An AI-Driven Diagnostic Decision Support System Using EP-Optimized RBF Neural Networks for Breast Cancer Detection

Vijaylaxmi Inamdar
Dayananda Sagar Academy of Technology and Management, Bengaluru, Karnataka, India
vijaylaxmi.inamdar@gmail.com

S. G. Shaila
School of Engineering, Dayananda Sagar University, Bengaluru, Karnataka, India
shaila-cse@dsu.edu.in (corresponding author)

Received: 18 September 2025 | Revised: 17 October 2025 | Accepted: 24 October 2025

 DOI: https://doi.org/10.48084/etasr.14887

ABSTRACT

Breast cancer remains one of the most prevalent diseases that affects women worldwide. Early diagnosis is crucial, as it significantly reduces the risk of cancer spreading to lymph nodes or distant organs. This study presents an approach that integrates Artificial Intelligence (AI) to develop an effective diagnostic decision support system based on features extracted from biopsy tissue samples. A hybrid model of Evolutionary Programming (EP) with Radial Basis Function (RBF) Neural Network (NN) was developed to enhance robustness and classification accuracy. Although NNs are capable of learning complex patterns from previously unseen data, traditional gradient-based learning methods are prone to getting trapped in local minima. To address this limitation, an Evolutionary Programming Adaptive Radial Basis Function (EPARBF) is incorporated to provide adaptive learning by optimizing kernel parameters and output layer weights of the RBF neural network using a Differential Evolution (DE)-based heuristic approach. The proposed approach uses a fully adaptive RBF network on Fine-Needle Aspiration (FNA) biopsy data and classifies breast cancer cases as benign or malignant. The model achieved high sensitivity and specificity across both training and test datasets, significantly outperforming static RBF models and Gradient-based adaptive networks, demonstrating strong generalization and classification capability.

Keywords-biopsy; breast cancer; benign; malignant; neural network; evolutionary programming; radial basis function

I. INTRODUCTION

Breast cancer is likely the most common and lethal cancer in women worldwide. It results from the occurrence of uncontrolled growth of cells in the breast gland, in most cases in the milk ducts (ductal carcinoma) or in the lobules (lobular carcinoma) that secrete the milk. The disease may be of variable aggressiveness, locally invasive within the breast, or metastasizing to lymph nodes and other distant parts [1]. Early diagnosis plays a significant role in improving cure rates and benefits of treatment. Various diagnostics and screening techniques, such as mammography, ultrasound, Magnetic Resonance Imaging (MRI), and biopsy, are utilized to identify and confirm the occurrence of breast cancer. Mammography is the most frequent screening device, but it has limitations in women with dense breasts [2].

Treatment is usually planned according to several clinical parameters, including size, grade, stage, hormone receptor status, and overall health of the patient. Treatment options include surgery, chemotherapy, radiotherapy, hormone therapy, and targeted immunotherapy, often used in combination for optimal effect. Although mammography has been established as the gold standard screening method, it is vulnerable to high false negative rates, particularly in patients with dense breasts. Additional procedures, such as diagnostic mammograms, ultrasound, MRI, and biopsy, are utilized to improve diagnostic accuracy. Despite advances in imaging devices, human interpretation of mammograms is prone to errors in diagnosis. Although biopsies are considered accurate and specific—where cytologists examine breast tissue samples—they provide a decision accuracy of 62% to 89% [3], leaving room for error in diagnosis and treatment planning. Thus, there is a demand for more objective, reproducible, and robust diagnostic instruments that can facilitate clinical decision-making. With an increase in the incidence rate and diagnostic complexity, the need for improved, automated, and intelligent diagnostic systems for the effective and early detection of breast cancer grows. The increasing demand for improved and automated methods for the detection of breast cancer has led to extensive research. Several classifiers, including Probabilistic Neural Networks

(PNN) and Support Vector Machines (SVM), have been tested in combination with feature extraction methods, such as Principal Component Analysis (PCA), to distinguish between malignant and benign tumors [4]. Genetic markers, such as single-nucleotide polymorphisms in the BRCA1, BRCA2, and TP53 genes, have also been tested for their ability to identify the presence of cancer [5]. Image diagnosis has also grown with the assistance of segmented regions in mammograms, especially the CranioCaudal (CC) and Mediolateral Oblique (MLO) views [6], and machine learning algorithms have been used to assess patients' reactions to neoadjuvant chemotherapy [7]. Neural models such as Multi-Layer Perceptron (MLP) with meta-plasticity [8] and comparative algorithm studies such as C4.5, SVM, Naïve Bayes (NB), and K-Nearest Neighbor (KNN) [9-11] also emphasize the diversity of techniques employed for classifying breast cancer. Recurrence prediction has also been performed using machine learning [12], genetic analysis via microarrays [13], and attribute selection to identify active genes [14].

Software such as WEKA has enabled clustering to be performed with minimal preprocessing, and weighted K-means SVMs have been used on breast and kidney cancer datasets [15, 16]. Other studies have presented survival outcome prediction methods using rule-based and machine learning methods [17], and ensemble methods such as SVM ensembles [18]. Other works utilized Feedforward Back-Propagation Neural Networks (FFBPN) [19, 20], evaluated breast cancer mortality [21], and highlighted the importance of preoperative MRI in minimizing positive resection margins in screen-detected cancer [22]. Deep learning architectures such as CNNs, when utilized with AlexNet and ResNet-50 architectures, have learned to recognize Toll-Like Receptors from mammograms, and thermography and ABUS imaging have improved early detection in dense breasts [23-25]. Broader issues include obstetric and oncologic problems in pregnant patients, multimodal treatments for primary breast cancer, and genetic risk assessments for inherited conditions, particularly the BRCA mutation [26-28]. In [29], Wiener filter-based denoising, log transform-based improvement, UNet++ segmenting with RMS Prop optimization, ConvNeXtTiny feature extraction, GCRNN-based prediction, and hyperparameter optimization using Aquila Optimizer (AO) were used. In [30, 31], automated benign and malignant breast tumor classification was proposed from mammogram images based on the INBreast and CBIS-DDSM datasets. Multi-level Otsu thresholding was employed for ROI segmentation, extracting 18 structural and morphological descriptors, and SVM and ANN classification, tested using stratified cross-validation, showed good performance.

