Al-Salam Journal for Engineering Science and Technology
Journal Homepage: http://journal.alsalam.edu.iq/index.php/ajest
ISSN: 2790-4822

Breast Cancer Diagnosis By IoB Implanted Tag Design And Machine Learning

Heba Mahdi Salih[1], Furkan Rabee [2]

[1,2] Computer Science Department, Faculty of Computer Science and Mathematics University of Kufa, Najaf, 00964, IRAQ

Corsponding Author: Heba Mehdi Salih,

DOI: https://doi.org/10.55145/ajest.2023.01.01.001
Received June 2022; Accepted September 2022; Available online September 2022

ABSTRACT: Breast cancer is one of the worst diseases in the world and the most common cancer affected by women. Early detection of cancers allows for faster treatments. Recent studies have focused on early breast cancer diagnosis utilizing non-invasive UWB technologies. This article proposed to use metamaterials as an Implantable antenna to detect breast cancer in filed of IOB. With non-toxic materials,and safity frequency range from 1 to 10 GH three different compact and comfortable sizes for metamaterial antennas have been used for implanted within the breast tissue. Two models for compressed breast tissue were created using the CST Microwave studio simulator. These models generated patient data set with differing dielectric properties similar to human tissue. These dataset are used to train several appropriate supervised machine learning algorithms: Decision tree (DT), support vector machine (SVM), and nearest neighbour (NN) in order to develop an intelligent classification model that can assist doctors in identifying malignant breast cells. As a result SVM can classify the breast data to detect the tumor affactivly with 93%accuracy .

Keywords:Machine learning,metamaterial ,implanted antenna,breast cancer,IOB,SAR

1. INTRODUCTION

Cancer is a collection of more than one hundred illnesses with the same defining characteristic: the uncontrolled spread of malignant cells throughout the body. Before the illness has progressed to an advanced level, people with cancer do not experience any discomfort or warning signals from the condition. So, it is the best way to find cancer in its early stages, which increases the chance that oncologic treatment will work. Microwave-based breast cancer imaging is one of the most promising technologies for identifying breast cancer in its early stages. The difference in dielectric characteristics of normal breast tissues (skin, fat, fibroglandular, etc.) and malignant breast tissues (tumor) is used in this procedure. [1]. Therefore, Microwave imagining is expected to be an x-ray alternative. Microwaves are electromagnetic waves that use the tissue’s dielectric characteristics caused by its water content. Because the malignant tumor contains much water, it has a greater dielectric constant than normal tissue, which causes much scattering when electromagnetic waves pass through it. The dielectric constant of a malignant tumor will be at least three times higher than the healthy tissue.[2]. This paper proposes an SRR chipless metamaterial antenna, which is implementable inside the human body with non-toxic materials and safe frequency bands for human tissue from 1 to 10 GHz [3, 4]. The antenna is good for detecting breast tumors early in women who have previously been infected and recovered from the illness or inherited a family history of this disease. Many suggested automated breast diagnostic systems include signal or image preprocessing, segmentation, and machine learning-based diagnosis. [5]. Such systems have shown to be beneficial in assisting doctors with diagnosis. Scientists have developed a wide range of machine learning-based methods for diagnosing illnesses. According to researchers, machine-learning algorithms effectively identify problems such as heart disease, diabetes, liver disease, dengue fever, and hepatitis. [6]. Microwaves with limited energy may be used to light up a breast when the patient is seated or lying down using microwave breast prototype devices. Signals resulting from backscattering are documented. The malignant tumor may be diagnosed by studying the backscattered signals and extracting the tumor signature contained within; In fact, prior research has shown that backscattered signals might vary depending on the size and location of breast cancers. In this paper, a study proposes using machine learning to identify the signals of breast cancers and make diagnoses.

Non-ionizing microwave signals are sent to the breast model, and the scattering parameters (S-parameters) are read. The read signals representing the raw data were fed into a machine model trained to detect a tumor’s existence. So, the proposed model has a good chance of being used in the future to find tumors early and save women’s lives.

*Corresponding author: hebam.almosawy@student.uokufa.edu.iq

2. RELATED WORK

We showed breast cancer prediction studies' work. Predicting whether a person will acquire cancer and machine learning may be useful for the medical field and individuals. Numerous researchers have employed various techniques or models to diagnose breast cancer; some of these are discussed in this section. Alibakhshikenari M [7] proposed a metamaterial idea that inspired this planar antenna array. The sub-wavelength resonant elements will be used in microwave medical imaging systems to identify malignancies in biological tissues. Selvaraj V [8] proposed a study for a more improved approach for characterizing breast tissue using a special microstrip antenna. Geetharamani, G. and T. Aathmanesan [9] proposed A Terahertz (THz) antenna for breast cancer diagnosis inspired by work on metamaterial. Still, it used dimension in nm, which is not practical for bending antenna and unsafe for human tissue. Alrayes, N. and M.I. Hussein.J. Boparai, and M. Popović [10] have researched Phantom breasts used to collect the information. Different kinds of materials have been used to create breast phantoms.V. Vijayasarveswari *et al.* [11] proposed studying the breast cancer sampling technique with dielectric properties for breast phantoms like real breasts. The feature extraction from UWB sensors employing different data normalization approaches, feature reduction algorithms, and classification stages are identified based on permittivity and conductivity.B. S. Bari *et al.*[12] The authors report creating an Artificial Neural Network-based "user-friendly," cost-effective UWB technique for the early detection of breast cancer. A feedback forward propagation Classification Method is utilized to determine the cancer's existence, size, and three-dimensional position. (3D).N. Ali, M. A. Sree, R. Uyguroglu, and A. Allam [13] examined how a woman's breast's electromagnetic properties can be used to determine if she has cancer; malignant cells are different from healthy cells in many respects. The second stage of a cancer diagnosis was determined by the size and quantity of tumors found in the breast and nearby lymph nodes. R. C. Conceição *et al.*[14] Researchers have categorized breast tumor models of various sizes and forms using a monostatic radar microwave imaging prototype technology and machine learning techniques for cancer detection. Azab, M.Y [15] A metamaterial biosensor design and analysis is presented. The proposed structure consists of two hexagonal gold loops on a polyamide substrate for early cancer detection. A. Celik, K. N. Salama, and A. M. Eltawil [16] proposed that the Internet of Bodies (IoBs) will soon play a significant role in the Internet of Things. Devices (e.g., worn, embedded, implanted, swallowed, etc.) are interconnected, with the human body creating a network. As a result, the IoB has many industries, including medicine, safety, security, wellness, entertainment, and more.

