Journal of Advanced Research in Applied Sciences and Engineering Technology

Journal homepage:
https://semarakilmu.com.my/journals/index.php/applied_sciences_eng_tech/index
ISSN: 2462-1943

UWB-Based Early Breast Cancer Existence Prediction Using Artificial Intelligence for Large Data Set

Ahmad Ashraf Abdul Halim[1,*], Vijayasarveswari Veeraperumal[1], Allan Melvin Andrew[2], Mohd Najib Mohd Yasin[1], Mohd Zamri Zahir Ahmad[1], Kabir Hossain[3], Bifta Sama Bari[4], Fatinnabila Kamal[5]

[1] Faculty of Electronic Engineering Technology (FTKEN), Universiti Malaysia Perlis, 02600, Arau, Perlis, Malaysia
[2] Faculty of Electrical Engineering Technology (FTKE), Universiti Malaysia Perlis, 02600, Arau, Perlis, Malaysia
[3] Norwegian University of Science and Technology (NTNU), N-2815 Gjøvik, Norway
[4] Faculty of Electrical and Electronic Engineering Technology (FTKEE), Universiti Malaysia Pahang, 26600, Pekan, Pahang, Malaysia
[5] Faculty of Applied and Human Sciences (FSGM), Universiti Malaysia Perlis, 01000, Kangar, Perlis Malaysia

ARTICLE INFO

Article history:
Received 6 October 2022
Received in revised form 7 November 2022
Accepted 8 December 2022
Available online 6 January 2023

Keywords:
Breast cancer detection; Feature selection; Machine learning

ABSTRACT

Breast cancer is the most often identified cancer among women and the main reason for cancer-related deaths worldwide. The most effective methods for controlling and treating this disease through breast screening and emerging detection techniques. This paper proposes an intelligent classifier for the early detection of breast cancer using a larger dataset since there is limited researcher focus on that for better analytic models. To ensure that the issue is tackled, this project proposes an intelligent classifier using the Probabilistic Neural Network (PNN) with a statistical feature model that uses a more significant size of data set to analyze the prediction of the presence of breast cancer using Ultra Wideband (UWB). The proposed method is able to detect breast cancer existence with an average accuracy of 98.67%. The proposed module might become a potential user-friendly technology for early breast cancer detection in domestic use.

1. Introduction

Breast cancer is the most significant factor in cancer death among women in the world [1-3]. Statistical analysis revealed a significant rise in the breast cancer death rate across all regions especially in lower-income region [4-6]. Early breast cancer detection or screening with an accurate diagnosis and treatment leads to a reduction in death rate while reducing treatment costs [7], [8]. Cancer is a disease where the body reproduces cells, and the reaction cells divide uncontrollably, leading to irregular cell development or a tumour that spreads into surrounding tissues. The tumour (also called neoplasm) can be classified as benign (noncancerous) or malignant (cancerous) [9]. Benign tumours stay in their primary location and do not spread to nearby tissues.

In contrast, a cancerous tumour gives cells that grow uncontrollably and spread to nearby tissues, harm those tissues, and impact other bodily parts. Chronic problems can occur if cancer cells move

[*] *Corresponding author.*
E-mail address: ashrafhalim@unimap.edu.my

https://doi.org/10.37934/araset.29.2.8190

to other organs. Therefore, it is important that early detection of the cancer cells presence is crucial to cure and prevent the cell from spreading to the other part of the body.

Early-stage breast cancer is diagnosed using a variety of current screenings and new technology. The two main categories of current breast cancer screening technologies are body imaging-based technology and microwave imaging-based technology, as depicted in Figure 1 [10], [11]. Body image-based technologies are conventional breast cancer screening, such as magnetic resonance imaging (MRI), mammography, and ultrasound [12-14]. These items may be found in most clinics and hospitals. As an alternative to costly and intrusive screening methods, microwave imaging-based technology is ready.

Two methods utilized in microwave imaging technologies include microwave tomography and radar-based imaging. Both methods used UWB signals to classify breast cancer according to its dielectric characteristics. This method has better precise ranges and is simple, inexpensive, safe, and free of ionizing radiation exposure.