The literature above has been plagued by data imbalances, where there are usually fewer malignant samples than benign ones, resulting in biased classification. Additionally, models such as SVM or PNN were plagued by overfitting, where high-dimensional medical data made the models memorize rather than generalize. SVM or KNN classifiers are very sensitive, and their effectiveness is usually based on carefully normalized and preprocessed data. In addition, the above approaches lost information when PCA eliminated biologically significant features in dimensionality reduction.

The primary goals of this study were:

1. Develop a system to assist pathologists and clinicians by providing systematic, automated, and data-based predictions, potentially decreasing diagnostic delays and errors.
2. Overcome the issue of local minima and attain high sensitivity and specificity in training and test sets.
3. Assure the classification performance and robustness of the model with respect to various dataset divisions, ensuring its reliability in real-world diagnosis cases.

Therefore, following these objectives, the contributions of this study are as follows.

1. Develops an AI-based system that uses features extracted from Fine Needle Aspiration (FNA) biopsy tissue samples for breast cancer classification.
2. Designs an adaptive RBFNN with kernel parameters and output weights tuned for stronger learning.
3. Overcomes the local minima problem of conventional gradient-based training by incorporating Evolutionary Programming (EP) to provide global optimization.
4. Presents a new hybrid scheme blending RBFNN with EP to transcend the limitations of conventional network training.
5. Evaluates the proposed model on an established FNA Biopsy dataset to classify benign and malignant tumors.

II. PROPOSED METHOD

Among various diagnostic approaches, machine learning algorithms, and especially NNs, have been effective in classifying malignant and benign breast tumors based on cytological characteristics of FNA biopsy samples. RBF neural networks are extremely suitable for medical diagnosis due to their capacity to approximate nonlinear relationships with a simple architecture. This study introduces a sophisticated diagnostic decision support system that combines AI methods to analyze features extracted from biopsy tissue samples. Although NNs are effective in learning intricate patterns from complex datasets, traditional gradient-based training methods are often hindered by local minima. To address this limitation, Evolutionary Programming (EP) is incorporated to facilitate adaptive learning by optimizing the kernel parameters and output weights of an RBF NN.

This is achieved using a Differential Evolution-based heuristic approach. The proposed EPARBF hybrid model combines adaptive RBF NNs with EP and was applied to FNA biopsy data to classify breast cancer cases as benign or malignant, improving classification accuracy and enhancing robustness. The outcome of EPARBF is compared with a predefined threshold to have a decision outcome. This obtained decision can be given back to a cytologist to compare his decision, and, if needed, the cytologist can go for a re-evaluation. The EPARBF model demonstrated high sensitivity and specificity across both training and test datasets,

significantly outperforming conventional Static RBF (SRBF) models and Gradient-based Adaptive RBFs (GARBF) networks. Its generalizability and classification performance were further validated for effectiveness and adaptability. Figure 1 presents the proposed architecture. The preprocessing of the features was performed through the normalization process defined as:

$$x_{nor,i} = \frac{(x_i - x_{min})}{(x_{max} - x_{min})} \quad (1)$$

where x_{max} is the maximum value and x_{min} is the minimum value of an FNA biopsy feature parameter.

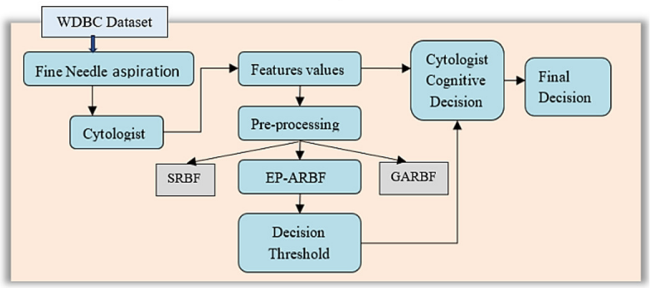

Fig. 1. Proposed approach to provide a decision aid to cytologists.

A. FNA Biopsy Features for Breast Cancer Analysis

FNA biopsy is an invasive diagnostic procedure used to extract cells or fluid from a suspicious lump or mass in the breast for microscopic examination. It is the primary diagnostic approach when a mass is detected during clinical examination, such as mammography or ultrasound imaging. A thin, hollow needle is inserted into the mass or lymph node, and cellular material is aspirated. The sample is then smeared on glass slides and examined under a microscope by a cytopathologist. From the aspirated sample, various morphological features of epithelial cells are analyzed, such as cell size, uniformity, clump thickness, nuclear size and shape, nucleoli visibility, mitosis count, and chromatin pattern. These are quantitatively encoded in the Wisconsin Diagnostic Breast Cancer (WDBC) dataset as radius, texture, concavity, etc. Considering practical diagnostic relevance, the proposed approach used the WDBC dataset [32] for breast cancer analysis. The FNA biopsy dataset consists of 569 patient cases, of which 212 samples are malignant and 357 are benign. In general, 30 features such as radius, texture, perimeter, area, smoothness, compactness, concavity, concave points, symmetry, and fractal dimension are considered from FNA of breast masses and are intended for binary classification—benign vs. malignant tumors. Therefore, FNA biopsy is a useful diagnostic intervention in the management of breast cancer because it is quick, safe, and readily available. Although it is restrictive in analyzing tissue architecture, its cell-level specificity, when digitized and translated, makes it suitable for AI-assisted diagnosis and investigation of early detection models.

