



OPEN ACCESS

EDITED BY

Domenico Mallardo,
G. Pascale National Cancer Institute
Foundation (IRCCS), Italy

REVIEWED BY

Sabarna Choudhury,
Qualcomm, United States
Chaima Elmejjari,
Universite Hassan II Casablanca Ecole
Nationale Supérieure de l'Enseignement
Technique, Morocco

*CORRESPONDENCE

Muhammad Shafiq
✉ shafiq786@hotmail.com

RECEIVED 16 October 2025

REVISED 26 January 2026

ACCEPTED 30 January 2026

PUBLISHED 18 February 2026

CITATION

Bilal A, Obidallah WJ, Wassan S,
Albathan M, Almakki R, Alshaikh Z and
Shafiq M (2026) Fusion of genomic and
pathological data for breast cancer
detection using BCDNN.
Front. Med. 13:1726223.
doi: 10.3389/fmed.2026.1726223

COPYRIGHT

© 2026 Bilal, Obidallah, Wassan,
Albathan, Almakki, Alshaikh and Shafiq.
This is an open-access article distributed
under the terms of the [Creative
Commons Attribution License \(CC BY\)](#).
The use, distribution or reproduction in
other forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which does
not comply with these terms.

Fusion of genomic and pathological data for breast cancer detection using BCDNN

Anas Bilal¹, Waeal J. Obidallah², Sobia Wassan³,
Mubarak Albathan², Riyad Almakki², Zeyad Alshaikh² and
Muhammad Shafiq^{4*}

¹College of Information Science and Technology, Hainan Normal University, Haikou, China, ²College of Computer and Information Sciences, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh, Saudi Arabia, ³School of Equipment Engineering, Jiangsu Urban and Rural Construction Vocational College, Changzhou, China, ⁴School of Computer Science, Shandong Xiehe University, Jinan, China

Background of study: Breast cancer is one of the leading causes of mortality among women worldwide. Early and accurate detection is crucial for improving treatment outcomes and survival rates. Recent advancements in Deep Learning (DL), Artificial and Intelligence (AI), have shown promising results in medical image analysis and cancer prediction.

Purpose: This study aims to develop and evaluate a BCDNN model that classifies tumors as benign or malignant using genomic and histopathological data. The research focuses on improving diagnostic accuracy through AI-driven methods.

Method: The proposed BCDNN model was implemented in MATLAB R2016. A publicly available breast cancer dataset from Kaggle was used, encompassing both genomic and pathological features. The dataset was pre-processed and feature selected before training the BCDNN with optimized hyperparameters.

Result: The proposed model achieved a mean classification accuracy of 93.84% during cross-validation, demonstrating stable, reliable performance in distinctive between benign and malignant cases.

Conclusion: The BCDNN model shows significant promise in supporting clinical decision-making for breast cancer diagnosis. Future work may enhance model generalizability and explore integration with real-time diagnostic systems, contributing to better health outcomes for women globally. The code for this study is available on [GitHub](#).

KEYWORDS

artificial intelligence, breast cancer detection, deep learning, early diagnosis, histopathological data

1 Introduction

Breast cancer is a major health issue in the world that continues to impact a large number of women and their families. Nevertheless, there is still no certain cure for breast cancer, despite the efforts made. However, the state-of-the-art innovations in deep Learning and AI have played a main role in enhancing the early detection of breast cancer. Early detection is of the essence in enhancing patient survival and effective therapy. Another promising feature extractor based on DL that can be used instead of the usual machine learning methods is a feature extractor. It has been successful in different medical uses. Deep Learning is a form of AI that assists the computer in making sense of the information

in the same like the human brain. The deep learning algorithms process various data to produce precise predictions. Deep Learning, also known as neural networks, consists of three layers: “input, hidden, and output”. The input layer obtains data, and the output layer produce the result. Deep Learning is better than traditional machine learning with complex data. The research Elkorany and Elsharkawy (1) is dedicated to the application of CNNs to detect breast cancer. Their study used a CNNs architecture based on a selection function depending on the variance variable and a combination of the Inception-V3, ResNet50, and AlexNet. These findings suggest that the methodology of this study performs better than the other studies that used the MIAS database in the accuracy of classification. Moreover, their method allows fast and accurate categorization of mammography. ROI patches are used in the diagnosis of breast cancer. The Jaincy and Pattabiraman (2), study being proposed is the BCDCNN framework, a deep learning model that uses MRI to perform optimized segmentation, adaptive feature learning, and a novel loss function to facilitate accurate and clinically effective detection of breast cancer.

