenna sensor and breast model is validated by testing with the help of VNA model

N9916A. Three sets of reflection characteristics are recorded with the gap of a week. Here, ultrasound gel is used as the surface matching substance to minimize the losses in between the resonator and breast model. The results from the HFSS software are compared with the average results obtained from the VNA. The reflection characteristics of four tissue types are depicted in Fig. 9. It is apparent by analyzing Figs. 9(d) and 9(a), 9(b), 9(c) that aberrant cells absorb more electromagnetic energy than healthy tissues. As a result of the malignant tissues absorbing more radiation, the tumor layer exhibits the largest return loss. Furthermore, it may be inferred from Figs. 9(e) and 9(f) that the reflection properties of a healthy breast and a tumorous breast differ significantly. The dielectric properties, permittivity, and conductivity of different layers are shown in Fig. 10.

4. SUPERVISED APPROACH FOR CLASSIFICATION

The machine learning approach is a captivating means to reduce the time spent in simulations that involve trial and error simulations when optimizing the geometrical parameters to meet expected design requirements. The supervised learning is a key category of machine learning in which learning model is built on a set of labeled data consisting independent and dependent variables to make predictions on unseen data. This learning approach is an attractive solution to solve regression and classification problem types. Here, the work is extended to identify the existence of the breast tumor using machine learning algorithms. Also, an attempt is made to improve the detection accuracy by applying different machine learning techniques. The machine learning models are built on the recorded

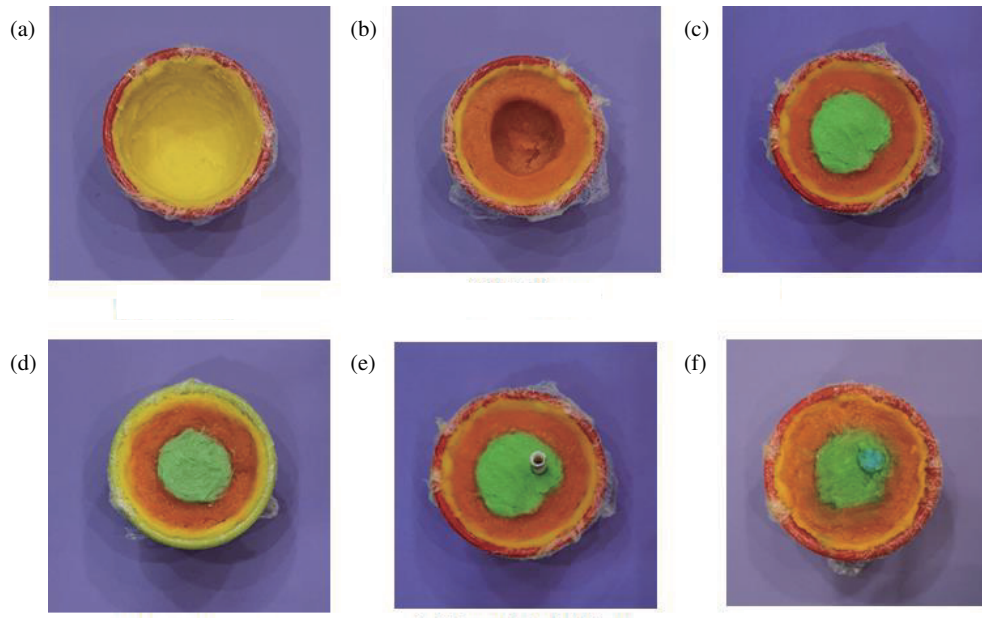


FIGURE 8. Modeling stages of breast phantom. (a) Skin layer. (b) Fat layer. (c) Gland layer. (d) Breast phantom with tumor. (e) Tumor embedding. (f) Breast phantom with tumor.

datasets from VNA to detect the healthy or cancerous tissues. The dataset is generated by recording three sets of reflection characteristics over a week-long period for the entire system, which includes an octagonal microstrip antenna and a breast model with and without a tumor. The detection classes are categorized as healthy tissue and tumorous tissue. Here, the notion that the left and right breasts are identical is employed for detecting the lesion. Consequently, by contrasting healthy breast with cancerous breast, the detection of a breast tumor can be examined. The different supervised classification algorithms used for detection are logistic regression, support vector machine, K-Nearest Neighbors, Random Forest, and Multilayer Perceptron. The dimension of the dataset used for predictive modelling is (804, 3) which includes the frequency with corresponding reflection characteristics data points from VNA and class labels. The stratified sampling technique with a 70 : 30 ratio is utilized to split the dataset into training and testing sets.