B. Mapping FNA Biopsy Attributes to SRBF, GARBF, and EPARBF Networks

Out of 30 available diagnostic FNA biopsy attributes/features, 11 different attributes of biopsy lesions were selected, including the patient's age. This choice is determined by their statistical significance and low redundancy. Recursive elimination and correlation methods were employed for selecting the most discriminative features. The use of the age parameter was inspired by its proven clinical utility in evaluating the risk for breast cancer. This subset of features offered a reasonable balance between model performance and interpretability, as illustrated in Figure 2, which provides 10 cytological features evaluated from breast biopsy data. Each is binary-encoded (0 or 1), and these features are used as inputs to the RBF networks. The universal approximation capability of NNs makes them well-suited for mapping a wide range of functions across domains. In this context, the RBF NNs are efficient due to their architectural simplicity and the incorporation of nonlinear kernel functions. The mathematical representation of an RBF can be presented as:

$$y(x) = \sum_{i=1}^{N} w_i . \phi(||x - c_i|| + b) \quad (2)$$

where $y(x)$ is the output of the RBF for the input vector x, N is the number of hidden neurons of the RBF, w_i is the weight connecting the i^{th} hidden neuron to the output, $\phi(\cdot)$ is the RBF (Gaussian), $\| x - ci \|$ is the Euclidean distance between input x and the center c_i of the i^{th} RBF, and b is the bias at the output layer.

The output represents the summation of the weighted value of mapped information through RBF. This study considered the Gaussian function, defined as:

$$\phi(||x - c_i||) = \exp\left(-\frac{||x - c_i||^2}{2\sigma_i^2}\right) \quad (3)$$

where σ_i is the width of the Gaussian function for the i^{th} neuron.

Considering (2) and (3), the centers of the RBF NNs are typically selected from the available dataset, and the width is assigned a fixed value. As a result, the learning process involves only adjusting the connection weights of the output layer. However, this static nature of the RBF NNs can lead to slow or suboptimal learning, particularly when dealing with complex or irregular data landscapes. To address this deficiency, this study employs an Adaptive RBF model, in which the centers and widths of the RBFs are optimized for the output weights. This additional learning aspect enables the network to adjust to the underlying data distribution, leading to faster convergence and enhanced learning performance. Therefore, the adaptation process is recast as an optimization problem, where the aim is to minimize the output error, as specified in (4). Adaptive optimization of this error results in improved accuracy and generalization for breast cancer classification tasks.

$$E = \frac{1}{2}\sum_{J=1}^{P}\left(y_J - \hat{y}_J\right)^2 \quad (4)$$

where E is the total error, P is the number of training samples, y_J is the actual target output, and $\hat{y}_J$ is the predicted output of the network for the j^{th} input.

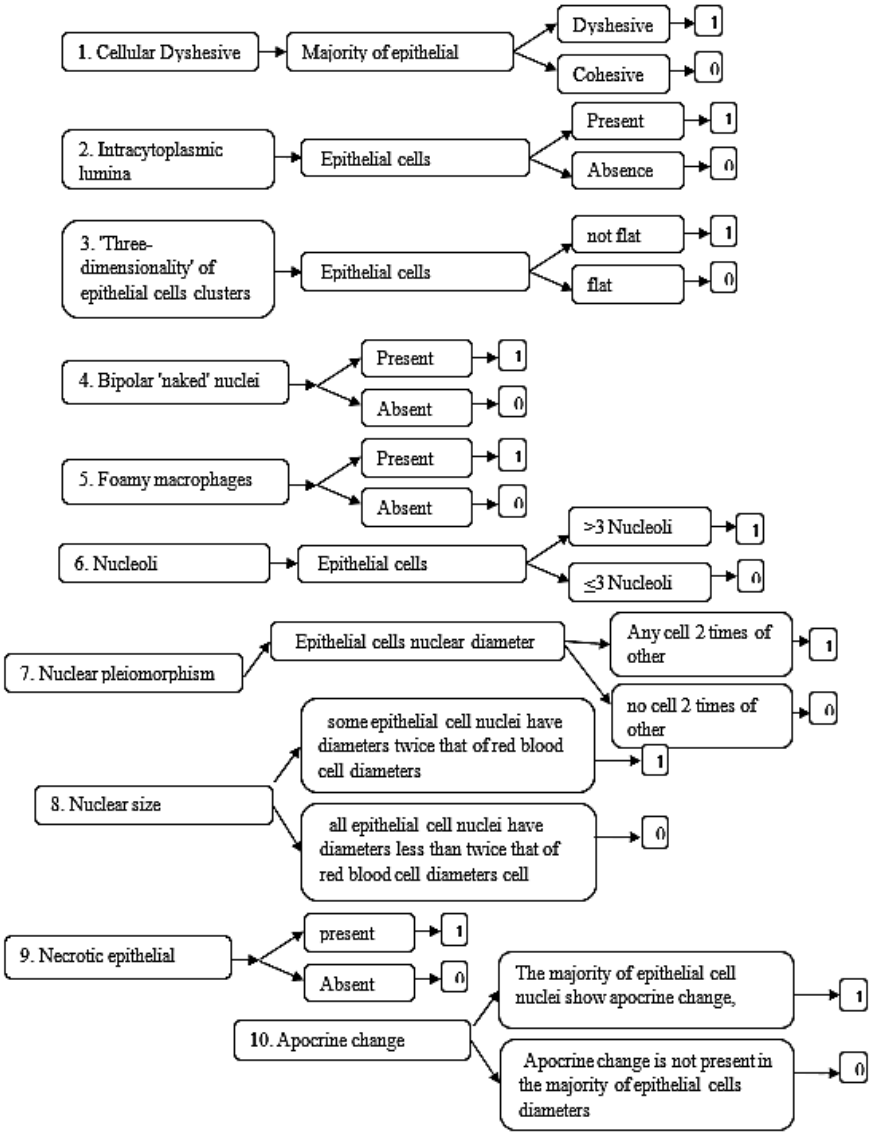

Fig. 2. FNA Biopsy attributes.

