Wearable IoT-Based Breast Tumor Classification Using Raspberry Pi and Hybrid ML/DL Models

Harshit Manik[1a], Mantra Gupta[1b], Shaurya Punj[1c], Dr. Amanpreet Kaur[1d]

[1]Department of Electronics and Communication Engineering, Thapar Institute of Engineering and Technology, Patiala, India

[1a]hmanik_be23@thapar.edu, [1b]mgupta5_be23@thapar.edu,
[1c]spunj_be23@thapar.edu, [1d]amanpreet.kaur@thapar.edu

Abstract*— Across the globe, breast cancer is a major health issue responsible for a significant number of deaths among women. Traditional screening methods rely on periodic mammograms, which may lead to delayed detection. This paper proposes a wearable IoT-based tumor detection system integrated into women innerwear, utilizing Raspberry Pi and hybrid deep learning models for an early a more accurate detection of breast cancer. The system employs multimodal wearable antennas for real-time tissue analysis and cloud-based monitoring. The proposed solution substantially increased accuracy from traditional machine learning models and produced precision of 97% thereby reducing false positives to an extent where it can be used for real-time classification applications i.e. monitoring and reducing dependence on hospital-based screenings.

***Keywords*—** *Breast Cancer, Wearable IoT, Deep Learning, Raspberry Pi, Tumor Classification,*

I. Introduction

Breast cancer remains one of the most prevalent and life-threatening diseases among women worldwide. Early detection significantly improves survival rates, yet conventional diagnostic methods such as mammography, ultrasound, and MRI require clinical visits, specialized radiologists, and costly equipment. These limitations create accessibility challenges, particularly in rural and low-resource regions where medical imaging facilities are scarce.

Early detection of breast cancer significantly improves patient survival rates and treatment outcomes. Traditional diagnostic methods such as mammography, ultrasound, and MRI require clinical visits and specialized radiologists, making them less accessible in remote or resource-limited settings. Recent advancements in the Internet of Things (IoT) using hardwares like Raspberry Pi and Jetson Nano Machine learning having advance regression and classification tools have introduced innovative approaches to breast cancer detection, enabling real-time, automated, and cost-effective screening solutions. IoT-based wearable devices integrated with smart sensors can continuously monitor breast tissue abnormalities, while deep learning algorithms analyse medical images to enhance diagnostic accuracy [1], [2].

Recent advancements, such as Sycai Medical's AI-based tool, have demonstrated the potential of non-invasive imaging techniques in early cancer detection[3]. The application of UWB technology in wearable devices has shown promise in accurately tracking and classifying physical activities, which could be extended to health monitoring applications[4].

Machine learning models, particularly Convolutional Neural New hybrid models, such as CNN combined with optimization techniques, significantly improve detection performance over traditional methods [3], [4]. Additionally, secure data transmission using blockchain and cloud computing ensures privacy in medical IoT applications, making these systems more viable for real-world deployment [5], [6]. (CNNs) and ensemble classifiers, have been extensively used to process thermal images, ultrasound scans, and other medical data for accurate tumour classification. Research has demonstrated that the datasets used in these studies vary widely, with many sourced from public repositories such as Kaggle and the UCI Machine Learning Repository. These datasets include labelled mammograms, histopathological images, and sensor- based readings from wearable devices. By combining diverse datasets and leveraging machine learning for feature extraction and classification, researchers have demonstrated promising results in breast cancer screening and diagnosis [4], [6].

Several studies have explored the potential of IoT-enabled breast cancer detection systems. A study presented an IoT-based smart healthcare system incorporating a Feedback Artificial Crow Search-based Shepherd CNN, optimizing classification accuracy in breast tumour identification [2]. Another research project integrated deep learning with IoT in breast cancer detection, demonstrating how real-time monitoring and automated alerts enhance early diagnosis [8]. The use of smart thermography systems based on bioheat transfer simulation has also been explored, providing non-invasive breast cancer detection using wearable sensors [7].