One of the easiest algorithms to build the predictive model is the logistic regression algorithm, which estimates the likelihood that a certain occurrence belongs to a particular class. By using a predetermined set of independent factors, model can predict the categorical dependent variable. In order to calculate the probability for the specified class, a sigmoid function is used. The projected values are converted to probabilities by using the sigmoid function which transforms any real value within a range of 0 and 1, forming “S-shaped” curve. The sigmoid function is represented by,

$$f(x) = \frac{1}{1 + e^{-x}} \quad (11)$$

where, x signifies weighted summation of independent variables.

The model provides detection accuracy of 90.47%. The logistic regression classifier fails to accurately detect the data points which fall near the decision boundary.

The most popular ML technique for classification is support vector machine (SVM). The SVM provides optimal hyperplanes which helped to overcome the mentioned drawback. The SVM divides the data points into two groups: those that fit inside a preset tube of width ε and are not penalized, and points which fall outside this boundary and are thus penalized shown in Fig. 11. The hyperplanes in the SVM classifiers try to give the margin between the closest datapoints of different classes. The discriminant function is given by,

$$g(X^{(s)}) = w_0^T X^{(s)} \pm b_0 \quad (12)$$

The algebraic distance r from the support vectors $X^{(s)}$ is given by,

$$r = \frac{g(X^{(s)})}{\|w_0\|} \quad (13)$$

The margin of separation ε between two classes is,

$$\varepsilon = 2r \quad (14)$$

The SVM model is implemented with Gaussian kernel. The model is tuned by optimizing the hyperparameters: kernel of radial basis function (RBF), kernel coefficient and regularization parameter to improve the accuracy. RBF can map an input space in infinite dimensional space. The tuned model parameters of kernel coefficient of 0.01 and regularization parameter of 0.0067 provide an accuracy of 91.32%.

The artificial neural network (ANN) comprises three layers: input, hidden, and output in which input signal is propagated in forward direction on layer-by-layer basis, this kind of architecture is commonly referred as Multi-layer Perceptron (MLP).