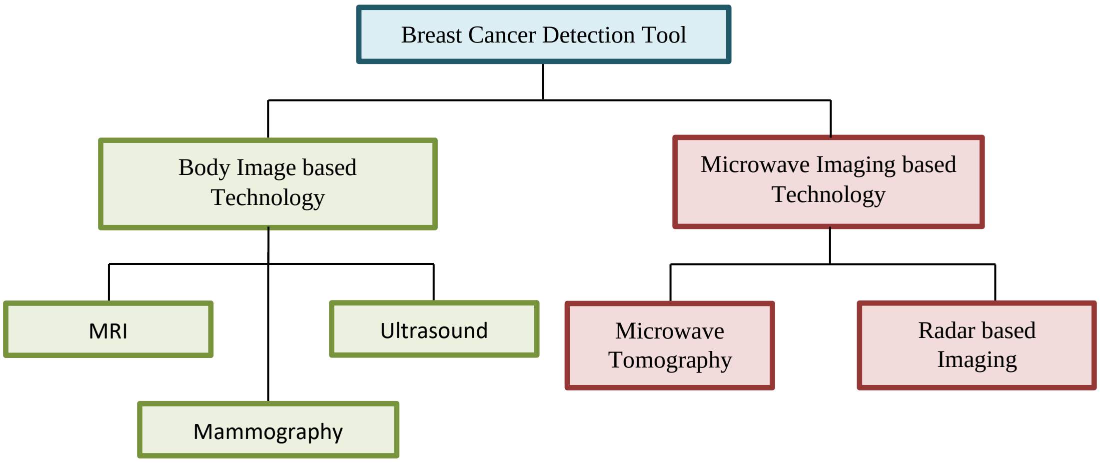

Fig. 1. Block diagram showing the different modalities in breast cancer detection [15]

Many other researchers have conducted studies on breast cancer detection using UWB. Vijayasarveswari *et al.,* [16] propose a breast cancer screening module for small data sets based on Artificial Neural Networks (ANN). This system is focused on detecting a tumour in the breast phantom. Bifta *et al.,* [17] proposed a UWB experimental setup. This method uses a small dataset to indicate tumours existence, size, and location. For the result, first, second, third, and fourth data sets, the detection efficiency of tumour existence, position (x, y, z), and size was around 87.72 percent, 87.24 percent, 83.93 percent, and 80.51 percent, respectively. The detection accuracy for the three-parameter is 100%, 92.43%, and 91.31%, respectively.

2. Methodology

2.1 Proposed System Architecture

The proposed system architecture consists of hardware and software modules. The hardware includes two antennae (transmitter and receiver) [18], a breast phantom, a tumour, and a UWB transceiver with a Personal Computer (PC) interface. The software comprises a data processing, classifier, and Graphical User Interface (GUI). The proposed development of the design system architecture is shown in Figure 2.

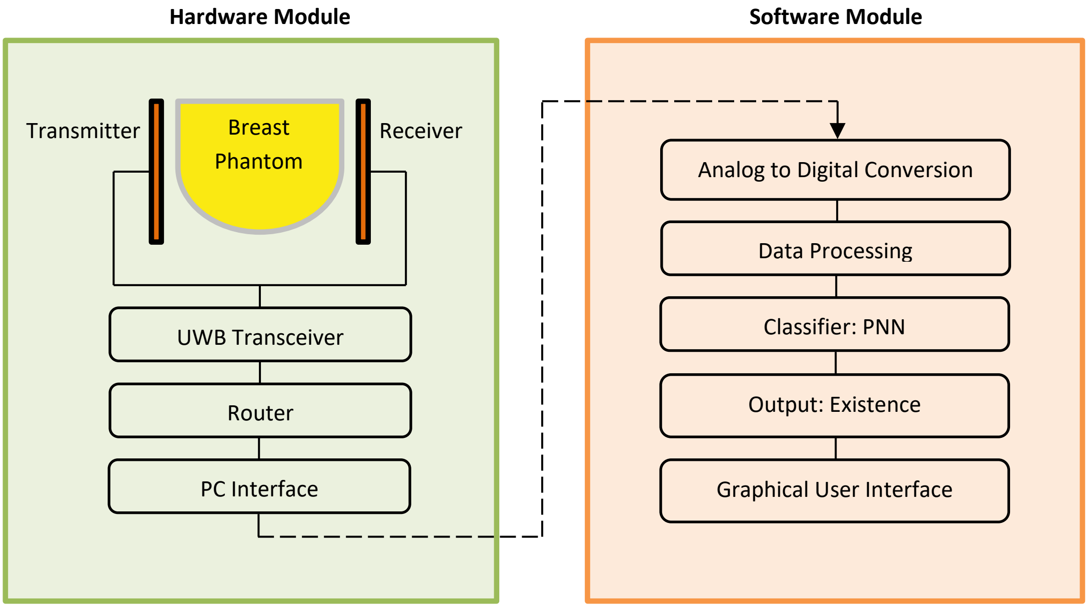

Fig. 2. Overall system design architecture

2.2 Breast Phantom and Tumour

Various breast phantoms have been proposed to explore the researchers capability to detect breast cancer [19], [20]. Most researchers create heterogeneous breast phantoms using inexpensive and non-chemical substances, such as petroleum jelly or a mixture of wheat flour, water, and soy oil. It is necessary to ensure that the permittivity and conductivity of the breast phantoms are comparable to those of actual breast tissue, as shown in Table 1. The breast phantom and tumour are shown in Figure 3.