This study proposes a predictive healthcare system that leverages IoT and Electronic Health Records (EHRs) for real-time, automated breast cancer detection. The framework integrates thermal imaging cameras with Convolutional Neural Networks (CNNs) to improve diagnostic precision. Researchers have also investigated 3D-printed breast prostheses capable of sensing and treating early-stage breast cancer, further expanding the applications of IoT-integrated therapeutic solutions [10].

Various machine learning techniques, including Logistic Regression, Random Forest, SVM, KNN, Decision Tree, Gradient Boosting, AdaBoost, and a Voting-based ensemble model, have been incorporated in this study. Additionally, a CNN+LSTM model has been implemented to further boost accuracy. The models were tested on a diverse breast tumour dataset from Kaggle, allowing for extensive performance evaluation in different clinical scenarios.

Thus this paper proposes a wearable breast cancer detection system leveraging Raspberry Pi for real-time processing and cloud-based data transmission for remote monitoring. This solution offers a low-cost and scalable framework that can function effectively across various settings, presenting a viable method for early tumor detection and ongoing patient health monitoring.

II. PROPOSED SYSTEM DESIGN

To overcome these limitations, our proposed approach incorporates the following enhancements:

- *Diverse Model Selection*: Instead of solely relying on complex models, we tested both lightweight and heavy machine learning models to balance accuracy, learning time, and memory efficiency. This approach ensured a robust comparative analysis between hybrid and simple models.
- *Advanced Data Preprocessing*: We applied advanced preprocessing techniques, including feature scaling and normalization, to enhance the model performance. Principal Component Analysis (PCA) was utilized to minimize feature redundancy and lower dimensionality, thereby optimizing the computational performance. *Hyperparameter Tuning*: To ensure optimal performance, hyperparameter tuning was conducted using Grid Search and Randomized Search techniques. This allowed us to find the best parameters for each model, enhancing accuracy and generalization.
- *Improved Feature Extraction*: By integrating CNN with LSTM, we utilized a hybrid deep learning architecture to better capture spatial and temporal dependencies. This approach significantly improved classification performance by capturing both spatial dependencies and temporal correlations within the dataset.

.

- *Implementation of Ensemble Learning*: Incorporating techniques such as Random Forest, Gradient Boosting, and XGBoost to improve classification robustness.
- *Cross-Dataset Generalization:* Evaluating the model on multiple datasets to assess its adaptability in real-world clinical applications.

By implementing these enhancements, our proposed model aims to achieve superior accuracy, better generalization, and improved real-world applicability for early breast cancer detection

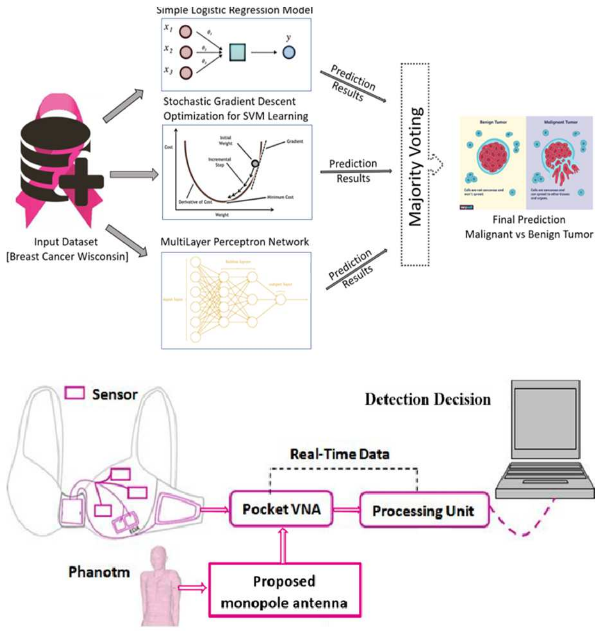


Fig-1 A wearable IOT based -breast monitoring system

-Machine Learning-Based Breast Tumor Classification

We implemented multiple traditional machine learning (ML) models to classify breast tumour data.