This paper proposed three appropriate supervised machine learning techniques to detect breast cancer working with a microwave ultra-wideband (UWB) non-ionizing device inside the breast tissue. Different tissue characteristics within two breast models have been used to create two data sets. Non-ionizing microwave signals are sent to the breast model. The scattering parameters signals read by the antenna Represent the raw data fed into a machine model. The implemented metamaterial antennas with two or three gold rings and a square substrate succeeded in detecting the tumor. They worked under the frequency range of wifi and IOB, which means we can use the device at home. Thus, the patient can chick at home safely with no cost and early tumor detection. The results of the algorithms are tested and visualized in the R.stodio.

3. THE PROPOSED METAMATERIAL ANTENNA DESIGN

Metamaterials are homogenous materials or structures with properties not present in conventional materials that produce a negative effective permittivity [17]. A periodic artificial substance called a metamaterial has electronic characteristics that come from the structure of the material rather than its component components. Metamaterials have played a significant role in sensing technologies in recent years. Small, highly efficient antennas for wearable or implemented medical IOB and IoT Wireless devices may be designed using metamaterial architectures. The metal is bent into a square form with one or more little gaps known as SRRs. Metal post structures and periodic split ring resonators (SRRs) can be utilized to create materials with specific permeabilities and dielectric constants. [18].

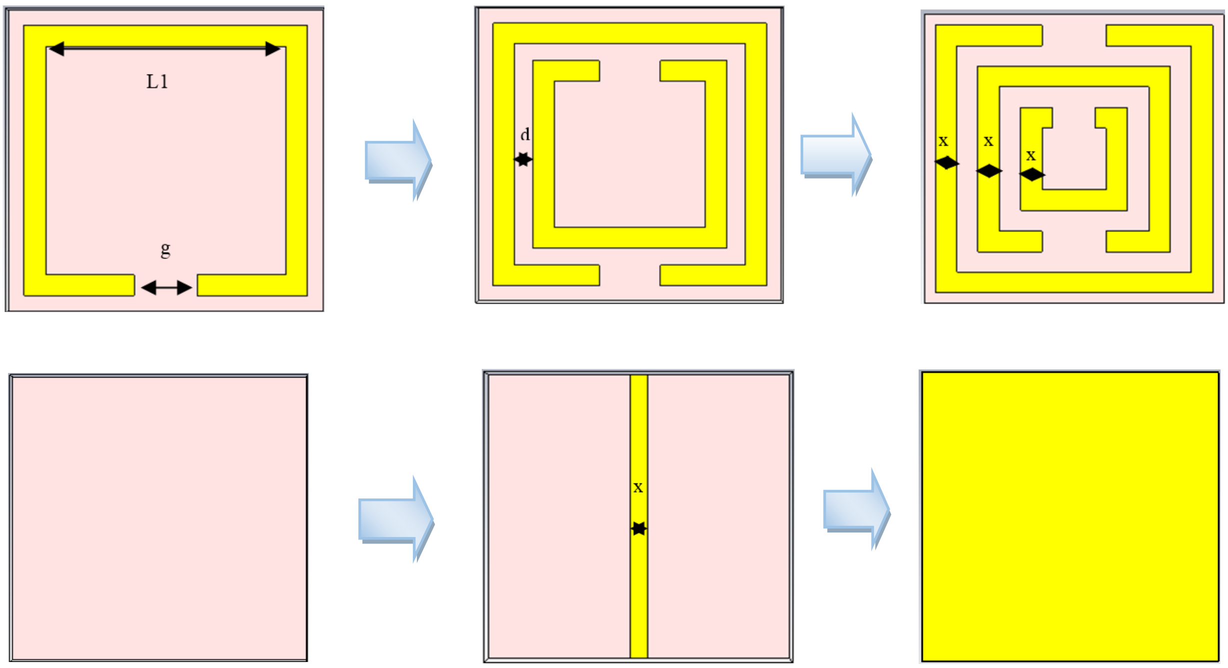
Recent research in designing antennas that match the biomedical application covers size reduction, bandwidth requirement, biocompatibility, radiation, and coupling effects. However, in this research, Multi-nontoxic materials are used in the antenna since it will be implanted in the breast tissue. The proposed metamaterial antenna was first created in open space using a Microwave studio simulator (CST) to check the antenna's performance and ensure it is working before being implanted inside the human body. Parameters are calculated and adjusted to keep the operating frequency within the UWB range.

3.1 ANTENNAS BEFORE IMPLANTING

Three distinct tiny sizes for metamaterial antennas were constructed in open space, each using a squared substrate and ring patches composed of gold with a thickness of 0.017mm. The thinnest substrate, 0.25 mm thick and with a permittivity of 11.68, was made of silicon. Thanks to the silicone, the antenna is now flexible and safe for human implantation. We use wire as the substrate ground for a pair of double-square SRR antennas (5mm x 5mm and 10mm x 10mm). The 7x7mm antenna was constructed using three rings and a square ground rather than wire. Fig 1 depicts these designs, and Tabel1 details the parameters used in their creation.