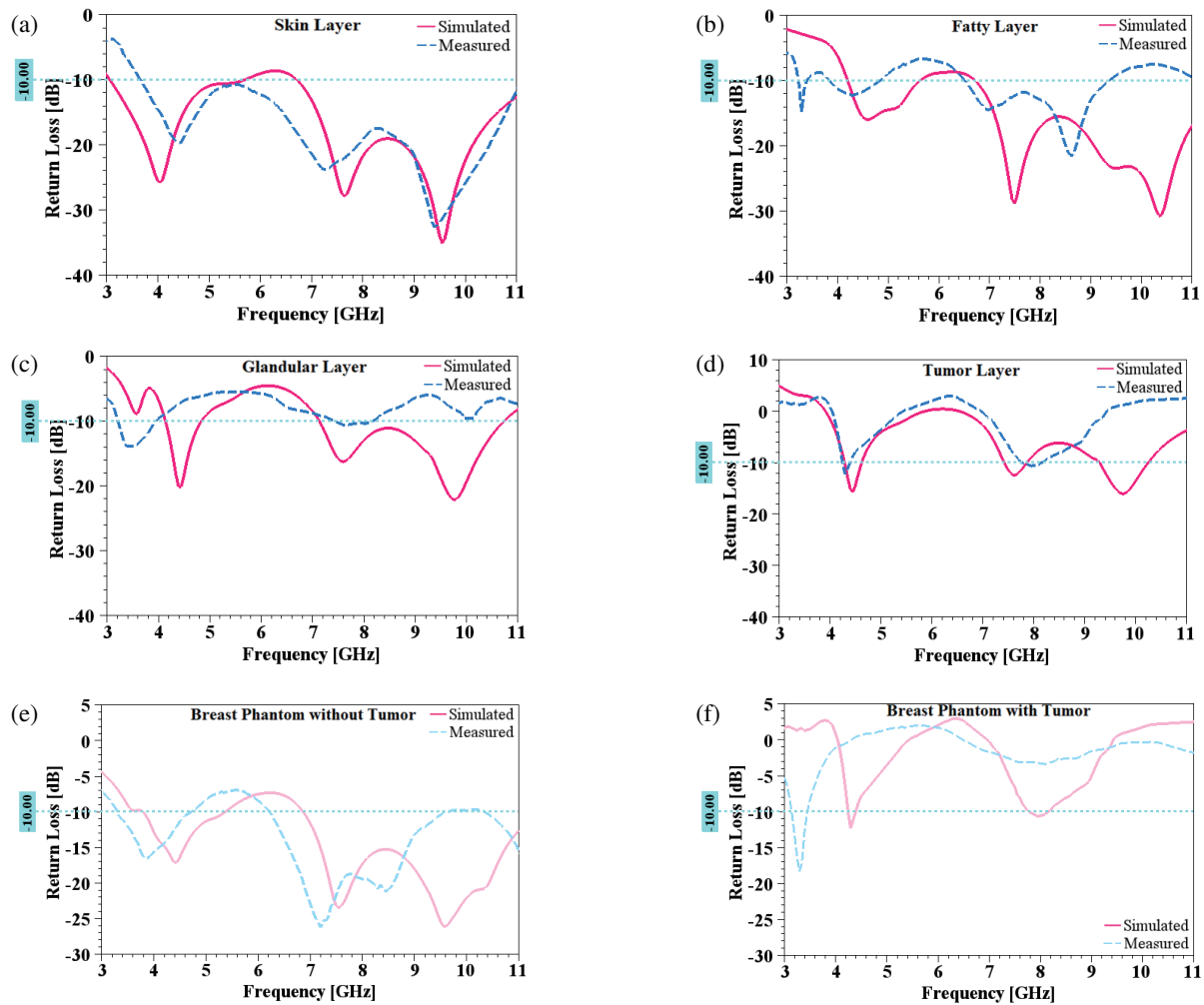


FIGURE 9. Reflection characteristics of different tissue types of breast model. (a) Skin layer. (b) Fatty layer. (c) Glandular layer. (d) Tumor layer. (e) System model without tumor. (f) System model with tumor.

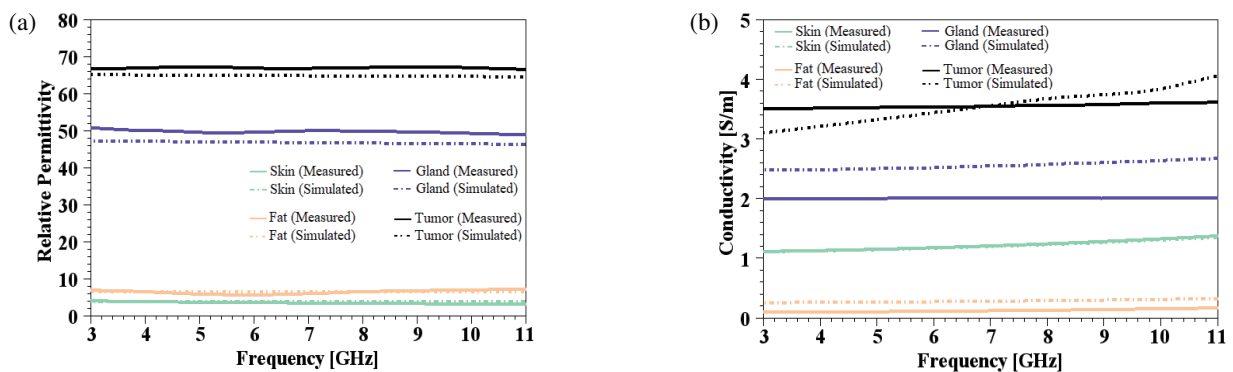


FIGURE 10. Dielectric properties of different breast tissues. (a) Relative permittivity. (b) Conductivity.