The deep learning model was trained and assessed using the Wisconsin Diagnostic Breast Cancer (WDBC) dataset, consisting of 569 samples of digitized fine needle aspiration (FNA) images of breast tumors. Each sample includes 30 numerical features derived from characteristics of the cell nuclei, such as radius, texture, perimeter, area, smoothness, compactness, concavity, and fractal dimension. The dataset also contains a binary classification label: 'M' for malignant (212 samples) and 'B' for benign (357 samples) [17]. The dataset is diverse and representative, encompassing a wide range of tumor characteristics in terms of shape, size, and structure. Each record corresponds to a distinct patient, ensuring data independence. To maintain consistency and reduce bias, all features were normalized before feeding into the model.

In order to test the model's generalization capability and minimize overfitting, a thorough evaluation was conducted., we employed **5-fold stratified cross-validation**. This approach ensures that each fold preserves the original proportion of benign and malignant samples. The model underwent training on four data subsets, with the remaining subset used for validation in each iteration. Performance metrics, including accuracy, precision, recall, and F1-score, were calculated for each fold and averaged to provide a robust estimate of the model's performance across all folds. The dataset was pre-processed, standardized, and divided into separate training and testing sets.. The methodology involved:

To ensure optimal model performance and mitigate bias, we performed several data preprocessing steps, including feature selection, encoding, scaling, and handling missing values. The steps are detailed below:

TABLE I

Learning Model	Accuracy During Training	Accuracy on Test Data
Logistic Regression	0.9868	0.9737
Random Forest	1.0000	0.9649
Support Vector Machine	0.9868	0.9737
K-Nearest Neighbors	0.9802	0.9474
Decision Tree	1.0000	0.9298
Gradient Boosting	1.0000	0.9561
AdaBoost	1.0000	0.9649
Voting Classifier	0.9912	0.9737

ACCURACY COMPARISON OF ML MODELS

Data preprocessing: Removal of unnecessary features and encoding of categorical labels. Feature scaling: Standardization using a Standard Scaler.

Feature Selection and Removal of Redundant Columns The dataset contained an id column, which was a unique identifier for each patient but had no relevance to classification. This column was removed as it did not contribute to predictive accuracy. Additionally, highly correlated features (correlation coefficient > 0.9) were analyzed to reduce multicollinearity and redundancy

Model training: Implementation of multiple ML models.

Performance evaluation: Accuracy comparison between training and testing sets.

A) Dataset and Preprocessing

We utilized the breast cancer dataset from Kaggle Breast Tumour dataset. The preprocessing steps included:

- Dropping the non-informative id column.
- Encoding the target variable (diagnosis) using Label Encoding.
- Dividing the dataset into 80% for training and 20% for testing. Standardizing features using StandardScaler.

B)

The models used for training and testing in this study are as follows:"

- *Logistic Regression*
- *Support Vector Machine (SVM)*
- *K-Nearest Neighbors (KNN)*
- *Decision Tree*
- *Random Forest*
- *Gradient Boosting*
- *AdaBoost*

A. Performance Evaluation

The models were evaluated based on training and testing accuracy, summarized in Table I.

A. Deep Learning Approach

1) Model Architecture: A recent study demonstrated that a CNN-LSTM model achieved superior accuracy (99.17% and 99.90% across two datasets) compared to standalone architectures like ResNet-50 and VGG-16. This hybrid model effectively balances accuracy and computational efficiency, making it suitable for real-time clinical applications. We implemented a Convolutional Neural Network (CNN) followed by a Long Short-Term Memory (LSTM) network to improve classification accuracy. The model consists of:

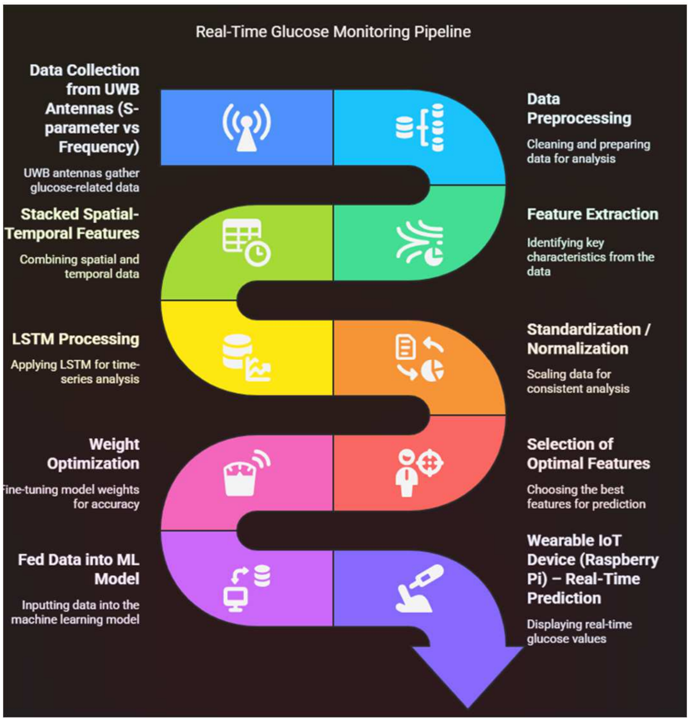


Fig. 2. System Architecture of the Wearable IoT-based Tumor Classification System

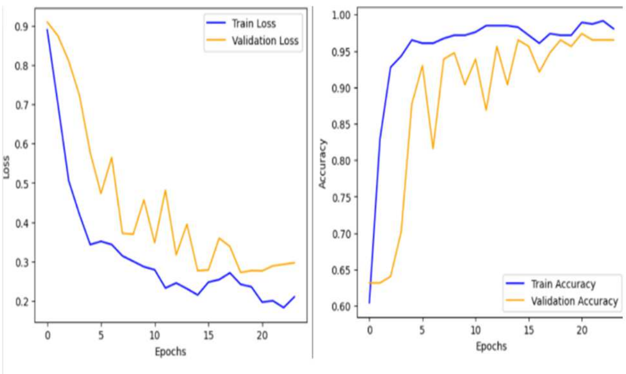
- Convolutional layers for feature extraction.
- LSTM layers to capture temporal dependencies.
- Fully connected layers for final classification.

2) Overfitting Prevention Strategies: To mitigate overfit- ting, we applied:

- Dropout layers (0.6) to reduce reliance on specific neurons.
- Batch normalization to stabilize training.
- Early stopping and learning rate reduction.

B. Performance Comparison

Figure 3 illustrates the classification performance of proposed models.



Classification Report:

	precision	recall	f1-score	support
0	0.97	0.97	0.97	72
1	0.95	0.95	0.95	42
accuracy			0.96	114
macro avg	0.96	0.96	0.96	114
weighted avg	0.96	0.96	0.96	114

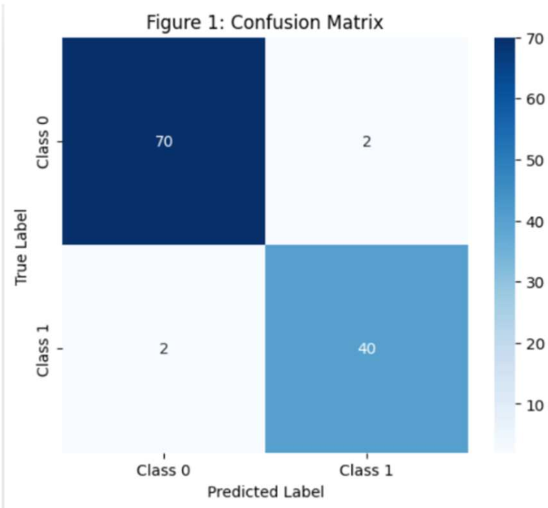


Fig. 3.A Summary of Classification Model Performance
Fig.3b Classification report of the Deep Learning model
Fig:3c-Confusion matrix of predicted vs actual label