Table 1. parameters for antenna design

parameters	dimensions		
	5(mm)	10(mm)	7(mm)
Ts	5	10	7
d	0.25	0.25	0.25
L1	4.5	9	6.5
t	0.017	0.017	0.017
x	0.35	0.5	0.5
g	1	2	1.5



Step 1 **step 2** **step 3**

FIGURE 1. steps for antennas front and back design step 2 for antenna size (5,10)mm and step3 for (7)mm antenna size

The structure in Step 2 represents two of the proposed antennas, 5x5mm and 10x10mm, while step 3 represents a 7x7mm antenna. The effectiveness of the proposed metamaterial antenna is evaluated in open space (before implanted), where the antenna's propagation behavior changes when the antenna is implanted inside biological tissues. For open space, The gain and voltage standing waves may be used to evaluate the antenna's efficiency (VSWR). Optimal efficiency is achieved with a VSWR between 0 and 1.5 [9]. With a VSWR of 1.2657377, the proposed 5×5 mm metamaterial antenna easily meets these requirements. The antenna has a gain of -9.137 dBi and bandwidth of 0.4135 GHz with an s11 of -18.615129 (less than -10 dB). The recommended 7×7mm antenna size has a VSWR of 1.334565, a bandwidth of 0.21397, and a gain of -29.62 dBi, while the 10×10mm antenna has a VSWR of 1.557702, a bandwidth of 0.21396, and a gain of -12.20 dBi. Fig 2 shows the simulation result for the three antennas in open space.

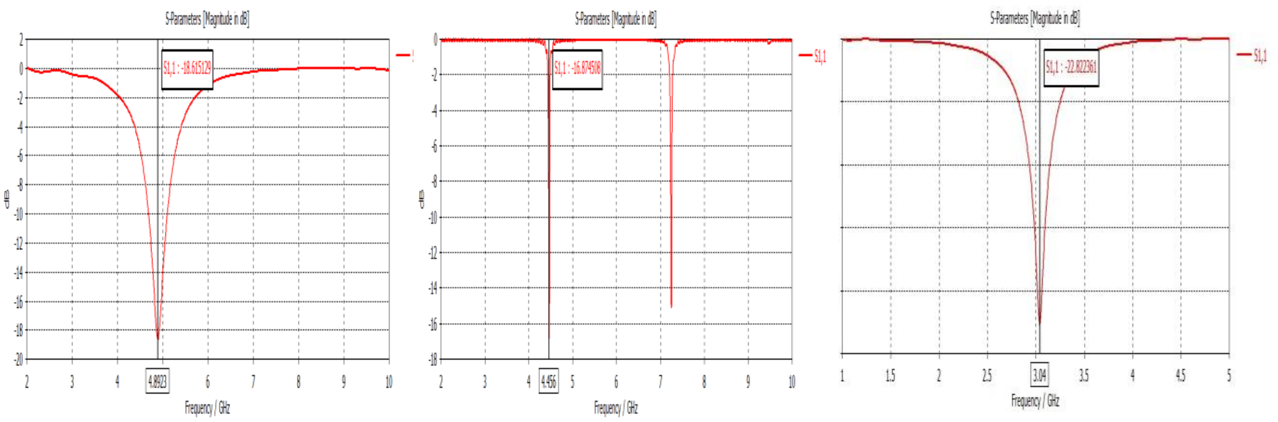


FIGURE 2. simulation result (s11 in open space) for (a) 5×5mm ;(b) 7×7mm ;(c) 10×10mm antennas

3.2 IMPLANTED ANTENNA IN BREAST PHANTOM MODELS

This section discusses the experimental setup for creating breasts and tumors, including an antenna. In addition, we cover the generation of the breast and tumor dataset necessary for ML processing. So the first stage of tumor detection is data creation.

When creating antennas, it is important to keep human tissues in mind, so this research used chipless metamaterial antenna for medical use with scientific characteristics that include comfortable shapes and dimensions that can be implanted inside the human body, as well as the use of non-toxic materials that do not have adverse effects on the body tissues and safe for the patients.

Diagnosis of breast cancer is affected because different biological tissues have varied conductivities and dielectric constants, leading to distinct RF (radio frequency) responses [19], The breast phantom is a communication medium that may be characterized by three electrical properties: relative permittivity, conductivity, and density. Layers of homogenous material are stacked on top of one another in the CST program to simulate an inhomogeneous environment. The breast is constructed using a layered model that includes skin, fat, and fibroglandular.

Since human tissues are frequency dependent, the major breast components’ dielectric characteristics will also differ. The range of characteristics is provided in the table2 of the breast’s main components with frequency band (200MG-5GHz)[20]. A specific absorption rate (SAR) is frequently used to evaluate heating problems since the spread of electromagnetic fields will increase tissue temperature in humans. SAR performance comparison criteria have been devised under safety regulations set by the Federal Communications Commission (FCC) and International Commission for Non-Ionizing Radiation Protection (ICNIRP). According to these standards, SAR values should not be more than 1,6 W/kg per one gram of tissue or 2 W/kg per ten grams of human tissue [21].

Standard low-power wireless devices do not generate sufficient amounts of power to be a source of worry for the SAR of the complete body. The SAR value is determined using the given equation 1:

$$SARi = \frac{Pi}{di} = \frac{\sigma|E|^2}{2di} \qquad (1)$$

(P) is the power loss density, which is measured in watts per meter cube; (di) is the density of human tissues, which is measured in kilograms per meter cube; (E) is the intensity of the electric field, which is measured in volts per meter; and σ is conductivity, which is measured in siemens per meter, in addition to this, and (i) denotes the number of tissue[22]. **Using the CST Microwave studio simulator,** two models were designed to create the patient data set for the breast phantoms model. Small model design with a breast radius (2.5cm) and a medium model design with a breast radius of (10 cm .the two models were designed Using a hemisphere breast model with three layers based on the conductivity (σ), density, and permittivity (ε), of the dielectric properties of breast tissues for each layer. The first is the skin layer, the second is the fat layer placed inside the skin layer, then the third layer is the fibroglandular placed inside the fat layer. For the small breast model, the skin layer with a center radius of 25 mm, the fat layer with a center radius of 23 mm, and the fibroglandular with a center radius of 15mm. In this model, 5mm, and 7mm antenna was used. We placed The antenna inside the fat layer, which rings towards the tumor. The gland is the layer that contains cancer; if any, a spherical shape of radius (2,3,4,6) mm is used as a tumor. Sar value for this model was 0.0583 w/k for 10 g of tissue