The learning landscape for neural networks tends to consist of several local optima, termed the point of convergence to the best possible solution, where gradient-based approaches tend to get stuck, resulting in suboptimal convergence or even prolonged training. Meta-heuristic algorithms can effectively overcome such constraints by increasing exploration with large jumps in the search space. Unlike gradient-based techniques, meta-heuristics are insensitive to discontinuities of the objective function and are not required to use fine-grained problem-specific formulations. The majority of meta-heuristics are based on natural computing phenomena and have a generic framework consisting of two main steps: mutation and selection. This is an iterative process that leads the population to optimal or nearly optimal solutions through subsequent generations, mathematically formulated as:

$$P(t+1) = f_s(f_y(P[t])) \quad (5)$$

In this case, the population $P[t]$ at generation t is mutated by a random process represented as f_y, and the selection of survivors to the next generation is carried out by the selection function f_s. The structure of these two operators—mutation and selection—is decisive in the efficiency of the evolutionary model. Optimal convergence requires a well-balanced combination of mutation and selection. Various classes of evolutionary computation employ different strategies for generating solutions. The model here suggests EP because it is dependent mostly on mutation alone for producing offspring. Of all mutation strategies, mutation through Gaussian distribution-based perturbation is most commonly used and considered very effective. This method poses variations around the mean with great probability close to the center and lesser probability with increasing distance from the mean, enabling controlled exploration of the search space. In the self-adaptive

version of EP, the mutation strategy provides the generation of new candidate solutions adaptively and in a controlled manner.

This is carried out utilizing a Gaussian distribution such that both the mutation step and the corresponding standard deviation change dynamically throughout the optimization. The mathematical formulations of this behavior are as follows:

$$\sigma_i' = \sigma_i . \exp(\tau . N(0,1)) \quad (6)$$

$$x_i' = x^i + \sigma_i' . N(0,1) \quad (7)$$

where x_i represents the current value of the i^{th} parameter, $x_i{}'$ represents the mutated value of the i^{th} parameter, σ_i represents the current standard deviation for x_i, $\sigma_i{}'$ represents the updated standard deviation after mutation, τ represents the learning rate parameter (often set to $1/\sqrt{2\sqrt{n}}$, where n is the dimensionality), and $N(0,1)$ represents a random number drawn from a standard normal distribution (mean 0, variance 1).

These equations ensure that both the solution vector and the mutation strength co-evolve during the optimization process, which enhances the algorithm's ability to adaptively balance exploration and exploitation. EP was employed to evolve the parameters of the RBF NN architecture. In the proposed configuration, the RBF network consists of 11 input nodes, 4 hidden nodes, and 1 output node. This structure results in a total of 52 parameters to be optimized, 44 corresponding to the weights connecting the input nodes to the hidden layer (i.e., the centers of the RBF), 4 values representing the spread (σ) of the Gaussian kernels for each hidden node, and 4 values corresponding to the weights from the hidden layer to the output node. Thus, each solution (individual) in the EP population is encoded as a numeric array of length 52. The array is organized such that the first 44 entries represent the hidden layer weights (w_h), the next 4 entries represent the spreads of the RBF (σ), and the final 4 entries represent the output layer weights (w_o), This parameter representation is depicted in Figure 3, which outlines the structure of an individual solution within the EP framework.

The complete functional block diagram of the proposed EPARBF (Evolutionary Programming-based Adaptive Radial Basis Function) model, including the flow of mutation and selection processes, is illustrated in the functional block diagram in Figure 4. Initially, a randomly generated population of candidate solutions is created, where each individual has a fixed length equal to the total number of parameters in the RBF architecture. For each solution, the centers, spreads, and output layer weights are extracted and assigned to the corresponding RBF network. Given a set of input features, the RBF network computes the output using these parameters. Subsequently, EP is applied to generate offspring by mutating current solutions. The fitness of each solution is evaluated, and a tournament selection strategy is used to select the best-performing individuals from the combined pool of parents and offspring. This process is repeated for all individuals in the population to build the next generation. The entire evolutionary cycle—comprising mutation, evaluation, and selection—is then repeated iteratively across generations until the optimization converges or a predefined termination condition is met.

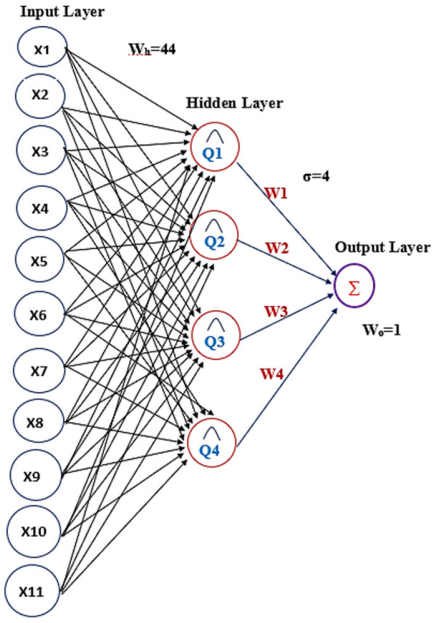

Fig. 3. Structure of EP with the RBF NN.

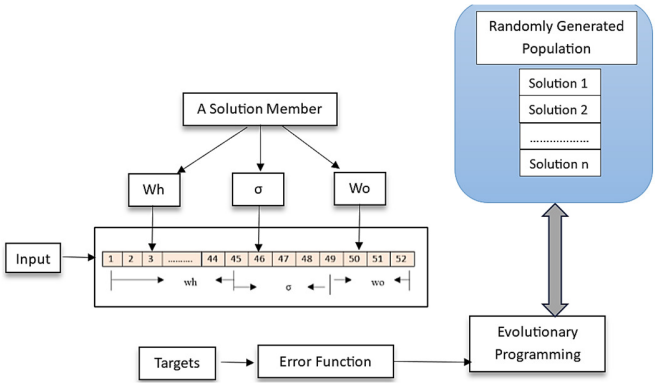

Fig. 4. Functional block diagram of EPARBF.

III. RESULTS AND DISCUSSION

The dataset consists of 11 diagnostic features for 569 patient examples, of which 212 are malignant and 357 are benign. Random sampling is first used to select 100 benign and 100 malignant samples for training and testing for a balanced selection, which preserves class symmetry by ensuring unbiased learning. The stratified random sampling approach was first adopted to preserve a balanced representation of classes, in which 100 benign and 100 malignant samples were chosen for the initial experiments to analyze convergence behavior and parameter sensitivity.