(a) (b)

Fig. 3. (a) The breast phantom, (b) The tumour

2.3 Data Collection

As illustrated in Figure 4, the breast phantom is between the transmitter and receiver. Using the Ethernet cable, the sensor is connected to the UWB transceiver (P400 RCM). The UWB pulses are created in the transceiver and transferred through the transmitting sensor to the receiver. The receiver antenna captured the signals at the center frequency of 4.3 GHz, passing through the router before all the data were analyzed using Matlab software [21], [22].

The steps for collecting data are as follows [23]:

i. The tumour is implanted in a breast phantom.
ii. The antenna transmits UWB signals, and the opposite antenna captures forward scattered UWB signals. Fifty repetitions are taken for each cycle.
iii. Step 1 to Step 2 is repeated for 100 cycles.
iv. Repeat Steps 1 to Step 3 but without the tumour. A total of 10 000 UWB signals are collected. Each signal sample has 1632 data points.

Table 1
The dielectric properties of developed breast phantom and tumour [17], [23], [24]

Breast Phantom Structure	Material	Permittivity	Conductivity (S/M)
Fatty tissue	Pure petroleum jelly	2.36	0.012
Glandular	Soy oil	2.7	0.061
Skin	Glass	3.5-10	Negligible
Tumour	A mixture of water and wheat flour	6.98	0.785

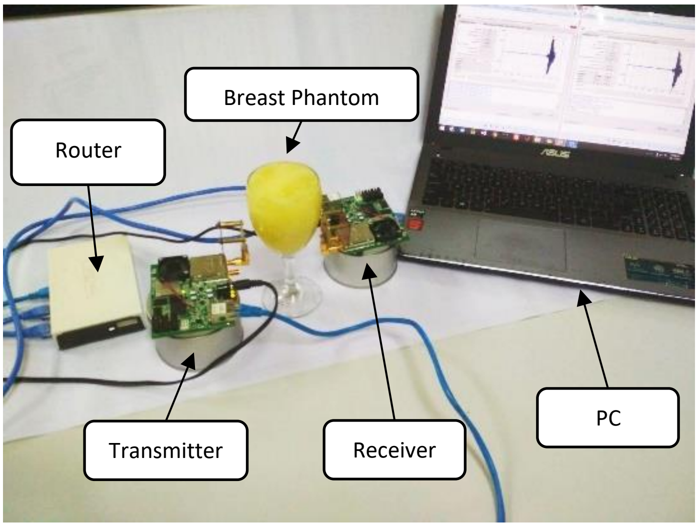

Fig. 4. Experimental setup for breast cancer detection

2.4 Data Pre-processing and Statistical Feature Method

In general, the signal is in the time domain. In the time domain, it is easier to visualize signal properties. Nevertheless, examining signal characterization in the frequency domain is also essential because it permits the observation of signal properties that are not visible in the time domain. Thus, the signals in the frequency domain collected from UWB transceivers are converted using the widely employed Fast Fourier Transform (FFT) [23], [25]. This system was trained and tested using a total of 10,000 data samples. Figure 5 depicts the experimental software process flowchart.