While the medium model employed a 10mm antenna in the fat layer, with a center radius of 48 mm and 8 mm for the thickness, the third layer, the fibroglandular, also has a center radius of 40 mm. The skin has a thickness of 2 mm and a center radius of 50 mm to cover the breast tissue completely. The spherical shape of radius (3,4,6,8) mm is used as a tumor to discover any tumors. We put the wave source (unit cell). The value of Sar for this model was 0.000852w/k for 10 g of tissue. as shown in Fig 3.

Table 2. **minimum and maximum boundaries for tissue information for the patient dataset**

Breast Phantom	Relative Permittivity ε	Conductivity σ
Fat tissue	1-10	0-1
Skin tissue	30-40	3-9
Glandular tissue	40-50	1.70-12

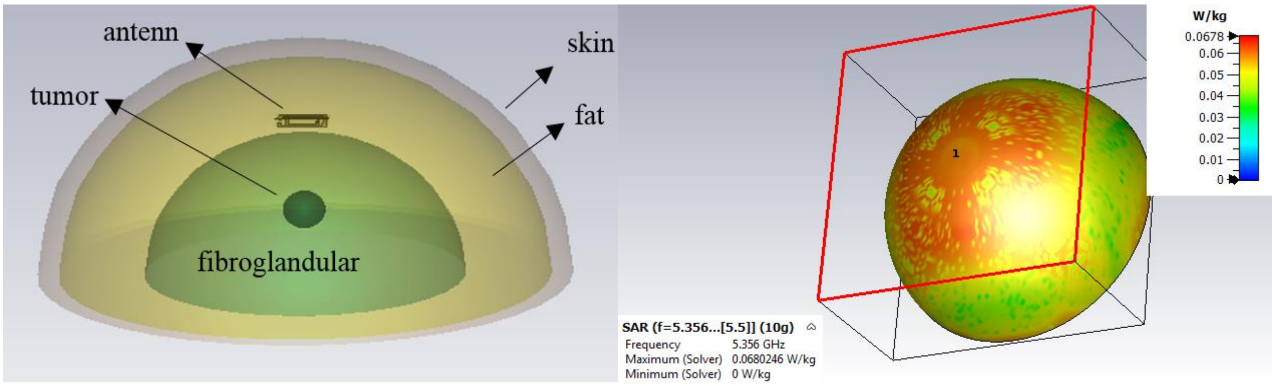


FIGURE 3. **breast model with one antenna inside the fat layer and SAR major**

4. SIMULATION RESULTS

This section shows some simulation results of s1-parameter signals before and after the tumor for both small and medium breast phantom models, as seen in fig 4.

Table 3. **result of the antenna implanted in small mode tissue (5cm)**

Tumor radius (mm)	Return loss S11(dB)	Resonant frequency(GHz)	Tumor Position (x,y,z)mm
Without tumor	-18.3353	5.473	---------
2	-17.5634	5.437	(0,0,-10)
3	-16.4872	5.365	(0,0,-10)
3	-16.6807	5.352	(0,0,-11)
3	-16.9391	5.352	(0,0,-13)
4	-15.7487	5.374	(0,0,-10)
6	-16.4144	5.401	(0,0,-10)

Table 4. **result of the antenna implanted in medium model tissue (10) cm**

Tumor radius(mm)	Return loss S11(dB)	Resonant frequency(GHz)	Tumor Position (x,y,z)mm
Without tumor	-35.5351	7.129	----------
2	-17.5634	5.437	(0,0,-10)
3	-38.7248	7.128	(0,0,-25)
3	-38.5501	7.128	(10,10,-30)
3	-38.6408	7.128	(8,20,-15)
4	-38.8707	7.128	(0,0,-25)
6	-38.8247	7.128	(0,0,-25)
8	-37.9421	7.128	(0,0,-25)