The model's effectiveness in classifying breast tumours is evaluated using multiple visualizations.

a)Confusion Matrix (Figure 3c)

From the confusion matrix, we observe 70 true negatives, 40 true positives, 2 false positives, and 2 false negatives, demonstrating strong classification performance with minimal mistakes.

b) Accuracy & Loss Over Epochs (Figure 3a)

The accuracy plot shows rapid improvement in both training and validation accuracy, stabilizing around 95-97% after 10 epochs, demonstrating strong generalization. The loss curves show steady reduction, with minor fluctuations in validation loss, suggesting effective learning with minimal overfitting.

c)Classification Report

The model demonstrates 97% precision, recall, and F1-score for benign cases., and 95% for malignant cases, with an overall accuracy of 96%, indicating a well-balanced and reliable performance.

III. IoT and Cloud Connectivity for continous Monitoring

IoT Integration Wearable device with multimodal antennas for continuous tumor monitoring. - Edge AI processing with Raspberry Pi for initial classification.

In the earlier phase of this work, a physical prototype was developed to validate the feasibility of breast tumor classification using thermal and electrical parameters. In the current study, the authors have transitioned to a simulation-based approach, enabling better control over input variables and facilitating comprehensive testing of the proposed deep learning models. The dataset features were simulated as sensor outputs and integrated with a Raspberry Pi-based embedded system, mimicking real-time data acquisition. The trained model was deployed on the Raspberry Pi to evaluate its performance in on-device inference, demonstrating the system's potential for real-world applications in low-cost and portable diagnostic platforms.

- Encrypting patient data storage in the cloud. - Secure API for real-time alerts and remote access by medical profession can increase the robustness of the system.

IV. SECTION-4

Analysis and Potential Developments

n this paper, a wearable IoT system for tumor classification was presented, incorporating machine learning and deep learning methods.. Future work includes improving model generalization, integrating additional biosensors, and real-time deployment in clinical trials.

- Data transmission from Raspberry Pi to cloud (AWS/Firebase).
- Remote access through a web or mobile dashboard.

TABLE II
Raspberry Pi 5 Performance Metrics for ML Models

Model	Inference Time (ms)	CPU Utilization (%)	Power (W)
Logistic Regression	30-40	70-75	5.2
Random Forest	80-120	85-95	6.5
SVM	90-150	90-95	6.8
KNN	60-100	85-90	6.0
Decision Tree	40-70	75-80	5.5
XGBoost	120-180	90-100	7.0
AdaBoost	100-150	85-95	6.7
Voting Classifier	120-200	90-100	7.2

A. ML/DL Model Performance

The Table 2 given above give performance of the raspberry pi while using different Machine learning models (parameters like inference time, CPU-utilization, Power are the key metrics here

The performance of various machine learning models for breast cancer classification is evaluated based on inference time, CPU utilization, and power consumption. While these metrics provide insight into computational efficiency, accuracy remains a critical factor in determining the effectiveness of these models.

Logistic Regression-having the lowest inference time (30-40 ms) and moderate CPU utilization (70-75%), is typically used for linear classification problems. However, its performance may be suboptimal if the dataset is non-linearly separable, leading to lower accuracy compared to more complex models.

Random Forest and Decision Tree- models demonstrate relatively balanced inference times and power consumption. Random Forest, benefiting from ensemble learning, often provides better generalization and robustness, leading to improved accuracy. Decision Trees, although computationally efficient (40-70 ms inference time), may suffer from overfitting, which can slightly reduce accuracy.

SVM and KNN algorithms- show inference times ranging from 60-150 ms, with high CPU utilization (90-95%). These models are known for their strong classification performance, especially in high-dimensional datasets. SVM, with its ability to find optimal hyperplanes, often delivers high accuracy, while KNN's performance depends heavily on the choice of neighbours and distance metric.