MLP algorithm is used to make the model denser. The accuracy of the model is increased by tuning the parameters of the algorithm: the dimensions of the hidden layer, learning rate for backpropagation, and an activation function. “S-shaped” hyperbolic tangent nonlinear activation function is used which maps the input data points into the range of $(-1, 1)$ for classifi-

cation. The neural network is built for the hidden layer of size (50, 50). The learning rate to update the weights is optimized to smaller value of 0.001. The MLP model provides a detection accuracy of 89.67%. The performance of the model observed is poor due to overfitting.

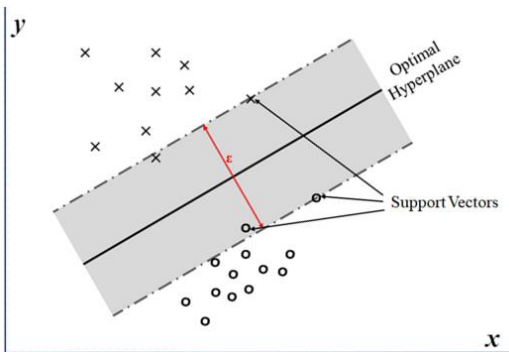


FIGURE 11. Illustration of Hyperplane.

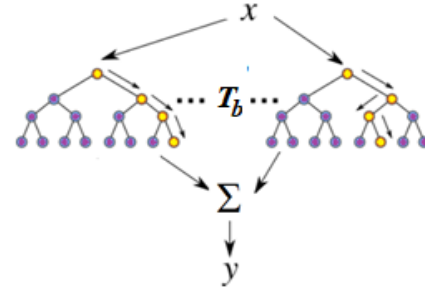


FIGURE 12. Representation of Random-Forest Tree.

TABLE 4. Performance metrics for multiple supervised machine learning algorithms.

Sr. No.	Algorithm	Accuracy	Precision	Sensitivity	Specificity	F1 Score	AUC Score
1	Logistic Regression	0.9049	0.8837	0.9344	0.875	0.9083	0.9047
2	Support Vector Machine	0.9132	0.8531	1.0	0.825	0.9208	0.9125
3	Multilayer Perceptron	0.8967	0.8540	0.9590	0.8334	0.9034	0.8962
4	K-Nearest Neighbors	0.9628	0.9593	0.9672	0.9583	0.9084	0.9627
5	Random Forest	0.9804	0.9672	0.9672	0.9667	0.9672	0.9805

Another non-parametric algorithm, K-Nearest Neighbors (KNN), is studied and implemented. This classifier incorporates vicinity to estimate predictions about the clustering of an individual data point and based on distance metric. Here, the measure of distance used to build the model is 'Minkowski' which results in the standard Euclidean distance when power parameter $p = 2$. The number of neighbors to classify the data point is defined by k-value. The selection of k-value is a tedious job as smaller value of k leads to high variance — low bias, and hefty value leads to low variance — high bias. The k value chosen to obtain improved accuracy is 3. This tuned model provides an accuracy of 96.28%.

Bagging or bootstrap technique is proven to be attractive to mitigate the variance of a projected prediction function. Trees are excellent candidates for bagging because, when being developed deeply enough, they have minimal bias and can recognize intricate interaction patterns in the data. Decision tree is a pattern classification algorithm used in genetic and bioinformatics. By constructing an ample number of de-correlated trees and averaging them, random forest is significantly impacting bagging. It is applicable to both classification and regression tasks, using distances between observations to assign new values. The Random Forest algorithm, among other well-known machine learning techniques, offers distinct predictive validity and model interpretability, making it a preferred choice for classification in complex data sets. During the training phase, the algorithm builds an enormous number of trees and outputs the class. The algorithm architecture is illustrated in Fig. 12.

For a batch of $b = 1$ to B ,

1. The random-forest trees T_b are built by randomly drawn bootstrap instances of size N and m estimators at each node.
2. Select the optimal split-point out of the m .
3. Split the node in half to its two offspring nodes.
4. Acquire the assortment of trees $\{T_b\}_1^B$.
5. At a newly discovered point x , to predict the class:

$$\hat{C}_{RF}^B(x) = \text{Majority vote} \left\{ \hat{C}_b(x) \right\}_1^B \quad (15)$$

where $\hat{C}_b(x)$ is the class prediction of the b th random-forest tree.