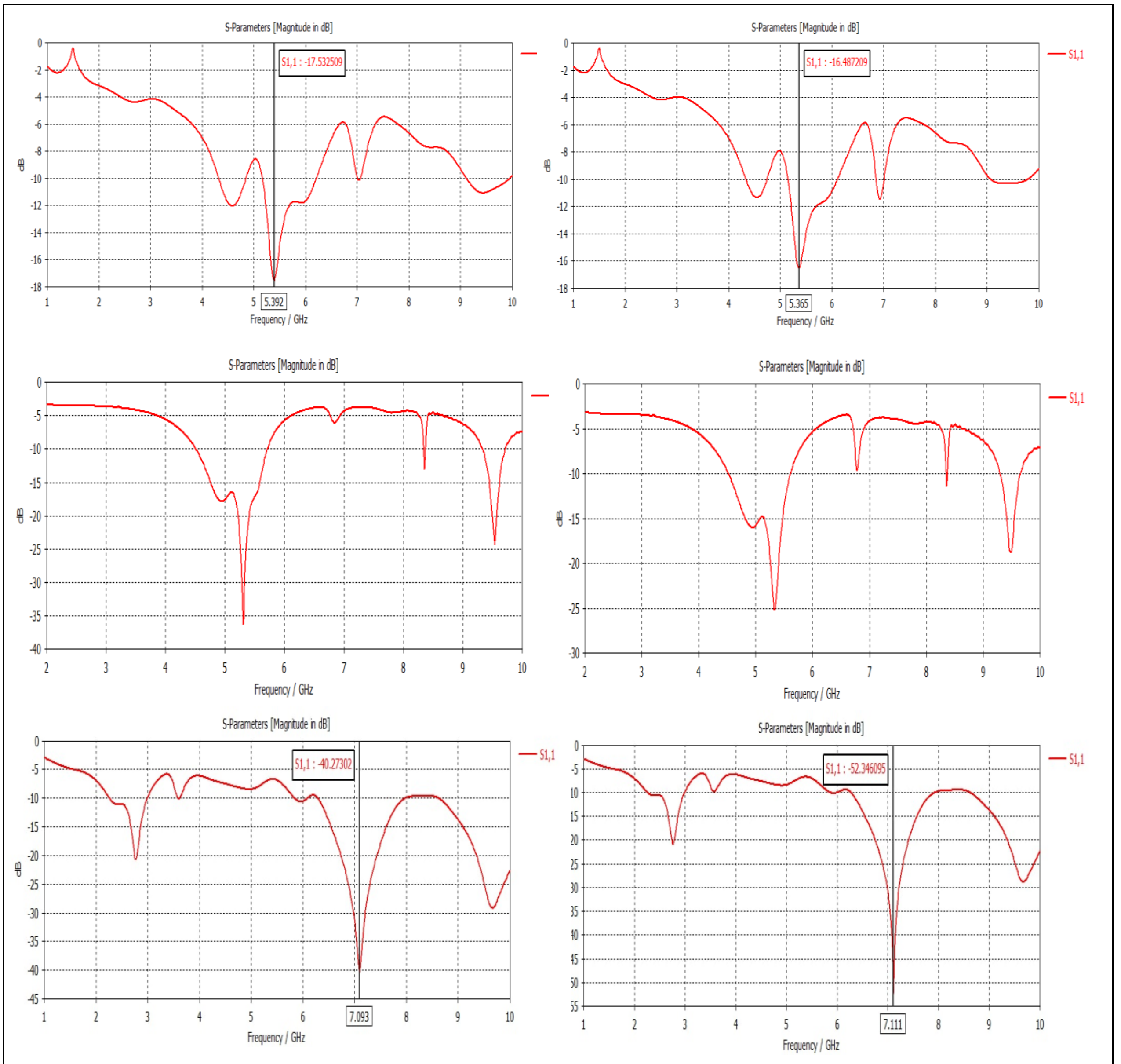


FIGURE 4. **simulation of s-parameters (Received signals reflected from the breast without tumor and breast containing the tumor**

The transmitter transmits signals and lights the phantom breast model, while the reception antenna gathers the breast phantom model's backscattered waves. Signals resulting from backscattering are documented because the difference between tumor-free and tumor-affected signals will mack the machine learning model able to diagnose the tumor signals.

After extracting the breast models simulation data, data signals need to be preprocessed and converted into exile files. The group of data patients has been created. This data has 150 patients, with 1001 attributes as the independent variable. In principle, a machine learning algorithm trained on a particular dataset should be able to be generalized to new, untrained datasets. The data used to train a model is typically different from the data used to test it. The training set should be as big as feasible to minimize variance in training the model. However, the unseen portion of the data should also be close to the original dataset, so the evaluation process is relevant. When dealing with ML classification, the crucial issue is not whether one learning algorithm is better than another but rather under what situations one technique might considerably outperform another on a specific application problem[23].

5. MACHIN LEARNING CLASSIFICATION

The process of detecting tumors has now moved on to the second stage. In the data processing stage, some machine learning approaches are used to create models for determining whether or not a breast tumor is present or absent using just the backscatter signals gathered from the antennas. divide the process of creating a pattern classification model into three stages: Data preparation(preprocessing), feature selection and extraction, and classification algorithms are all steps in the classification process. [14].

5.1 DATA PREPROCESSING

Multiple uses of machine learning may be found in the healthcare industry. Healthcare data are generally large, disorganized, and complex. Before they can be utilized to train machine learning systems, data must be properly mapped and preprocessed. This step is an important point in machine learning model development since the quality of the model is heavily reliant on the data's correctness in reflecting clinical reality. [24]. Data first needs to be properly labeled and processed because this data will be used to train a machine learning system. The labels must correctly represent clinical reality. Regardless matter how much work is put into enhancing the machine learning algorithm, any mistake in labeling will greatly restrict the accuracy that may be achieved. So the raw data samples for both groups labeled the category label of a tumor-contained signal as one and a tumor-free signal as zero. Most machine learning methods need to normalize the data. [25] . Data normalization is a technique for standardizing the range of characteristics without affecting the data's dimension. Important is the data normalization procedure since it is necessary to choose the best characteristics without removing relevant information from the curated data. [11]. it solves two major data problems that restrict the learning process of machine learning algorithms: dominating features and outliers [26]. Several techniques have been developed for Data normalization. This study uses Z-score (zs) and min-max (mm) to standardize raw data samples. The train and test datasets are created using a 3:1 division of the normalized data.

5.2 FEATURE EXTRACTION AND SELECTION

the algorithms for feature selection and extraction are quite important. These approaches, if used correctly, may choose the most useful, non-redundant data from electromagnetic waves, making the task of training the classifier faster. The objective is to choose variables or combine them into features, which will minimize the quantity of data that must be processed (known as "dimensional reduction") while retaining the original database's significance.

The principal component analysis is the breast tumor community's most widely utilized feature extraction approach. The data are transformed using PCA into a new subdomain, where the principal components with the highest variance are the new orthonormal basis. It locates the components that make it easier to tell if a breast tumor is present. Dimensional reduction is possible by selecting the most important primary components since PCA organizes them according to their relative significance., i.e., the features with the most variation [27]. When doing dimensional reduction, one of the criteria that must be considered is the number of features used. (function **pc** from R studio is utilized for feature extraction, reducing the dataset's dimensions from 1001 features to 121. PCA datasets always reveal better performance than the original dataset.