Lee et al. (3) the development of a powerful deep neural network based on transformers has fully overturned deep Learning. Their study is better than the traditional 3D section-by-section models and allows a better diagnosis of breast cancer by using the context information on the neighboring image sections. This work brings out the changing nature of the deep learning methods other than the conventional CNNs. Wassan et al. (4) research examines privacy-sensitive biomedical image classification, that is, by combining federated Learning and differential privacy with neural networks and Gaussian processes, we can securely and accurately predict the medical results (4). The author Mall et al. (5) discusses the wide range of medical image processing to use artificial intelligence and deep neural networks to make automated diagnoses related to different acute diseases, as well as provides an overview of the currently available sources of data and the prospects of future research. The number of neurons in the input and output layers of a neural network, as in the BCDNN, is an

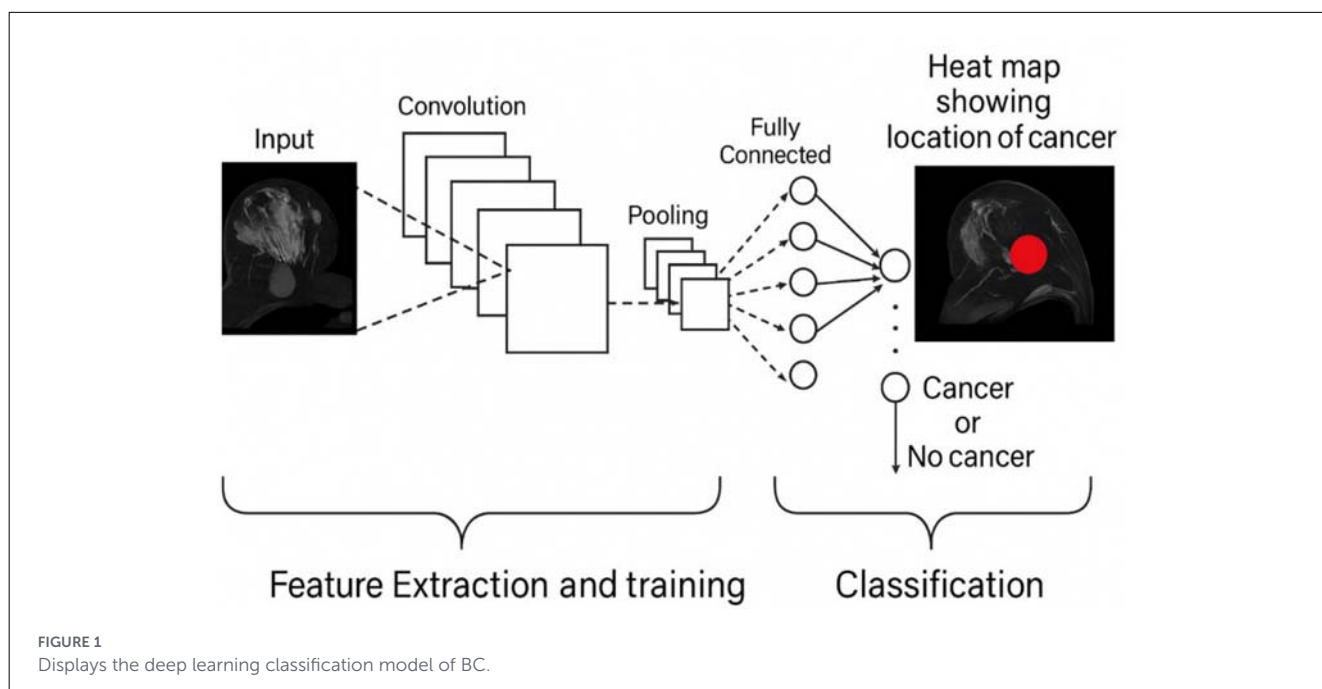
important parameter to design the model and varies according to the data and the purpose of the classification. The number of neurons in the input layer should be equal to the number of features obtained from the genomics and histopathological data. For example, there would be 10 input features that would have 10 neurons, each of which is a particular feature of the data. The classification task defines the output layer structure. In the case of binary breast cancer detection (benign vs. malignant), there is often a single neuron, and an output of more than 0.5 is considered to be a malignancy, and a result below 0.5 is considered to be a benign condition. In multiclass classification, the number of neurons in the output layer needs to be equivalent to the number of classes, and each of the neurons corresponds to a class, and the ultimate prediction is associated with the neuron that has the greatest output probability. DNNs are a subdivision of ANNs that include more than two layers in between the input and the output layers, which allows them to acquire hierarchical and complex information representations that are similar to human cognition.