Boosting algorithms, including XGBoost and AdaBoost, show the highest inference times (100-200 ms), but they leverage iterative learning to improve accuracy. XGBoost, known for its efficiency in handling structured data, generally outperforms other models in terms of precision and recall. AdaBoost also enhances weak classifiers but may be sensitive to noisy data.

Finally, the Voting Classifier, which aggregates multiple models, provides a trade-off between accuracy and computational cost. By combining the strengths of various models, it aims to enhance prediction reliability, often resulting in superior classification accuracy at the expense of increased inference time and resource utilization.

D .Conclusions :

The integration of UWB microwave imaging with the Internet of Things (IoT) presents new possibilities for smart healthcare monitoring. This paper presents a wearable IoT-based breast tumour classification system leveraging Raspberry Pi and hybrid deep learning models. The proposed approach aims to provide real- time, non-invasive tumor detection with cloud-based remote monitoring.

Future IoT devices may be integrated as follows: Using IoT for real-time diagnosis by deploying the model on edge platforms such as Raspberry Pi or Jetson Nano.for real-time breast cancer detection and remote monitoring.

Our deep learning approach improved classification accuracy while mitigating overfitting. Future work includes optimizing hyperparameters and integrating real-time deployment using edge devices.

Our analysis demonstrates that traditional ML models can effectively classify breast tumours with high accuracy. However, the slight variations in testing accuracy highlight the need for model selection based on robustness against overfitting. Future work will explore optimizing hyperparameters and integrating deep learning with ensemble methods for improved classification reliability. Also it is highly effective to use s-parameter vs frequency data for better classification accuracy and non-invasiveness.

1. Sensitivity to Tissue Heterogeneity: Breast tissues exhibit varying dielectric properties between healthy and malignant tissues. S-parameters, measured across a frequency spectrum, capture these differences effectively. Studies have demonstrated that malignant tissues show distinct S-parameter signatures compared to benign tissues, enhancing classification accuracy.[15]

2. Non-Invasive and Real-Time Assessment: Utilizing S-parameters allows for non-invasive probing of breast tissues. Techniques such as microwave imaging leverage S-parameter data to construct real-time images, facilitating immediate and accurate tumor detection.[16]

ACKNOWLEDGMENT

The authors wish to convey their heartfelt thanks. to the faculty and mentors at Thapar Institute of Engineering and Technology for their continuous guidance and support. We also acknowledge the contributions of our project team members and research collaborators who assisted in data collection and prototyping. Special thanks to the Electronics and Communication Engineering Department for providing technical resources and lab facilities. This work was carried out as part of an undergraduate research initiative focused on AI and wearable healthcare technologies.

V. REFERENCES

[1] IJISAE, ”Advancing Breast Cancer Detection: Integrating IoT and Deep Learning,” IJISAE, 2023. Available: https://ijisae.org/index.php/IJISAE/ article/view/5109.

[2]NCBI PubMed, ”Smart IoT in Breast Cancer Detection Using Optimal Deep Learning,” PubMed, 2023. Available: https://pubmed.ncbi.nlm.nih. gov/37221422/.

[3]“Sycai Medical’s AI-Based Cancer Detection Secures €3 Million in Funding,” CINCODIAS.ELPAIS.COM, Jan. 2025. [Online]. Available: https://www.cincodias.elpais.com. [Accessed: Feb. 24, 2025].

[4] DOI: 10.1109/JSEN.2022.3205116

[3]ResearchGate,” Integrating IoT and Machine Learning for Advanced Breast Cancer Detection: A Secure Framework for Enhanced Diagnostic Accuracy,” ResearchGate, 2023. Available: https://www.researchgate. net/publication/385872133.

[4] PMC NCBI, ”Deep Learning in Different Ultrasound Methods for Breast Cancer Diagnosis,” PMC NCBI, 2023. Available: https://pmc.ncbi.nlm. nih.gov/articles/PMC10296633/.