5.3 CLASSIFICATION ALGORITHMS

The purpose of classification algorithms is to study how to connect processed data with each class to determine the proper class when new data has to be evaluated. This study has data sets representing a different group of patients has 150 instances with 1001 attributes as the independent variable and one as the dependent variable for the analysis .k-nearest neighbor (KNN), decision trees (DT), and support vector machines (SVM) were used to train these data.

- support vector machine (SVM)

The term “SVM” refers to a particular supervised machine learning method. (data utilized is labeled). SVM creates a hyper-plane with a margin that is as broad as is feasible for splitting various data categories or keeping comparable data from one class on one side of the margin and similar data from another class on the other [28].

SVM can be used to make predictions and classify data. When predicting a breast tumor, disease classes can be split in a hyperplane by the chances of developing tumor tissue on one side of the margin and normal tissue (without a tumor) on the other side. There are two types of SVM: linear and nonlinear SVM. Data can be separated using a line in linear SVM, but nonlinear SVM is used when this is impossible. For nonlinear SVM, a kernel function is used. Different classes of data that may be divided nonlinearly are transferred onto higher - dimensional when applying a kernel function to make them linearly separable. Equation (2) is the kernel function.

$$k\left(z_i, z_j\right) = \emptyset(z_i)\emptyset\left(z_j\right) \qquad (2)$$

Sigmoid kernel function, polynomial, linear, and radial basis functions are examples of several kinds of kernel functions. Kernel functions take incoming data and turn it into a needed format. Kernel functions have the ability to convert an input space with fewer dimensions into one with many dimensions. [28].When multi-collinearity and a nonlinear connection between the input and output variables, SVMs can perform very effectively. In most cases, the performance of SVMs is significantly improved when they are allowed to work with many dimensions and continuous variables. [23]

- K-Nearest Neighbor (KNN)

KNN, for short, is a supervised method for machine learning that may be applied to classification and regression issues. A straightforward procedure saves all samples that are provided and separates them into new classes by applying a measure of similarity (e.g., distance functions).[29]. Here is how the algorithm works. The majority vote of a case’s neighbors determines its class. When compared the case to its K closest neighbors, as determined by a distance function, it is placed in the most prevalent class. The case is placed in the class of the case that is the most similar to it if K = 1. For distance computations and determining the nearest data points, the Euclidean, Manhattan, and Minkowski separation distances are frequently used.[29]. However, there is no assurance that a big K value is more accurate since it decreases overall noise.). If a data spot’s target value is unavailable in KNN, the k most similar data spots in the training set are found, and their average value is provided. Despite the noise and size of the data, It could offer accurate outcomes and predictions. According to the majority vote of KNN’s K-nearest neighbors, new data (new backscatter signal) is classified by KNN based on the knowledge of its known classes [30]. These important parameters must be selected and optimized while developing the model: The distance utilized to determine which neighbors are extremely close is D, and K is the number of nearest neighbors. [31].

- Decision Trees (DT)

A supervised learning method that may be used to categorize and perform regression on the data is a decision tree. The root is divided into sub-trees or additional branches based on the maximum information gain and certain criteria once the entropy value for each feature is calculated.[32]. With more features, the procedure recursively proceeds, and the leaf node will provide the eventual outcome.

$$entropy(z) = -\sum_{i=1}^{m} p(z_i) log_n\, p(z_i) \qquad (3)$$

Class entropy is determined using equation (3), where m represents a set of classes in Z, Z represents the current data set to be computed, and p(z) represents the ratio of the number of items in class m to the set of items in set z.

$$IN\ (Gain) = E\ (class) - E\ (features) \qquad (4)$$

Where IN represents information and E represents entropy. For information gain, each branch’s entropy is subtracted from the entropy of the class in equation (4). Two types of data may be utilized in decision trees: continuous and categorical. The tree gave good accuracy since every characteristic was evaluated and analyzed.

Models are simple to learn, and rules are simple to develop, able to replace missing values in variables with the value that has the highest probability of being correct. If a tree contains many branches, this technique might suffer from over-fitting, making it difficult to calculate and understand. [33].

The overview of the classification system may be observed in Fig 5.

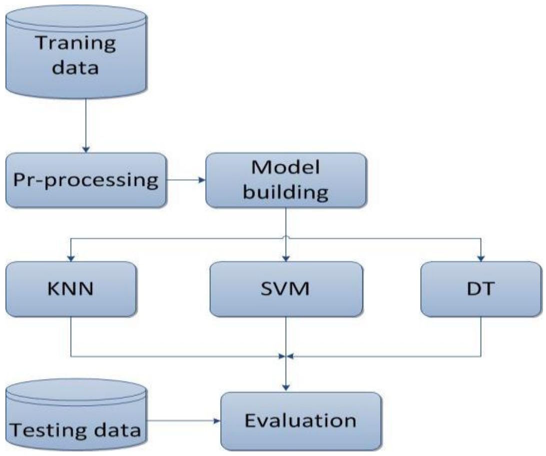


FIGURE 5. An overview of the proposed system

Accurate, specific, and sensitive measurements—which must be computed to avoid categorization errors—are necessary for accurate breast cancer early detection. When a breast is checked for malignancies at a preliminary phase, a significant rate of misdiagnosis is probably possible. If malignancies are there, but the algorithm fails to detect them, or if there are no malignancies. However, the algorithm still determines that a tumor is present, known as misdiagnosis. The system's overall efficiency would be harmed if this possibility exists; hence it must be avoided or decreased. So for the result, we calculate the flowing:

The Kappa Statistic: is a performance measure that contrasts actual and expected accuracy [23].
Accuracy was calculated using a confusion matrix, which shows where the model is getting confused [34], as shown in Fig 6.