A. XOR Classification for RBF Networks

The XOR classification problem is a well-known benchmark for measuring classifier performance because of its nonlinear separability. Although NN-based classifiers have nonlinear mapping capabilities, they tend to be unable to converge when the number of hidden nodes is set at two, as conventional learning algorithms are unable to find good weight settings. In this study, an RBF network with

architecture [2, 2, 1], with 2 input nodes, 2 hidden nodes, and 1 output node, was tested using SRBF and GARBF. GARBF investigated variations in learning rates (0.2, 0.5, and 0.9) to determine their effects on convergence. The proposed EPARBF model employed EP for parameter optimization, with an RBF architecture of [11, 4, 1], with 11 input nodes, 4 hidden nodes, and 1 output node. All algorithms were subjected to 10 separate trials to provide stochastic variability, with up to 1000 iterations in each. The findings are presented in Tables I through III. Interestingly, the SRBF did not consistently converge for most trials, as shown in Table I. The GARBF improved convergence for lower learning rates (lr=0.2) but exhibited inconsistency at higher rates (lr=0.4 and 0.9), as shown in Table II. In contrast, the proposed EPARBF demonstrated consistently optimal and stable convergence, producing clear classification boundaries and robust results as represented in Table III.

TABLE I. PERFORMANCEOF SRBF OVER 10 TRIALS

Trials	SRBF			
	Inputs			
	[0,0]	[0,1]	[1,0]	[1,1]
1	0.1234	0.9825	0.0178	0.1361
2	x	x	x	x
3	0.1312	0.0166	0.9744	0.1333
4	0.1354	0.9889	0.0156	0.1145
5	x	x	x	x
6	0.1342	0.9819	0.0178	0.1363
7	x	x	x	x
8	0.1239	0.9424	0.0125	0.1789
9	0.1278	0.9475	0.0157	0.1791
10	x	x	x	x

x represents non-convergence

TABLE II. PERFORMANCE OF GARBF OVER 10 TRIALS WITH DIFFERENT LEARNING RATES (LR)

Trials	GARBF (lr=02)				GARBF (lr=0.4)				GARBF (lr=0.9)			
	Inputs				Inputs				Inputs			
	[0,0]	[0,1]	[1,0]	[1,1]	[0,0]	[0,1]	[1,0]	[1,1]	[0,0]	[0,1]	[1,0]	[1,1]
1	0.0041	1.0000	0.9998	0.0062	x	x	x	x	0.7292	1.3169	1.0105	0.001
2	0.0053	0.9999	0.9998	0.0073	0.000	0.6408	0.3051	0.0823	0.7315	1.3149	1.0122	0.000
3	0.0058	1.0000	1.0000	0.0061	x	x	x	x	x	x	x	x
4	0.1717	1.1342	1.0429	0.0057	0.0001	0.6403	0.3042	0.0820	0.7316	1.3147	1.0131	0.0002
5	0.0065	0.9974	0.9726	0.0069	0.0022	1.0003	0.9998	0.0020	x	x	x	x
6	x	x	x	x	x	x	x	x	0.0029	0.9998	1.0003	0.0023
7	0.0081	0.9466	0.9351	0.1781	0.0000	0.6400	0.3046	0.0821	0.0014	1.0000	1.0000	0.0016
8	0.0083	0.9522	0.9552	0.1823	0.0001	0.6401	0.3047	0.0820	x	x	x	x
9	0.0091	0.9532	0.9756	0.1886	0.0001	0.6399	0.3045	0.0820	0.0073	0.5563	0.3068	0.1280
10	0.0097	0.9990	0.9999	0.0099	x	x	x	x	0.0074	0.5562	0.3069	0.1279

TABLE III. PERFORMANCE OF EPARBF OVER 10 TRIALS

Trials	EPARBF			
	Inputs			
	[0,0]	[0,1]	[1,0]	[1,1]
1	0.0002	0.9997	1.0001	0.0003
2	0.0007	1.0000	1.0001	0.0001
3	0.0006	1.0001	0.9997	0.0003
4	0.0004	1.0004	1.0006	0.0003
5	0.0003	1.0007	0.9997	0.0011
6	0.0002	1.0008	0.9998	0.0012
7	0.0001	0.9993	1.0001	0.0001
8	0.0005	1.0004	0.9999	0.0001
9	0.0001	0.9996	1.0002	0.0001
10	0.0001	1.0032	1.0000	0.0001

B. Breast Cancer Recognition with the Proposed EPARBF Model

In addition, to broaden the experiments, the whole dataset was taken into account, and a 10-fold cross-validation process was used to provide an unbiased and thorough assessment to enhance generalization and resilience. In this context, the dataset was split into ten folds with the class ratios maintained; the model was trained on nine folds and tested on one, until each fold had been utilized for testing. To reduce random effects, the whole cross-validation procedure was run five times with varying random seeds. Model performance was measured using Sensitivity, Specificity, Accuracy, Precision, F1-Score, and ROC-AUC metrics to provide a comprehensive performance evaluation. Sensitivity evaluates the model's capacity to identify true positives accurately, whereas specificity evaluates its capability to identify true negatives accurately. The ROC-AUC (Receiver Operating Characteristic–Area Under Curve) was calculated to assess the overall discriminative ability of the model using the area under the plot of True Positive Rate (TPR) against False Positive Rate (FPR). The higher the value of ROC-AUC, the better the classification performance and stability at various thresholds. These measures, described in (8)-(12), are well-accepted in medical diagnostics, and values close to 100% are regarded clinically reliable.

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP+FN}} \quad (8)$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN+FP}} \quad (9)$$

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \quad (10)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \quad (11)$$

$$\text{F1} - \text{score} = 2 \times \frac{\text{Precision} \times \text{Sensitivity}}{\text{Precision+Sensitivity}} \quad (12)$$

where TP (True Positive) represents correctly identified malignant cases, TN (True Negative) represents correctly identified benign cases, FP (False Positive) denotes benign

samples misclassified as malignant, and FN (False Negative) denotes malignant samples misclassified as benign.