The randomized search is used to obtain the best estimators for model building. The grid is optimized to obtain the best parameters such as estimators ranging from 100 to 1000, max features, max depth from 5 to 35, min samples leaf from 1 to 20, and min samples split from 2 to 20 with 50 iterations. With the best parameters obtained from randomized search, the random forest (RF) algorithm provides detection accuracy of 98.04%.

The mentioned algorithms are analyzed using accuracy, precision, sensitivity, specificity, F1 score, and auc score. These performance metrics for mentioned algorithms are listed in Table 4. The random forest algorithm is demonstrated to have an excellent fit for the model on the measured data, which leads to improved detection accuracy. The different performance metrics for Logistic Regression, Support Vector Machine, K-Nearest Neighbors, Random Forest, and Multilayer Perceptron algorithms are listed in Table 4. The performance of these classification models with respect to true and false positive rate is depicted in Fig. 13.

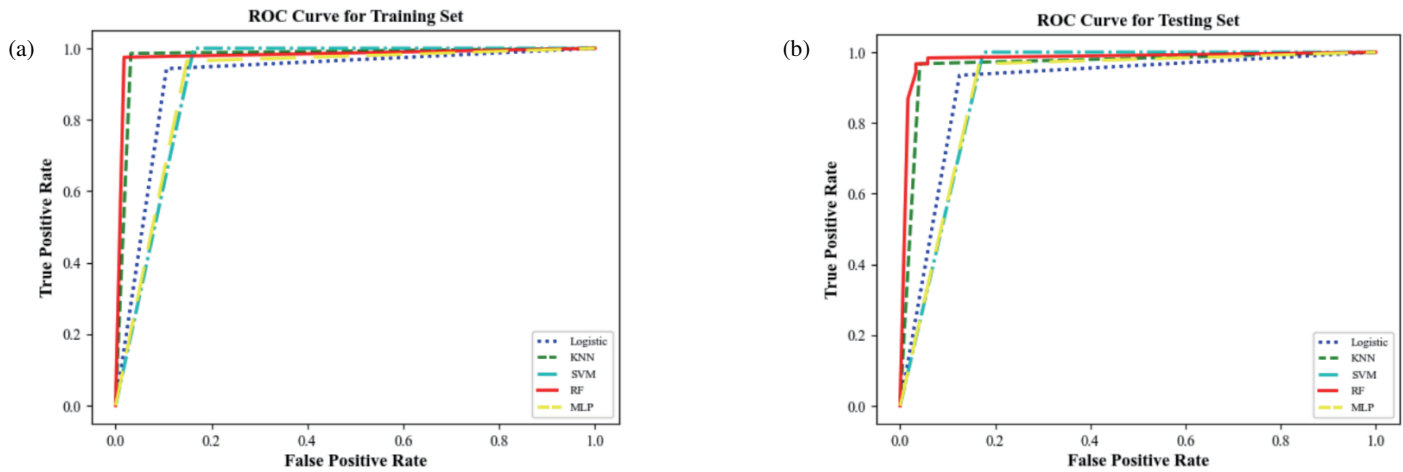


FIGURE 13. ROC curve across multiple machine learning techniques on training and testing data. (a) ROC curve for training samples. (b) ROC curve for testing samples.

TABLE 5. Comparison of invented UWB antennas with proposed antenna.

Ref. No.	Antenna Configuration	Antenna Size (mm ²)	No. of Antenna Sensors	Band (GHz)	Gain (dB)	Dataset	Algorithm	Performance Metric (Accuracy)
[28]	Breast Cancer Diagnosis by Experimental Comparison of Classifiers	-	-	-	-	WBC	NB	95.99%
						WDBC	NB	96.66%
						WBC	J48	95.13%
						WDBC	SMO	97.71%
						WDBC	IBK	95.95%
[29]	Vivaldi antenna with circular holographic for microwave imaging	56×56	2	0.4–6.4	-	UWB transceivers	LDA	87.10%
						Prototype	QDA	89.29%
[30]	Multi-static Radar System operated in Time-Domain for Microwave Breast Screening	-	16 element arrays (2 switching matrix)	2–4	+35	UWB	LDA	70.30%
						Prototype	SVM	73.64%
[31]	Ultrawide-Band transceivers for detection of breast tumor using neural networks.	-	2	3–11	-	UWB transceivers antenna setup	NN trainscg	94.42%
							NN traingdm	88.42%
[32]	UWB-RMSA for Breast Cancer Detection using microwave sensing.	20.47×10.56	2	3–18	4.27	UWB RMSA Prototype	RF	94.4%
PA	Simple octagonal microstrip antenna for breast tumor detection.	40×40	1	2.76–11.14	5.3	UWB OMSA Prototype	LR	90.49%
							SVM	91.32%
							MLP	89.67%
							KNN	96.28%
							RF	98.04%