1.1 Deep neural networks

In recent years, DNNs made a significant leap in different fields and areas, including computer vision, medical diagnosis, speech recognition, and NLP. The fact that they have the ability to resolve complex issues and extrapolate valuable information on massive datasets is enhanced by the fact that they have an autonomous ability to derive hierarchical aspects of the data. To identify internal cancer, genetic and histological data may be used with the assistance of a DL method. The primary goal is to forecast the outcome of breast cancer by using the BCDNN model. One of the purposes of the research improvement is the development of a customized BCDNN algorithm that can be employed to categorize breast cancer as benign or malignant cases. It is necessary to identify the problem, collect the data, and pre-process it with the help of credible sources, such as Kaggle. Then, the paper critically evaluates the precision prediction of the BCDNN model. In addition, this literature contributes to our knowledge about the way DL can be used to improve the detection of breast cancer and may result in more reliable diagnoses and a positive patient outcome. The schematic illustration of the BC Deep Learning classification Model is shown in Figure 1.

DNNs are robust machine learning algorithms that have multiple hidden layers that facilitate the successive derivation of complex, abstract data representations. DNNs are fundamentally made up of neurons (or units) that operate by computing weighted sums of inputs, performing activation functions to bring about non-linearity, and passing the output to other layers. The learning ability of the network is dependent on modification of weights and biases in the process of training, and it is mainly done through backpropagation, a process of gradient descent that minimizes a loss function. To analyze images, different types of DNNs are customized to particular tasks, including CNNs, RNNs working with sequential data, and Transformers working with natural language. The problems that are encountered when training deep networks include vanishing or exploding gradients, which are mitigated by methods such as proper weight initialization and the

Abbreviations: AI, artificial intelligence; DL, deep learning; ML, machine learning; BC, breast cancer; BCDNN, breast cancer deep neural network; BCDCNN, breast cancer deep convolutional neural network; CNN, convolutional neural network; DNN, deep neural network; ANN, artificial neural network; RNN, recurrent neural network; LSTM, long short-term memory; RELU, rectified linear unit; SOFTMAX, soft maximum function; AUC, area under the curve; PR AUC, precision–recall area under the curve; ROC, receiver operating characteristic; RMSE, root mean squared error; MSE, mean squared error; R², coefficient of determination; CV, cross-validation; TP, true positive; TN, true negative; FP, false positive; FN, false negative; FNA, fine needle aspiration; CAD, computer-aided diagnosis; MRI, magnetic resonance imaging; WSI, whole slide image; VIT, vision transformer; GRAD-CAM, gradient-weighted class activation mapping; L1, L1 regularization; L2, L2 regularization; SGD, stochastic gradient descent; GD, gradient descent; LR, learning rate; CE Loss, cross-entropy loss; GPU, graphics processing unit; TPU, tensor processing unit; CSV, comma-separated values; SD, standard deviation; SE, standard error; BUSI, breast ultrasound images dataset; MIAS, mammographic image analysis society dataset; CBIS-DDSM, curated breast imaging subset of DDSM; DDSM, digital database for screening mammography; BREAKHIS, breast cancer histopathological image dataset; BACH, breast cancer histology dataset; TCGA, the cancer genome atlas; TCGA-BRCA, breast invasive carcinoma dataset (TCGA); MATLAB, matrix laboratory.