Table IV shows the comparative performance of the SRBF, GARBF, and proposed EPARBF models over five random seeds using 10-fold stratified cross-validation. The results show the mean and standard deviation of all five runs, thus maintaining statistical consistency. Throughout the folds, the proposed EPARBF model attained an average 98.25±0.09 accuracy, 97.50±0.13 sensitivity, 98.00±0.10 specificity, 98.40±0.09 precision, 97.95±0.09 F1-score, and a ROC-AUC of 0.992±0.005. The proposed EPARBF model achieved higher sensitivity, specificity, accuracy, precision, and F1-score with less variation throughout the trials, which shows it to be stable, robust, and having good generalizability. The exploration ability of EPARBF allowed the network to find more optimal configurations of weight and kernel parameters, reducing the chances of being trapped in a local minima. Figure 5 shows ROC curves that compare the classification performance of SRBF, GARBF, and EPARBF models. The EPARBF model exhibits the highest ROC-AUC value (≈0.992), indicating superior discriminative capability and robustness across varying decision thresholds. The GARBF model demonstrates moderate performance due to improved adaptability through gradient-based learning, whereas the SRBF curve remains closest to the diagonal, reflecting weaker separability between malignant and benign classes. The steeper and more convex shape of the EPARBF curve near the upper-left corner confirms its enhanced sensitivity and lower false positive rate**,** establishing its effectiveness for reliable breast cancer diagnosis.

TABLE IV. COMPARATIVE ANALYSIS OF SRBF, GARBF, AND EPARBF MODELS ACROSS FIVE RANDOM SEEDS USING 10-FOLD STRATIFIED CROSS-VALIDATION

Model	Random seed	Sensitivity (%)	Specificity (%)	Accuracy (%)	Precision (%)	F1-score (%)
SRBF	Seed 1	81.94	83.15	84.12	82.46	82.19
	Seed 2	82.10	82.87	84.35	83.22	82.66
	Seed 3	81.72	83.04	84.05	82.56	82.13
	Seed 4	82.28	82.91	84.41	83.37	82.82
	Seed 5	81.95	83.20	84.26	82.48	82.26
	Mean±SD	**81.99±0.22**	**83.03±0.14**	**84.24±0.14**	**82.82±0.36**	**82.41±0.28**
GARBF	Seed 1	91.85	94.25	92.96	93.11	92.46
	Seed 2	92.14	94.68	93.22	93.55	92.82
	Seed 3	91.75	94.37	93.01	93.27	92.51
	Seed 4	92.36	94.42	93.26	93.62	92.87
	Seed 5	91.98	94.33	93.17	93.35	92.64
	Mean±SD	**92.02±0.22**	**94.41±0.16**	**93.12±0.11**	**93.38±0.18**	**92.66±0.17**
EPARBF	Seed 1	97.42	97.85	98.10	98.23	97.82
	Seed 2	97.68	97.92	98.32	98.45	98.06
	Seed 3	97.35	98.12	98.21	98.51	97.92
	Seed 4	97.59	98.06	98.19	98.37	97.97
	Seed 5	97.46	97.87	98.36	98.46	98.01
	Mean±SD	**97.50±0.13**	**98.00±0.10**	**98.25±0.09**	**98.40±0.09**	**97.95±0.09**

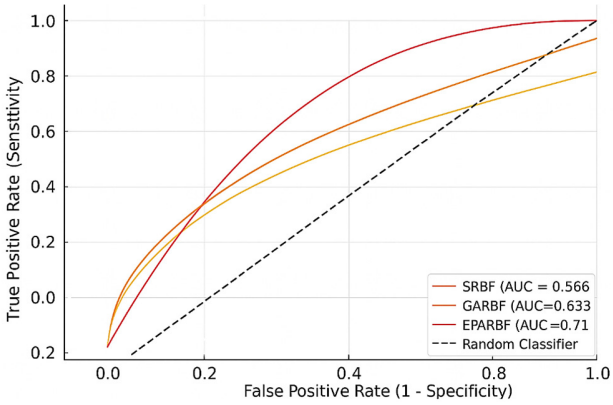

Fig. 5. ROC-AUC curves for SRBF, GARBF, and EPARBF models.

The proposed EPARBF model was compared with recent RBF-related methods, as shown in Table V. In [33], an RBF-kernel SVM on MammoWave microwave frequency responses achieved an accuracy of 91.0% (sensitivity 84.4%, specificity 97.2%), as proof of the applicability of RBF techniques in non-conventional sensing data. In [34], an RBFN was optimized using the Mayfly meta-heuristic along with CLAHE preprocessing on breast MRI, achieving an accuracy of 97.54% (this study reports accuracy and lacks a complete set of per-class metrics). In [35], ensembles containing RBFN on WDBC tabular features achieved an accuracy of 95.34% and a sensitivity of 93.05%, giving a direct comparison on a similar feature set with that used in this work. In [36], an accuracy of 96.7%, a sensitivity of 94.9%, and a specificity of 97.8% were obtained.

TABLE V. COMPARATIVE PERFORMANCE ANALYSIS

Models	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F1-score (%)
[33]	91.00	84.41	97.22	88.58	86.40
[34]	97.54	96.10	98.10	96.50	96.30
[35]	95.33	93.05	96.50	94.02	93.50
[36]	96.71	94.90	97.82	95.63	95.32
Proposed EPARBF	**98.25**	**97.50**	**98.00**	**98.40**	**97.95**

Comparatively, the optimized EPARBF yielded the best overall accuracy (98.25%) with well-balanced sensitivity and specificity, substantiating the efficacy of evolution-based parameter optimization over heuristic or static RBF variants in lowering vulnerability to local minima and enhancing generalization over folds. This comparison demonstrates the

superior convergence stability, adaptability, and generalization capability of the proposed EPARBF approach within computer-aided diagnosis frameworks.