[5] MDPI, ”Breast Cancer Detection with an Ensemble of Deep Learning Models,” MDPI, 2023. Available: https://www.mdpi.com/2313-433X/9/ 11/247.

[6] Semantic Scholar, ”Early Detection of Breast Cancer with IoT: A Promising Solution,” Semantic Scholar, 2023. Available: https://www.semanticscholar.org/paper/Early-Detection-of-Breast-Cancer-with-IoT.

[7] Wiley Online Library, ”A Flexible Wearable Thermography System Based on Bioheat Transfer Simulation for Breast Cancer Detection,” Wiley Online Library, 2023. Available: https://onlinelibrary.wiley.com/ doi/10.1155/2022/5921691.

[8] International Publications, ”Integrating IoT and Machine Learn- ing for Advanced Breast Cancer Detection,” International Publi- cations, 2023. Available: https://internationalpubls.com/index.php/pmj/article/download/2045/1302/3704.

[9] Wikipedia, ”Conformable Ultrasound Breast Patch for Deep Tis- sue Scanning and Imaging,” Wikipedia, 2023. Available: https://en.wikipedia.org/wiki/Canan Da%C4%9Fdeviren.

[10] Reuters, ”Health Rounds: Experimental Breast Prosthesis Detects, Treats Cancer in Mice,” Reuters, 2023. Available: https://www.reuters.com/business/healthcare-pharmaceuticals/health-rounds-experimental-breast-prosthesis-detects-treats-cancer-mice-2024-10-09. 10

.

[11] B. Author et al., ”Deep Learning for Breast Cancer Classification,” IEEE Transactions on Medical Imaging, 202X. 99

[12]D. N. Elsheakh, R. A. Mohamed, O. M. Fahmy, K. Ezzat, and A. R. Eldamak, “Complete Breast Cancer Detection and Monitoring System by Using Microwave Textile-Based Antenna Sensors,” Biosensors (Basel), vol. 13, no. 1, p. 87, Jan. 2023. DOI: 10.3390/bios13010087.

[13] R. Dessai, D. Singh, M. Sonkki, J. Reponen, T. Myllyla¨, S. Myllyma¨ki, and M. Sa¨resto¨niemi, “A Breast Tumor Monitoring Vest with Flexible UWB Antennas—A Proof-of-Concept Study Using Realistic Breast Phantoms,” Micromachines, vol. 15, no. 9, p. 1153, Sep. 2024. DOI: 10.3390/mi15091153.

[14] Breast tissue tumor detection using “S” parameter analysis with an UWB stacked aperture coupled microstrip patch antenna having a “ + ” shaped defected ground structure | International Journal of Microwave and Wireless Technologies | Cambridge Core

[15]

Islam, N., Hasib, K.M., Mridha, M.F. et al. Fusing global context with multiscale context for enhanced breast cancer classification. Sci Rep 14, 27358 (2024). https://doi.org/10.1038/s41598-024-78363-w

[16] Mokhtari Dowlatabad H, Mamdouh A, Yousefpour N, Mahdavi R, Zandi A, Hoseinpour P, Moosavi-Kiasari SMS, Abbasvandi F, Kordehlachin Y, Parniani M, Mohammadpour-Aghdam K, Faranoush P, Foroughi-Gilvaee MR, Abdolahad M. High-Frequency (30 MHz-6 GHz) Breast Tissue Characterization Stabilized by Suction Force for Intraoperative Tumor Margin Assessment. Diagnostics (Basel). 2023 Jan 4;13(2):179. doi: 10.3390/diagnostics13020179. PMID: 36672989; PMCID: PMC9857665.

[17] Wolberg, W., Mangasarian, O., Street, N., & Street, W. (1993). Breast Cancer Wisconsin (Diagnostic) [Dataset]. UCI Machine Learning Repository. https://doi.org/10.24432/C5DW2B.