Predicted / Real	FREE-TUMOR	WITH TUMOR
FREE-TUMOR	TN	FP
WITH TUMOR	FN	TP

FIGURE 6. the structure of the confusion matrix

$$Accuracy = \frac{((TP + TN)}{(TP + TN + FP + FN)} \quad (5)$$

Figure 6 shows the machine model's predicted class and real class, where TP stands for "true positive," which denotes that cancer has been correctly categorized. FN stands for "false negative," which denotes that the prediction of not having breast cancer was wrong. At the same time, FP stands for "false positive," meaning breast cancer was misclassified. TN stands for "true negative," which shows the right way to classify cancer that does not exist.

Recall (Sensitivity): Recall tells us how many true positives are correctly described. Equation (6) can be used to define recall.

$$Recall\,(sensitivity) = \frac{TP}{TP + FN} \qquad (6)$$

Precision (Specificity): Precision tells us how many of the correctly predicted identifications there are. We can explain what precision is by using equation (7).

$$Precision\ (Specificity) = \frac{TP}{TP + FP} \qquad (7)$$

The F-score, sometimes referred to as the F1-score, is a measure of a model's accuracy on a certain dataset. It has been applied to evaluate binary classification methods that categorize occurrences into good and bad categories[35]. As a method to combine both, as in equation (8), it is defined as the harmonic mean of the model's precision and recall.

$$F - score = \frac{2PR}{P + R} \qquad (8)$$

6. RESULT AND DISCUSSION

The comparison outcomes of the classification algorithms are displayed in Table 5, which shows an obvious variation in the accuracy between these algorithms for the same patient's dataset.

Table 5. result information for the patient data set

Algorithm performance	SVM				KNN	DT
	linear	radial	polynomial	sigmoid		
accuracy	**93%**	63 %	70%	**%93**	63 %	73 %
kappa	0.85	0.046	0.129	0.85	0.046	0.4872
precision	0.95	0.947	1	0.95	0.947	0.611
recall	0.95	0.642	0.689	0.95	0.642	0.916
f-score	0.95	0.765	0.815	0.95	0.765	0.733

All algorithms have been evaluated and compared based on how well they work. Accuracy, kappa, precision, recall, and f-score are terms used to describe the results of the three algorithms. since it is known that the larger the kappa precision, recall, and F-scores are, the better the outcome, we can observe that SVM has the best result for detecting breast cancer, with an accuracy of 93%, F-score coming out to be 95%, and kappa is 85% for sigmoid, radial, and linear SVM. This result means that the antenna can detect the tumor in a tissue in which it was previously implanted, even when data signals were overlapped.

SVM has performed better than all the other algorithms in terms of accuracy, F-score, and other evaluations. So, according to the results obtained, we have concluded that the SVM algorithm with the three kernels (sigmoid, radial, and linear) predicts this research's best result. SVM's highest result may have been its ability to handle high-dimensional data. The polynomial SVM classifier is less than in other cases. It can be seen that the accuracy outcome is 70%, and F-score is 81%, which means it is not very suitable for separating the data. The DT saw the next good result with an accuracy of 73%, but it is still not very good, while KNN does not work well in this data

7. CONCLUSION

This research proposes an effective early breast cancer detection method. The non-ionizing implemented antennas that successfully detected the tumor worked under the frequency range of wifi and IOB, which means the device can be used at home. This method uses a unique feature extraction technique under UWB microwave and two distinct breast model types to create patient data sets. This study created acceptable data for the machine intelligence system using the UWB-implemented antenna. Many applicable supervised approaches for machine learning, such as closest neighbor and support vector machines, Decision trees, are trained and investigated using this data.

The machine learning method (SVM) delivered excellent accuracy results at 93%. The group of patients was taken from frequencies commensurate with the antenna. The explanation for SVM's highest performance may have been because it performs well with high-dimensional data. That implies that the antenna used to detect the tumor worked efficiently.

REFERENCES

[1] T. Kim, J. Oh, B. Kim, J. Lee, S. Jeon, and J. Pack, "A study of dielectric properties of fatty, malignant and fibro-glandular tissues in female human breast," in *2008 Asia-Pacific Symposium on Electromagnetic Compatibility and 19th International Zurich Symposium on Electromagnetic Compatibility*, 2008, pp. 216-219: IEEE.

[2] K. Nahalingam and S. K. Sharma, "An investigation on microwave breast cancer detection by ultra-widebandwidth (UWB) microstrip slot antennas," in *2011 IEEE International Symposium on Antennas and Propagation (APSURSI)*, 2011, pp. 3385-3388: IEEE.

[3] U. O. Matthew and J. S. Kazaure, "Chemical polarization effects of electromagnetic field radiation from the novel 5G network deployment at ultra high frequency," *Health and Technology,* vol. 11, no. 2, pp. 305-317, 2021/03/01 2021.

[4] M. M. Dawoud, "High frequency radiation and human exposure," in *Proceedings of the International Conference on Non-Ionizing Radiation at UNITEN (ICNIR 2003)*, 2003, pp. 1-7.

[5] B. L. Oliveira, D. Godinho, M. O'Halloran, M. Glavin, E. Jones, and R. C. Conceição, "Diagnosing Breast Cancer with Microwave Technology: Remaining challenges and potential solutions with machine learning," *Diagnostics,* vol. 8, no. 2, p. 36, 2018.

[6] M. Fatima and M. Pasha, "Survey of machine learning algorithms for disease diagnostic," *Journal of Intelligent Learning Systems and Applications,* vol. 9, no. 01, p. 1, 2017.

[7] M. Alibakhshikenari *et al.*, "Metamaterial-inspired antenna array for application in microwave breast imaging systems for tumor detection," *IEEE Access,* vol. 8, pp. 174667-174678, 2020.

[8] V. Selvaraj, P. Srinivasan, D. Baskaran, and R. Krishnan, "Characterisation of breast tissue using compact microstrip antenna," *International Journal of Biomedical Engineering and Technology,* vol. 31, no. 2, pp. 161-175, 2019.

[9] G. Geetharamani and T. Aathmanesan, "Metamaterial inspired THz antenna for breast cancer detection," *SN Applied Sciences,* vol. 1, no. 6, pp. 1-9, 2019.