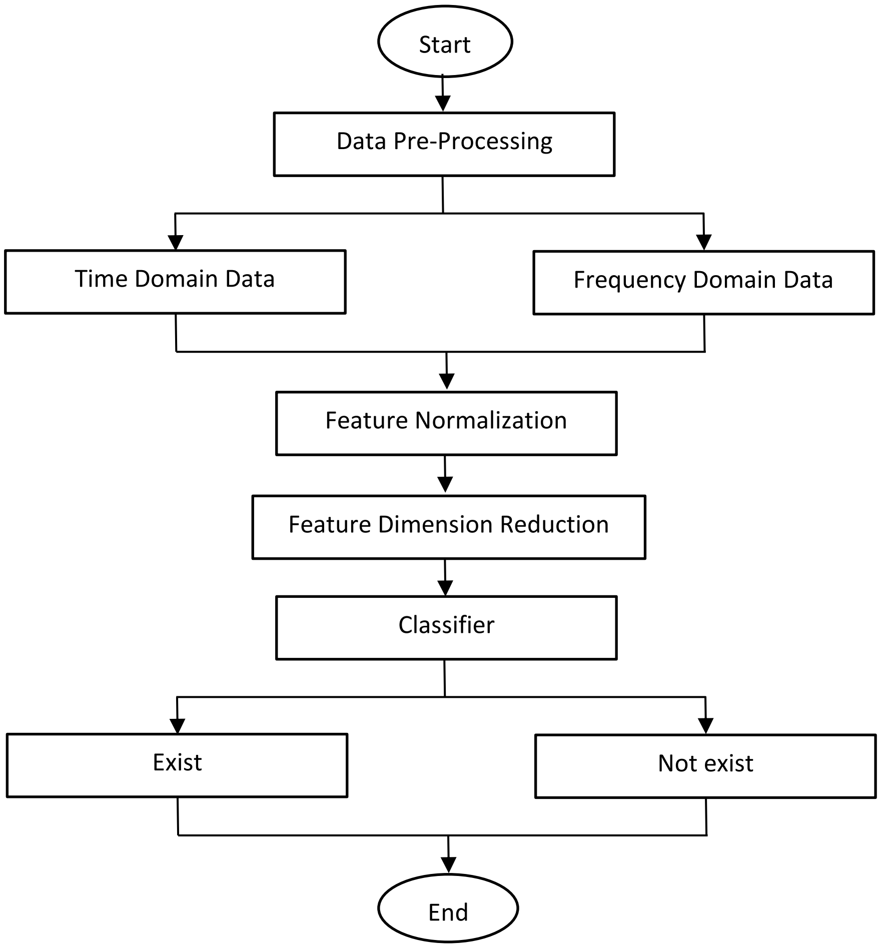

Fig. 5. Experimental software process flowchart

Feature normalization is organizing and standardizing the data without eliminating useful information from the preprocessed data. Once the data is normalized, the normalized data is dimensionally reduced to remove redundant and statistically insignificant data.

Sixty percent of data samples are used for training, while forty percent are utilized for testing. The received signals were processed to generate 1632 discrete data points per sample. Before feeding the signals to the classifier, they are preprocessed to obtain critical features (five features: binary normalization, decimal scaling, linear scaling, min-max, and z-score) in Eq. (1) to Eq. (5) and data dimension reduction using Singular Value Decomposition (SVD) [24].

Binary normalization ensures the maximum accuracy of a number for a given range of bits.

$$BN = (0.8 \times (v - min_D)./(max_D - min_D) + 0.1) \tag{1}$$

Where v is the instantaneous value of feature D, $maxD$ and $minD$ are the maximum and minimum values of D.

The decimal scaling method normalizes the data by moving the decimal points. The number of decimal points depends on the maximum absolute value of the data sample, D.

$$DS = \frac{v}{10^j} \tag{2}$$

Where v is the instantaneous value of feature D, and j is the smallest integer that can obtain a maximum v' with a value less than 1.

Linear scaling is the special case of the min-max normalization method. It normalizes the data to a [0, 1] range

$$LS = \frac{v - min_D}{max_D - min_D} \quad (3)$$

Where v is the instantaneous value of feature D and $maxD$, and $minD$ are the maximum and minimum values of D, respectively.

Min-max normalization method rescales the data from one range to a new range which is [-1, 1] range

$$MM = \frac{v - min_D}{max_D - min_D}(new\_max_D - new\_min_D) + new\_min_D \quad (4)$$

Where v is the instantaneous value of feature D, $maxD$ and $minD$ are the maximum and minimum values of D, respectively, new_maxD is one and new_minD is -1.

The data is normalized by converting the particular value to a common scale with zero mean and unity standard deviation

$$ZS = \frac{v - \mu_D}{\sigma_D} \quad (5)$$

Where v is the instantaneous value of feature D and μD, and σD are the mean and standard deviation of feature D, respectively.

2.5 Probabilistic Neural Network (PNN)

Specht introduced standard back-propagation neural networks in 1998 for classification and pattern recognition applications. PNN is a scalable alternative to these networks. Standard neural networks require extensive forward and reverse calculations. In contrast, PNN requires less computational time and power. Additionally, they can manage a variety of training data. These networks apply probability theory to classification tasks in order to reduce misclassifications. In this study, a PNN classifier was used to determine the presence of breast cancer based on input parameters. The proposed system was founded on gathered data. The PNN architecture is shown in Figure 6 [26].