The two-dimensional space underlying the total Receiver Operating Characteristic (ROC) curve spanning (0, 0) to (1, 1) is determined by the Area under the ROC Curve. The ROC curve is a probabilistic graphical illustration that demonstrates how well a classification model performs across different thresholds.

$$TPR = \frac{TP}{TP + FN} \quad (16)$$

$$FPR = \frac{FP}{FP + TN} \quad (17)$$

where TP : true positives; FP : false positives; TN : true negatives; FN : false negatives.

5. COMPARISON WITH LITERATURE

The proposed octagonal microstrip antenna designed to operate in UWB frequency spectrum and to analyze the dielectric traits of the different layers of the breast model. Breast tumor detection accuracy is improved with the application of machine learning algorithms. Thereby, the antenna structure is compared with the UWB antennas used for the application of breast tumor detection in the literature with respect to the size of antenna, number of sensors, operating frequency spectrum, dataset, algorithms, and detection accuracy with accuracy score, and comparison is listed in Table 5.

6. CONCLUSION

An octagonal radiating patch configuration for detecting the existence of breast tumor using supervised machine learning approach is presented. The high-frequency structure simulator 2022 R2 software is used to optimize the proposed structure. The antenna is made on a double sided copper FR4 epoxy glass substrate and tested using vector network analyzer model N9916A. The proposed structure provides the bandwidth of 8.38 GHz from 2.76–11.14 GHz, and the maximum gain observed at 8 GHz is 5.3 dB. The designed antenna offers triple tuned ultra-wideband behavior as it resonates at 4.1, 7.8, 10.2 GHz frequencies below -10 dB of reflection coefficient values. The heterogenous breast phantom having four breast tissues is fabricated by using safflower oil, distilled water, propylene glycol, agar-agar gelatin powder, formalin, xanthum gum, and liquid detergent. The safflower oil, distilled water, and propylene glycol are proven to be effective in obtaining the needed dielectric properties to give phantom more realism. Agar-agar gelatin, xanthum gum, and liquid detergent are beneficial in stabilizing and preserving the shape of the phantom. Machine learning is a powerful method for optimizing geometrical parameters and reduces simulation time. This approach is used to detect breast tumors using various algorithms. The dataset is gathered by testing the antenna structure and breast phantom with and without tumor. The ultrasound gel is used as matching medium between antenna and breast model. The three sets of reflection characteristics from a week-long of different breast layers, phantoms with and without tumor are recorded. Different algorithms are applied on the recorded dataset and used to categorized tissues as healthy or tumorous. The accuracy obtained by like logistic regression, support vector machine, Multilayer Perceptron, K-Nearest Neighbors, and Random Forest is 90.49%, 91.32%, 89.67%, 96.28%, and 98.04%, respectively. The highest auc score among the mentioned algorithms is obtained by RF algorithm of 98.05% and hence proven to be the best fit on the recorded dataset to improve the detection accuracy.