IV. CONCLUSION

This study introduced a diagnostic decision-making framework powered by an Evolutionary Programming Adaptive Radial Basis Function (EPARBF) network to classify breast cancer precisely and reliably from cytological features extracted from FNA biopsy data. The novelty of this study is the integration of EP and an adaptive RBF NN to optimize kernel centers, spreads, and weights globally to avoid local minima and convergence problems that are common in training RBFs. A comprehensive experiment with the WDBC dataset using 10-fold stratified cross-validation with five random seeds confirmed the outperformance of the developed EPARBF model. The framework achieved an accuracy of 98.25%, a sensitivity of 97.50%, a specificity of 98.00%, a precision of 98.40%, and an F1-score of 97.95%, which is superior compared to state-of-the-art RBF-based solutions. The enhancement of classification performance and fold consistency validate the robustness, adaptability, and diagnostic reliability of the developed method. The EPARBF model can offer computer-aided diagnosis systems with intelligent functionality by integrating global evolutionary optimization and adaptive kernel learning to achieve high generalizability and clinical interpretability.

REFERENCES

[1] "Cancer," *World Health Organization*. https://www.who.int/health-topics/cancer.

[2] H. Hussein *et al.*, "Supplemental Breast Cancer Screening in Women with Dense Breasts and Negative Mammography: A Systematic Review and Meta-Analysis," *Radiology*, vol. 306, no. 3, Mar. 2023, Art. no. e221785, https://doi.org/10.1148/radiol.221785.

[3] E. D. Pisano *et al.*, "Fine-Needle Aspiration Biopsy of Nonpalpable Breast Lesions in a Multicenter Clinical Trial: Results from the Radiologic Diagnostic Oncology Group V," *Radiology*, vol. 219, no. 3, pp. 785–792, Jun. 2001, https://doi.org/10.1148/radiology.219.3.r01jn28785.

[4] A. Osareh and B. Shadgar, "Machine learning techniques to diagnose breast cancer," in *2010 5th International Symposium on Health Informatics and Bioinformatics*, Ankara, Turkey, 2010, pp. 114–120, https://doi.org/10.1109/HIBIT.2010.5478895.

[5] S. Silva, O. Anunciação, and M. Lotz, "A Comparison of Machine Learning Methods for the Prediction of Breast Cancer," in *Evolutionary Computation, Machine Learning and Data Mining in Bioinformatics*, vol. 6623, C. Pizzuti, M. D. Ritchie, and M. Giacobini, Eds. Springer Berlin Heidelberg, 2011, pp. 159–170.

[6] R. Ramos-Pollán *et al.*, "Discovering Mammography-based Machine Learning Classifiers for Breast Cancer Diagnosis," *Journal of Medical Systems*, vol. 36, no. 4, pp. 2259–2269, Aug. 2012, https://doi.org/10.1007/s10916-011-9693-2.

[7] S. Mani *et al.*, "Machine learning for predicting the response of breast cancer to neoadjuvant chemotherapy," *Journal of the American Medical Informatics Association*, vol. 20, no. 4, pp. 688–695, Jul. 2013, https://doi.org/10.1136/amiajnl-2012-001332.

[8] J. Fombellida, S. Torres-Alegre, J. A. Piñuela-Izquierdo, and D. Andina, "Artificial Metaplasticity for Deep Learning: Application to WBCD Breast Cancer Database Classification," in *Bioinspired Computation in Artificial Systems*, vol. 9108, J. M. Ferrández Vicente, J. R. Álvarez-Sánchez, F. De La Paz López, Fco. J. Toledo-Moreo, and H. Adeli, Eds. Springer International Publishing, 2015, pp. 399–408.

[9] H. Asri, H. Mousannif, H. A. Moatassime, and T. Noel, "Using Machine Learning Algorithms for Breast Cancer Risk Prediction and Diagnosis," *Procedia Computer Science*, vol. 83, pp. 1064–1069, 2016, https://doi.org/10.1016/j.procs.2016.04.224.

[10] T. Saba, "Recent advancement in cancer detection using machine learning: Systematic survey of decades, comparisons and challenges," *Journal of Infection and Public Health*, vol. 13, no. 9, pp. 1274–1289, Sep. 2020, https://doi.org/10.1016/j.jiph.2020.06.033.

[11] S. A. Mohammed, S. Darrab, S. A. Noaman, and G. Saake, "Analysis of Breast Cancer Detection Using Different Machine Learning Techniques," in *Data Mining and Big Data*, vol. 1234, Y. Tan, Y. Shi, and M. Tuba, Eds. Springer Singapore, 2020, pp. 108–117.

[12] P. H. Abreu, M. S. Santos, M. H. Abreu, B. Andrade, and D. C. Silva, "Predicting Breast Cancer Recurrence Using Machine Learning Techniques: A Systematic Review," *ACM Computing Surveys*, vol. 49, no. 3, Jul. 2016, Art. no. 52, https://doi.org/10.1145/2988544.

[13] B. Bektaş and S. Babur, "Machine learning based performance development for diagnosis of breast cancer," in *2016 Medical Technologies National Congress (TIPTEKNO)*, Antalya, Turkey, Oct. 2016, pp. 1–4, https://doi.org/10.1109/TIPTEKNO.2016.7863129.

[14] N. Kolay and P. Erdoğmuş, "The classification of breast cancer with Machine Learning Techniques," in *2016 Electric Electronics, Computer Science, Biomedical Engineerings' Meeting (EBBT)*, Istanbul, Turkey, Apr. 2016, pp. 1–4, https://doi.org/10.1109/EBBT.2016.7483683.

[15] S. Kim, "Weighted K-means support vector machine for cancer prediction," *SpringerPlus*, vol. 5, no. 1, Dec. 2016, Art. no. 1162, https://doi.org/10.1186/s40064-016-2677-4.

[16] M. Montazeri, M. Montazeri, M. Montazeri, and A. Beigzadeh, "Machine learning models in breast cancer survival prediction," *Technology and Health Care*, vol. 24, no. 1, pp. 31–42, Jan. 2016, https://doi.org/10.3233/THC-151071.

[17] M. W. Huang, C. W. Chen, W. C. Lin, S. W. Ke, and C. F. Tsai, "SVM and SVM Ensembles in Breast Cancer Prediction," *PLOS ONE*, vol. 12, no. 1, 2017, Art. no. e0161501, https://doi.org/10.1371/journal.pone.0161501.