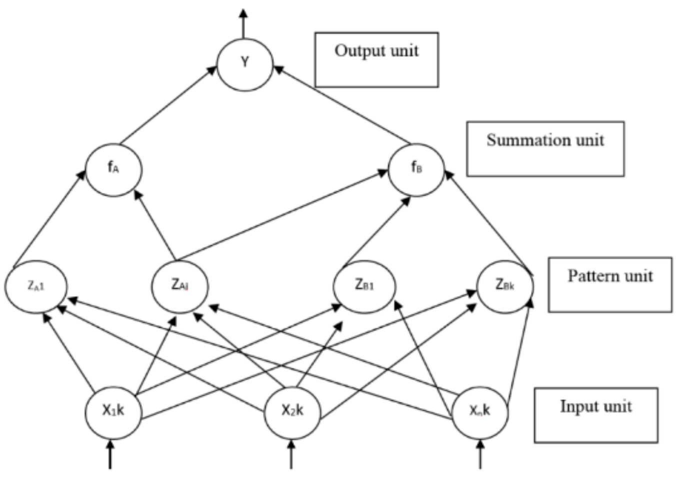

Fig. 6. PNN architecture [26]

The overall process flow of the proposed multi-stage feature selection and fusion for both datasets is shown in Figure 7.

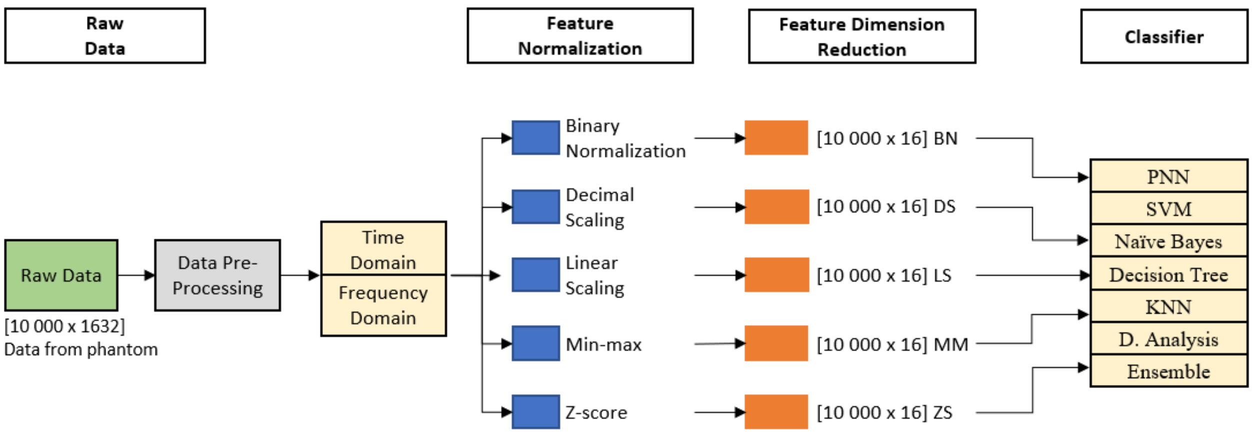

Fig. 7. Block diagram of the proposed method

3. Results

The proposed algorithm can assist the classifier in detecting the existence of the tumour in the heterogeneous breast phantom. The proposed algorithm is tested using seven classifiers to ensure a robust model. The result for the time and frequency domain is shown in Table 2 and Table 3. The result frequency domain shows better results compared to the time domain.

Table 4 compares the proposed system with other existing methods. Vijayasarveswari *et al.,* [16] and Bifta *et al.,* [17] used small data sets compared to this project, which used 20 times larger than an existing project. Even though the proposed systems performance is slightly lower than the previous researcher, it still shows that this system can precisely classify more than 98% of 10 000 datasets. This is important for building analytic models with more extensive datasets using machine learning.

Table 2
Classifiers result in the time domain

Frequency Domain	BN	DS	LS	MM	ZS
SVM	85.81318681	85.71428571	85.35164835	85.98901099	86.31868132
NB	69.86813187	33.72527473	38.3956044	34.92307692	34.45054945
KNN	79.43956044	88.30769231	79.13186813	87.81318681	82
DT	77.73626374	85.58241758	77.1978022	86.23076923	82.14285714
DA	85.87912088	85.83516484	85.57142857	85.52747253	85.65934066
Ensemble	83.93406593	89.38461538	84.52747253	88.96703297	86.37362637
PNN	86.021978 (0.009)	90.395604 (0.009)	86.351648 (0.01)	90.021978 (0.05)	86.4835165 (0.9)