REFERENCES

- [1] Mathur, P., K. Sathishkumar, M. Chaturvedi, P. Das, K. L. Sudarshan, S. Santhappan, V. Nallasamy, A. John, S. Narasimhan, F. S. Roselind, and I.-N.-N. I. Grp, "Cancer statistics, 2020: Report from national cancer registry programme, india," *JCO Global Oncology*, Vol. 6, 1063–1075, Jun. 2020.
- [2] Garcia-Figueiras, R., S. Baleato-Gonzalez, A. R. Padhani, A. Luna-Alcala, J. A. Vallejo-Casas, E. Sala, J. C. Vilanova, D.-M. Koh, M. Herranz-Carnero, and H. A. Vargas, "How clinical imaging can assess cancer biology," *Insights into Imaging*, Vol. 10, No. 1, Springer, 2019.
- [3] Heywang-Koebrunner, S. H., A. Hacker, and S. Sedlacek, "Advantages and disadvantages of mammography screening," *Breast Care*, Vol. 6, No. 3, 199–207, 2011.
- [4] Fear, E. C., P. M. Meaney, and M. A. Stuchly, "Microwaves for breast cancer detection?" *IEEE Potentials*, Vol. 22, No. 1, 12–18, 2003.
- [5] Washington, D. C., "First report and order, revision of part 15 of the commissions rule regarding ultra-wideband transmission systems," *Federal Communications Commission*, Apr. 2002.
- [6] Lazebnik, M., E. L. Madsen, G. R. Frank, and S. C. Hagness, "Tissue-mimicking phantom materials for narrowband and ultra-wideband microwave applications," *Physics in Medicine and Biology*, Vol. 50, No. 18, 4245–4258, Sep. 2005.
- [7] Zhang, H., S. Y. Tan, and H. S. Tan, "A novel method for microwave breast cancer detection," in *2008 Asia Pacific Microwave Conference (APMC 2008)*, Vol. 1-5, 2342–2345, Hong Kong, China, IEEE, 2008.
- [8] Islam, M. T., M. Samsuzzaman, S. Kibria, and M. T. Islam, "Experimental breast phantoms for estimation of breast tumor using microwave imaging systems," *IEEE Access*, Vol. 6, 78 587–78 597, 2018.
- [9] Ray, K. P., "Design aspects of printed monopole antennas for ultra-wide band applications," *International Journal of Antennas and Propagation*, Vol. 2008, 1–8, 2008.
- [10] Dastranj, A., A. Imani, and M. Naser-Moghaddasi, "Printed wide-slot antenna for wideband applications," *IEEE Transactions on Antennas and Propagation*, Vol. 56, No. 10, 3097–3102, Oct. 2008.
- [11] Shaalan, A. A. and M. I. Ramadan, "Design of a compact hexagonal monopole antenna for ultra-wideband applications," *Journal of Infrared Millimeter and Terahertz Waves*, Vol. 31, No. 8, 958–968, Aug. 2010.
- [12] Zhu, X.-F. and D. Su, "A study of a compact microstrip-fed UWB antenna with an open T-slot," *Progress In Electromagnetics Research Letters*, Vol. 13, 181–189, 2010.
- [13] Kahar, M., A. Ray, D. Sarkar, and P. P. Sarkar, "An UWB microstrip monopole antenna for breast tumor detection," *Microwave and Optical Technology Letters*, Vol. 57, No. 1, 49–54, Jan. 2015.
- [14] Porter, E., E. Kirshin, A. Santorelli, M. Coates, and M. Popovic, "Time-domain multistatic radar system for microwave breast screening," *IEEE Antennas and Wireless Propagation Letters*, Vol. 12, 229–232, 2013.
- [15] Porter, E., E. Kirshin, A. Santorelli, and M. Popovic, "Microwave breast screening in the time-domain: Identification and compensation of measurement-induced uncertainties," *Progress In Electromagnetics Research B*, Vol. 55, 115–130, 2013.
- [16] Afyf, A., L. Bellarbi, A. Achour, F. Riouch, and A. Errachid, "A novel low cost UWB antenna for early breast cancer detection," *American Journal of Electromagnetics and Applications*, Vol. 3, No. 5, 31–37, 2015.
- [17] Cheng, Y. and M. Fu, "Dielectric properties for non-invasive detection of normal, benign, and malignant breast tissues using microwave theories," *Thoracic Cancer*, Vol. 9, No. 4, 459–465, Apr. 2018.
- [18] Selvaraj, V. and P. Srinivasan, "Interaction of an EM wave with the breast tissue in a microwave imaging technique using an ultra-wideband antenna," *Biomedical Research*, Vol. 28, No. 3, 1025–1030, 2017.
- [19] Subramanian, S., B. Sundarambal, and D. Nirmal, "Investigation on simulation-based specific absorption rate in ultra-wideband antenna for breast cancer detection," *IEEE Sensors Journal*, Vol. 18, No. 24, 10 002–10 009, 2018.
- [20] Amdaouch, I., O. Aghzout, A. Naghar, A. V. Alejos, and F. J. Falcone, "Breast tumor detection system based on a compact UWB antenna design," *Progress In Electromagnetics Research M*, Vol. 64, 123–133, 2018.
- [21] Padmavathy, T. V., P. Venkatesh, D. Bhargava, and N. Sivakumar, "Design of I-shaped dual C-slotted rectangular microstrip patch antenna (I-DCSRMPA) for breast cancer tumor detection," *Cluster Computing-the Journal of Networks Software Tools and Applications*, Vol. 22, No. 6, 13 985–13 993, Springer, 2019.