[18] L. Abdel-Ilah and H. Šahinbegović, "Using machine learning tool in classification of breast cancer," in *CMBEBIH 2017*, 2017, pp. 3–8, https://doi.org/10.1007/978-981-10-4166-2_1.

[19] J. P. Leone *et al.*, "Factors associated with late risks of breast cancer-specific mortality in the SEER registry," *Breast Cancer Research and Treatment*, vol. 189, no. 1, pp. 203–212, Aug. 2021, https://doi.org/10.1007/s10549-021-06233-4.

[20] J. J. J. Gommers *et al.*, "The Impact of Preoperative Breast MRI on Surgical Margin Status in Breast Cancer Patients Recalled at Biennial Screening Mammography: An Observational Cohort Study," *Annals of Surgical Oncology*, vol. 28, no. S3, pp. 432–432, Dec. 2021, https://doi.org/10.1245/s10434-021-10001-5.

[21] A. A. Hekal, A. Elnakib, and H. E. D. Moustafa, "Automated early breast cancer detection and classification system," *Signal, Image and Video Processing*, vol. 15, no. 7, pp. 1497–1505, Oct. 2021, https://doi.org/10.1007/s11760-021-01882-w.

[22] R. Karthiga and K. Narasimhan, "Medical imaging technique using curvelet transform and machine learning for the automated diagnosis of breast cancer from thermal image," *Pattern Analysis and Applications*, vol. 24, no. 3, pp. 981–991, Aug. 2021, https://doi.org/10.1007/s10044-021-00963-3.

[23] I. Allajbeu *et al.*, "Automated Breast Ultrasound: Technical Aspects, Impact on Breast Screening, and Future Perspectives," *Current Breast Cancer Reports*, vol. 13, no. 3, pp. 141–150, Sep. 2021, https://doi.org/10.1007/s12609-021-00423-1.

[24] O. M. William Wolberg, "Breast Cancer Wisconsin (Diagnostic)." UCI Machine Learning Repository, 1993, https://doi.org/10.24432/C5DW2B.

[25] S. Altintas and W. Tjalma, "Pregnancy-Associated Breast Cancer (PABC) and Fertility Issues in Young Women After Breast Cancer," in *Breast Cancer Essentials*, M. Rezai, M. A. Kocdor, and N. Z. Canturk, Eds. Springer International Publishing, 2021, pp. 657–665.

[26] A. K. Kattepur and K. S. Gopinath, "Management of Hereditary Breast Cancer: An Overview," in *Breast Cancer*, S. C. Sharma, A. Mazumdar, and R. Kaushik, Eds. Springer Nature Singapore, 2022, pp. 353–397.

[27] P. Hays, "Great Strides in Precision Medicine: Personalized Oncology, Immunotherapies, and Molecular Diagnostics," in *Advancing Healthcare Through Personalized Medicine*, Springer International Publishing, 2021, pp. 141–417.

[28] V. Inamdar, S. G. Shaila, and M. K. Singh, "FNAB-Based Prediction of Breast Cancer Category Using Evolutionary Programming Neural Ensemble," in *Computational Vision and Bio-Inspired Computing*, 2021, pp. 653–663, https://doi.org/10.1007/978-981-33-6862-0_51.

[29] M. Sreevani and R. Latha, "A Deep Learning with Metaheuristic Optimization-Driven Breast Cancer Segmentation and Classification Model using Mammogram Imaging," *Engineering, Technology & Applied Science Research*, vol. 15, no. 1, pp. 20342–20347, Feb. 2025, https://doi.org/10.48084/etasr.9406.

[30] V. R. Gurudas, S. G. Shaila, and A. Vadivel, "Breast Cancer Detection and Classification from Mammogram Images Using Multi-model Shape Features," *SN Computer Science*, vol. 3, no. 5, Jul. 2022, Art. no. 404, https://doi.org/10.1007/s42979-022-01290-y.

[31] V. R. Gurudas, S. G. Shaila, and A. Vadivel, "Morphological and Textural Data Fusion for Breast Cancer Classification Based on Inter and Intra group Variances," *International Journal of Intelligent Engineering and Systems*, vol. 17, no. 3, pp. 682–695, Jun. 2024, https://doi.org/10.22266/ijies2024.0630.53.

[32] "Breast Cancer Wisconsin (Diagnostic) Data Set." Kaggle, [Online]. Available: https://www.kaggle.com/datasets/uciml/breast-cancer-wisconsin-data.

[33] S. P. Rana *et al.*, "Radial Basis Function for Breast Lesion Detection from MammoWave Clinical Data," *Diagnostics*, vol. 11, no. 10, Oct. 2021, Art. no. 1930, https://doi.org/10.3390/diagnostics11101930.

[34] A. Ponraj and A. Canessane, "Radial Basis Function Networks and Contrast-Limited Adaptive Histogram Equalization Filter Based Early-Stage Breast Cancer Detection Techniques," *Journal of Computer Science*, vol. 19, no. 6, pp. 760–774, Jun. 2023, https://doi.org/10.3844/jcssp.2023.760.774.

[35] H. A. Essa, E. Ismaiel, and M. F. A. Hinnawi, "Feature-based detection of breast cancer using convolutional neural network and feature engineering," *Scientific Reports*, vol. 14, no. 1, Sep. 2024, Art. no. 22215, https://doi.org/10.1038/s41598-024-73083-7.

[36] C. Beltran-Perez, H. L. Wei, and A. Rubio-Solis, "Generalized Multiscale RBF Networks and the DCT for Breast Cancer Detection," *International Journal of Automation and Computing*, vol. 17, no. 1, pp. 55–70, Feb. 2020, https://doi.org/10.1007/s11633-019-1210-y.

AUTHORS PROFILE

Vijaylaxmi Inamdar is a Ph.D scholar in the Department of Computer Science and Engineering, having 8 years of experience in teaching in the field. Her research areas are health care, data mining and imageprocessing

S. G. Shaila is a Professor and Chairperson of CSE (Data Science) at DSU with a PhD from NIT Trichy. She has 17 years of teaching and research experience, over 45 publications, and 13 patents. Her research areas include Data Mining, Information Retrieval, Image Processing, and Computational Neuroscience.