Table 3
Classifiers result in the frequency domain

Frequency Domain	BN	DS	LS	MM	ZS
SVM	85.6043956	85.86813187	85.79120879	86.1978022	85.65934066
NB	72	72.41758242	71.96703297	72.63736264	65.56043956
KNN	92.15384615	91.73626374	91.62637363	92.1978022	95.6043956
DT	94.06593407	94.06593407	93.53846154	94.45054945	93.02197802
DA	93.93406593	84.87912088	84.57142857	85.42857143	85.63736264
Ensemble	91.703297	91.2527473	95.82417582	92.340659	95.74725275
PNN	96.50549451 (0.001)	96.04395604 (0.05)	91.197802 (0.001)	96.67032967 (0.001)	98.671438 (0.9)

Table 4
Comparison with the previous researcher

Researcher	Data Sample	Method	Test and Trains Sets	Existence Accuracy (%)
Vijayasarveswari *et al.,* 2021 [16]	125	FFBPNN	K-Fold Cross Validation	100
Bifta *et al.,* 2020 [17]	448	FFBPNN with feedforward net function	70% Training 15% Validation 15% Testing	100
Proposed System	10000	PNN classifier	K-Fold Cross Validation	98.67

Figure 8 shows breast cancer detection visualization using a developed GUI. This is to make it easier for the end-user to check for any breast abnormalities.

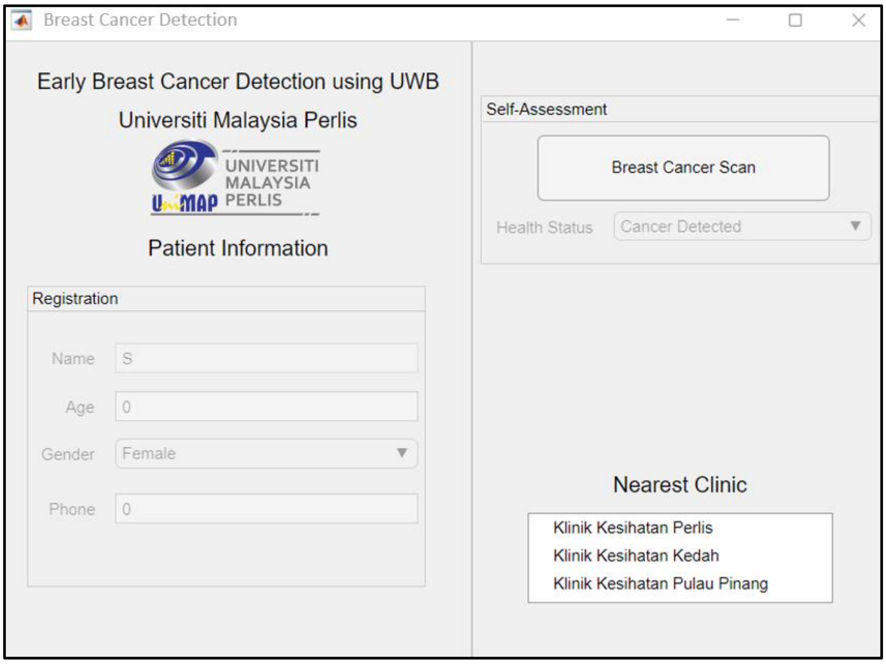

Fig. 8. GUI for early breast cancer detection

4. Conclusions

This paper proposes a practical technique for breast cancer detection modules based on PNN. The module is developed by combining the k-fold cross-validation method with the PNN classifier. The developed module is tested for breast cancer detection applications. Ten thousand data sample is used to classify the existence of breast cancer. The proposed model detection performance accuracy is above 98%, even though large data samples were fed into this model. By integrating a GUI, medical practitioners and end users can use the system at home for breast health self-screening.

Acknowledgment

The authors like to thank Universiti Malaysia Perlis (UniMAP) and Universiti Malaysia Pahang (UMP) for their generous assistance in providing research facilities for this study. This research was funded by Research Materials Fund (RESMATE) UniMAP (RESMATE Grant. 900100626).

References

[1] Kamarudin, Mohamad Saddam, Ishkrizat Taib, Intan Syafinaz, Ahmad Mubarak Tajul Arifin, and Mohd Noor Abdullah. "Analysis of Heat Propagation on Difference Size of Malignant Tumor." *Journal of Advanced Research in Applied Sciences and Engineering Technology* 28, no. 2 (2022): 211-221. https://doi.org/10.37934/araset.28.2.211221

[2] Alanazi, Saad Awadh, M. M. Kamruzzaman, Md Nazirul Islam Sarker, Madallah Alruwaili, Yousef Alhwaiti, Nasser Alshammari, and Muhammad Hameed Siddiqi. "Boosting breast cancer detection using convolutional neural network." *Journal of Healthcare Engineering* 2021 (2021). https://doi.org/10.1155/2021/5528622