- [22] Islam, M. T., M. Samsuzzaman, M. Faruque, M. J. Singh, and M. T. Islam, "Microwave imaging based breast tumor detection using compact wide slotted UWB patch antenna," *Optoelectronics and Advanced Materials-Rapid Communications*, Vol. 13, No. 7-8, 448–457, Jul. 2019.
- [23] Aydin, E. A. and M. K. Keles, "UWB rectangular microstrip patch antenna design in matching liquid and evaluating the classification accuracy in data mining using random forest algorithm for breast cancer detection with microwave," *Journal of Electrical Engineering & Technology*, Vol. 14, No. 5, 2127–2136, Springer, 2019.
- [24] Rao, P. K. and R. Mishra, "Elliptical shape flexible MIMO antenna with high isolation for breast cancer detection application," *IETE Journal of Research*, Vol. 69, No. 1, 325–333, Jan. 2023.
- [25] Hossain, A., M. T. Islam, A. F. Almutairi, M. S. J. Singh, K. Mat, and M. Samsuzzaman, "An octagonal ring-shaped parasitic resonator based compact ultrawideband antenna for microwave imaging applications," *Sensors*, Vol. 20, No. 5, 1354, Mar. 2020.
- [26] Ponnappalli, V. L. N. P., S. Karthikeyan, and J. L. Narayana, "A circular slotted shaped UWB monopole antenna for breast cancer detection," *Progress In Electromagnetics Research Letters*, Vol. 104, 57–65, 2022.
- [27] Sri, P. A. V. and K. K. Naik, "U-slotted elliptical shape patch antenna for UWB on-body communications," *Progress In Electromagnetics Research Letters*, Vol. 104, 77–85, 2022.
- [28] Salama, G. I., M. B. Abdelhalim, and M. A.-e. Zeid, "Experimental comparison of classifiers for breast cancer diagnosis," in *2012 Seventh International Conference on Computer Engineering & Systems (ICCES)*, 180–185, Ain Shams Univ, Fac Engn, Nov. 2012.
- [29] Conceicao, R. C., H. Medeiros, M. O'Halloran, D. Rodriguez-Herrera, D. Flores-Tapia, and S. Pistorius, "Initial classification of breast tumour phantoms using a UWB radar prototype," in *2013 International Conference on Electromagnetics in Advanced Applications (ICEAA)*, 720–723, IEEE, 2013.
- [30] Santorelli, A., E. Porter, E. Kirshin, Y. J. Liu, and M. Popovic, "Investigation of classifiers for tumor detection with an experimental time-domain breast screening system," *Progress In Electromagnetics Research*, Vol. 144, 45–57, 2014.
- [31] Aydin, E. A. and M. K. Keles, "UWB rectangular microstrip patch antenna design in matching liquid and evaluating the classification accuracy in data mining using random forest algorithm for breast cancer detection with microwave," *Journal of Electrical Engineering & Technology*, Vol. 14, No. 5, 2127–2136, Springer, 2019.
- [32] Alshehri, S. A., S. Khatun, A. B. Jantan, R. S. A. R. Abdullah, R. Mahmud, and Z. Awang, "3D experimental detection and discrimination of malignant and benign breast tumor using nn-based uwb imaging system," *Progress In Electromagnetics Research*, Vol. 116, 221–237, 2011.
- [33] Gupta, K. C., R. Garg, I. Bahl, and P. Bhartia, *Microstrip Lines and Slotlines*, 2nd ed., Artech House, 1996.