[10] J. Boparai and M. Popović, "Heterogeneous Skin Phantoms for Experimental Validation of Microwave-Based Diagnostic Tools," *Sensors,* vol. 22, no. 5, p. 1955, 2022.

[11] V. Vijayasarveswari *et al.*, "Multi-stage feature selection (MSFS) algorithm for UWB-based early breast cancer size prediction," *PloS one,* vol. 15, no. 8, p. e0229367, 2020.

[12] B. S. Bari *et al.*, "Ultra wide band (uwb) based early breast cancer detection using artificial intelligence," in *InECCE2019*: Springer, 2020, pp. 505-515.

[13] N. Ali, M. A. Sree, R. Uyguroglu, and A. Allam, "Stage II cancer diagnosis using printed antenna implemented on hemispherical model for human breast," *Journal of Instrumentation,* vol. 15, no. 09, p. P09016, 2020.

[14] R. C. Conceição *et al.*, "Classification of breast tumor models with a prototype microwave imaging system," *Medical Physics,* vol. 47, no. 4, pp. 1860-1870, 2020.

[15] M. Y. Azab, M. F. O. Hameed, A. M. Nasr, and S. Obayya, "Highly sensitive metamaterial biosensor for cancer early detection," *IEEE Sensors Journal,* vol. 21, no. 6, pp. 7748-7755, 2021.

[16] A. Celik, K. N. Salama, and A. M. Eltawil, "The internet of bodies: A systematic survey on propagation characterization and channel modeling," *IEEE Internet of Things Journal,* vol. 9, no. 1, pp. 321-345, 2021.

[17] B. Nasiri, A. Errkik, J. Zbitou, A. Tajmouati, L. El Abdellaoui, and M. Latrach, "A compact miniature microstrip low pass filter based on dumbbell DGS and SRR unit cell," *International Journal of Microwave and Optical Technology,* vol. 13, no. 1, pp. 32-39, 2018.

[18] A. Sabban, "New compact wearable metamaterials circular patch antennas for IoT, medical and 5G applications," *Applied System Innovation,* vol. 3, no. 4, p. 42, 2020.

[19] H. Bahrami, B. Gosselin, and L. A. Rusch, "Realistic modeling of the biological channel for the design of implantable wireless UWB communication systems," in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2012, pp. 6015-6018: IEEE.

[20] S. S. Tiang, M. S. Hathal, T. F. Zanoon, M. F. Ain, and M. Z. Abdullah, "Radar sensing featuring biconical antenna and enhanced delay and sum algorithm for early stage breast cancer detection," *Progress in electromagnetics research B,* vol. 46, pp. 299-316, 2013.

[21] S. Kiani, P. Rezaei, and M. Fakhr, "A CPW-fed wearable antenna at ISM band for biomedical and WBAN applications," *Wireless Networks,* vol. 27, no. 1, pp. 735-745, 2021.

[22] A. A. Aly and M. Piket-May, "FDTD computation for SAR induced in human head due to exposure to EMF from mobile phone," *Advanced Computing,* vol. 5, no. 5/6, p. 1, 2014.

[23] F. Osisanwo, J. Akinsola, O. Awodele, J. Hinmikaiye, O. Olakanmi, and J. Akinjobi, “Supervised machine learning algorithms: classification and comparison,” *International Journal of Computer Trends and Technology (IJCTT)*, vol. 48, no. 3, pp. 128-138, 2017.

[24] K. Y. Ngiam and W. Khor, “Big data and machine learning algorithms for health-care delivery,” *The Lancet Oncology*, vol. 20, no. 5, pp. e262-e273, 2019.

[25] S. Nersisyan *et al.*, “ExhauFS: exhaustive search-based feature selection for classification and survival regression,” *PeerJ*, vol. 10, p. e13200, 2022.

[26] D. Singh and B. Singh, “Investigating the impact of data normalization on classification performance,” *Applied Soft Computing*, vol. 97, p. 105524, 2020.

[27] E. O. Omuya, G. O. Okeyo, and M. W. Kimwele, “Feature selection for classification using principal component analysis and information gain,” *Expert Systems with Applications*, vol. 174, p. 114765, 2021.

[28] R. Katarya and S. K. Meena, “Machine learning techniques for heart disease prediction: a comparative study and analysis,” *Health and Technology*, vol. 11, no. 1, pp. 87-97, 2021.

[29] R. A. Nugrahaeni and K. Mutijarsa, “Comparative analysis of machine learning KNN, SVM, and random forests algorithm for facial expression classification,” in *2016 International Seminar on Application for Technology of Information and Communication (ISemantic)*, 2016, pp. 163-168: IEEE.

[30] P. Flach, *Machine learning: the art and science of algorithms that make sense of data*. Cambridge university press, 2012.

[31] K. Murphy, “Machine learning: a probabilistic perspective. Cambridge, Mass.[ua],” ed: MIT Press, 2013.

[32] A. S. Karthiga, M. S. Mary, and M. Yogasini, “Early prediction of heart disease using decision tree algorithm,” *International Journal of Advanced Research in Basic Engineering Sciences and Technology (IJARBEST)*, vol. 3, no. 3, pp. 1-17, 2017.

[33] S. Ray, “A quick review of machine learning algorithms,” in *2019 International conference on machine learning, big data, cloud and parallel computing (COMITCon)*, 2019, pp. 35-39: IEEE.

[34] F. J. Ariza-López, J. Rodriguez-Avi, and M. Alba-Fernandez, “Complete control of an observed confusion matrix,” in *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*, 2018, pp. 1222-1225: IEEE.

[35] S. Mian Qaisar and S. F. Hussain, “An effective arrhythmia classification via ECG signal subsampling and mutual information based subbands statistical features selection,” *Journal of Ambient Intelligence and Humanized Computing*, pp. 1-15, 2021.