[3] Kamarudin, Saddam, Ishkrizat Taib, Nurul Fitriah Nasir, Zainal Ariff Abidin, Hazimuddin Halif, A. M. T. Arifin, and Mohd Noor Abdullah. "Comparison of Heat Propagation Properties in Different Sizes of Malignant Breast Tumours using Computational Fluid Dynamics." *Journal of Advanced Research in Applied Sciences and Engineering Technology* 28, no. 3 (2022): 368-375. https://doi.org/10.37934/araset.28.3.368375

[4] Azamjah, Nasrindokht, Yasaman Soltan-Zadeh, and Farid Zayeri. "Global trend of breast cancer mortality rate: a 25-year study." *Asian Pacific journal of cancer prevention: APJCP* 20, no. 7 (2019): 2015. https://doi.org/10.31557/APJCP.2019.20.7.2015

[5] Bevacqua, Martina Teresa, Simona Di Meo, Lorenzo Crocco, Tommaso Isernia, and Marco Pasian. "Millimeter-waves breast cancer imaging via inverse scattering techniques." *IEEE Journal of Electromagnetics, RF and Microwaves in Medicine and Biology* 5, no. 3 (2021): 246-253. https://doi.org/10.1109/JERM.2021.3052096

[6] Tan, King Fang, Farzaana Adam, Rohayu Hami, Noorsuzana Mohd Shariff, and Noor Mastura Mohd Mujar. "Review of breast cancer in young women." *Med. J. Malays* 16 (2020): 8-22.

[7] Mann, Ritse M., Christiane K. Kuhl, and Linda Moy. "Contrast-enhanced MRI for breast cancer screening." *Journal of Magnetic Resonance Imaging* 50, no. 2 (2019): 377-390. https://doi.org/10.1002/jmri.26654

[8] Abdollahi, Jafar, Atlas Keshandehghan, Mahsa Gardaneh, Yasin Panahi, and Mossa Gardaneh. "Accurate detection of breast cancer metastasis using a hybrid model of artificial intelligence algorithm." *Archives of Breast Cancer* (2020): 22-28. https://doi.org/10.32768/abc.20207118-24

[9] Kajala, Aditi, and V. K. Jain. "Diagnosis of breast cancer using machine learning algorithms-a review." In *2020 International Conference on Emerging Trends in Communication, Control and Computing (ICONC3)*, pp. 1-5. IEEE, 2020. https://doi.org/10.1109/ICONC345789.2020.9117320

[10] Sani, Lorenzo, Alessandro Vispa, Riccardo Loretoni, Michele Duranti, Navid Ghavami, Daniel Alvarez Sánchez-Bayuela, Stefano Caschera et al. "Breast lesion detection through MammoWave device: Empirical detection capability assessment of microwave images' parameters." *Plos one* 16, no. 4 (2021): e0250005. https://doi.org/10.1371/journal.pone.0250005

[11] Aldhaeebi, Maged A., Khawla Alzoubi, Thamer S. Almoneef, Saeed M. Bamatraf, Hussein Attia, and Omar M. Ramahi. "Review of microwaves techniques for breast cancer detection." *Sensors* 20, no. 8 (2020): 2390. https://doi.org/10.3390/s20082390

[12] Mehmood, Mavra, Ember Ayub, Fahad Ahmad, Madallah Alruwaili, Ziyad A. Alrowaili, Saad Alanazi, M. Humayun Muhammad Rizwan, Shahid Naseem, and Tahir Alyas. "Machine learning enabled early detection of breast cancer by structural analysis of mammograms." *Comput. Mater. Contin* 67 (2021): 641-657. https://doi.org/10.32604/cmc.2021.013774

[13] Houssein, Essam H., Marwa M. Emam, Abdelmgeid A. Ali, and Ponnuthurai Nagaratnam Suganthan. "Deep and machine learning techniques for medical imaging-based breast cancer: A comprehensive review." *Expert Systems with Applications* 167 (2021): 114161. https://doi.org/10.1016/j.eswa.2020.114161

[14] Mann, Ritse M., Nariya Cho, and Linda Moy. "Breast MRI: State of the Art." *Radiology* 292, no. 3 (2019): 520–36. https://doi.org/10.1148/radiol.2019182947

[15] Abdul Halim, Ahmad Ashraf, Allan Melvin Andrew, Mohd Najib Mohd Yasin, Mohd Amiruddin Abd Rahman, Muzammil Jusoh, Vijayasarveswari Veeraperumal, Hasliza A. Rahim, Usman Illahi, Muhammad Khalis Abdul Karim, and Edgar Scavino. "Existing and Emerging Breast Cancer Detection Technologies and Its Challenges: A Review." *Applied Sciences* 11, no. 22 (2021): 10753. https://doi.org/10.3390/app112210753

[16] Vijayasarveswari, V., M. Jusoh, T. Sabapathy, R. A. A. Raof, S. Khatun, and I. Iszaidy. "Reliable early breast cancer detection using artificial neural network for small data set." In *Journal of Physics: Conference Series*, vol. 1755, no. 1, p. 012037. IOP Publishing, 2021. https://doi.org/10.1088/1742-6596/1755/1/012037

[17] Bari, Bifta Sama, Sabira Khatun, Kamarul Hawari Ghazali, Md Fakir, Wan Nur Azhani W. Samsudin, Mohd Falfazli Mat Jusof, Mamunur Rashid, Minarul Islam, and Mohd Zamri Ibrahim. "Ultra wide band (uwb) based early breast cancer detection using artificial intelligence." In *InECCE2019*, pp. 505-515. Springer, Singapore, 2020. https://doi.org/10.1007/978-981-15-2317-5_43

[18] Bari, Bifta Sama, Sabira Khatun, Kamarul Hawari Ghazali, Md Fakir, Mohd Hisyam Mohd Ariff, Mohd Faizal Jamlos, Mamunur Rashid, Minarul Islam, Mohd Zamri Ibrahim, and Mohd Falfazli Mat Jusof. "Bandwidth and gain enhancement of a modified ultra-wideband (UWB) micro-strip patch antenna using a reflecting layer." In *InECCE2019*, pp. 463-473. Springer, Singapore, 2020. https://doi.org/10.1007/978-981-15-2317-5_39

[19] Vispa, Alessandro, Lorenzo Sani, Martina Paoli, Alessandra Bigotti, Giovanni Raspa, Navid Ghavami, Stefano Caschera, Mohammad Ghavami, Michele Duranti, and Gianluigi Tiberi. "UWB device for breast microwave imaging: phantom and clinical validations." *Measurement* 146 (2019): 582-589. https://doi.org/10.1016/j.measurement.2019.05.109

[20] Vijayasarveswari, V., M. Jusoh, and S. Khatun. "Experimental UWB based efficient breast cancer early detection." *Indian J. Sci. Technol* 10 (2017): 1-6. https://doi.org/10.17485/ijst/2017/v10i12/113016

[21] Hassan, Nouralhuda A., Amr H. Yassin, Mazhar B. Tayel, and Moustafa M. Mohamed. "Ultra-wideband scattered microwave signals for detection of breast tumors using artifical neural networks." In *2016 Third International Conference on Artificial Intelligence and Pattern Recognition (AIPR)*, pp. 1-6. IEEE, 2016. https://doi.org/10.1109/ICAIPR.2016.7585226

[22] Singh, Nalini, Ambarish G. Mohapatra, and Gurukalyan Kanungo. "Breast cancer mass detection in mammograms using K-means and fuzzy C-means clustering." *International Journal of Computer Applications* 22, no. 2 (2011): 15-21. https://doi.org/10.5120/2557-3507

[23] Vijayasarveswari, V., A. M. Andrew, M. Jusoh, T. Sabapathy, R. A. A. Raof, M. N. M. Yasin, R. B. Ahmad, S. Khatun, and H. A. Rahim. "Multi-stage feature selection (MSFS) algorithm for UWB-based early breast cancer size prediction." *PloS one* 15, no. 8 (2020): e0229367. https://doi.org/10.1371/journal.pone.0229367

[24] Halim, Ahmad Ashraf Abdul, Allan Melvin Andrew, Wan Azani Mustafa, Mohd Najib Mohd Yasin, Muzammil Jusoh, Vijayasarveswari Veeraperumal, Mohd Amiruddin Abd Rahman, Norshuhani Zamin, Mervin Retnadhas Mary, and Sabira Khatun. "Optimized Intelligent Classifier for Early Breast Cancer Detection Using Ultra-Wide Band Transceiver." *Diagnostics* 12, no. 11 (2022): 2870. https://doi.org/10.3390/diagnostics12112870

[25] Shao, Wenyi, Arezou Edalati, Todd R. McCollough, and William J. McCollough. "A time-domain measurement system for UWB microwave imaging." *IEEE Transactions on Microwave Theory and Techniques* 66, no. 5 (2018): 2265-2275. https://doi.org/10.1109/TMTT.2018.2801862

[26] Andrew, Allan Melvin, Ammar Zakaria, Shaharil Mad Saad, and Ali Yeon Md Shakaff. "Multi-stage feature selection based intelligent classifier for classification of incipient stage fire in building." *Sensors* 16, no. 1 (2016): 31. https://doi.org/10.3390/s16010031