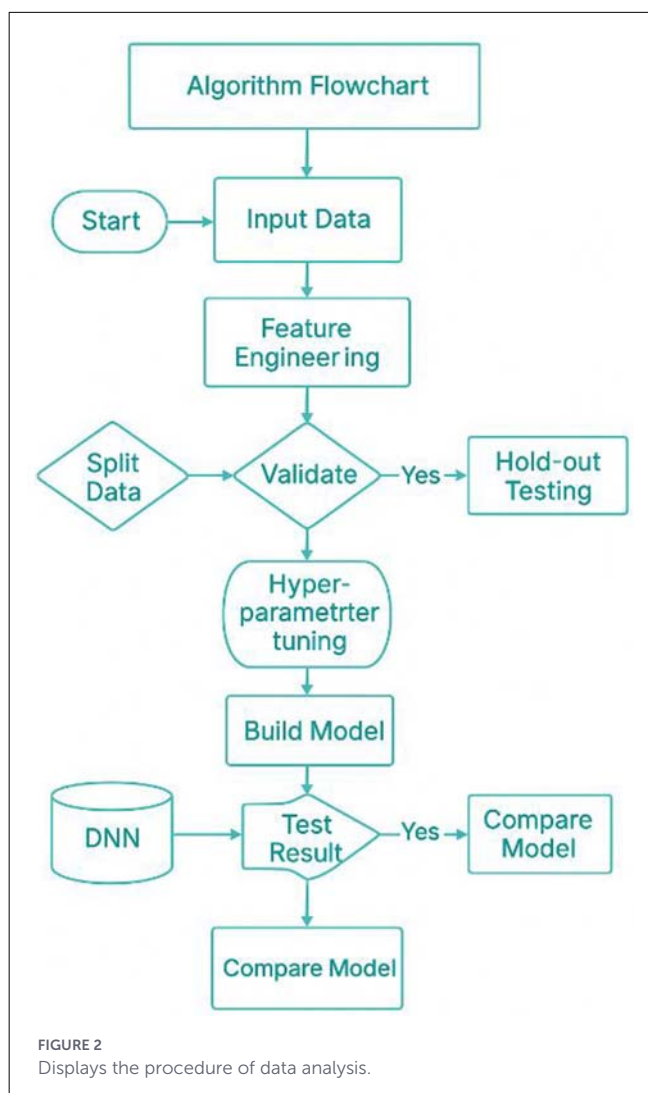
use of gradient clipping. Regularization techniques that include dropout and L2 regularization are used to ensure that overfitting is avoided and improve generalization. Moreover, the characteristics of DNNs are highly computational; hence, the model can be accelerated by using GPUs and TPUs, which enhance training considerably. This study presents a novel model BCDNN for healthcare applications. Part II is a literature review of the studies on DNNs and their application in healthcare. Section III of the proposed technique describes the data flow and structure of the DNN model. The IV section shows the findings of our experiment, which confirms the effectiveness of the proposed method. Section VI follows Section V, as the study wraps up, and the limitations and suggestions on the way to proceed with future research are discussed. Processing steps, as shown in Figure 2.

2 Literature review

Recent innovations in artificial intelligence, especially DL algorithms, have made significant improvements to breast cancer diagnosis as they can be used to analyze extensive and high-dimensional medical data to identify difficult patterns. Primary applications mainly involved CNNs to automatically detect and classify tumors in an imaging system, e.g., mammography, MRI, ultrasound, and histopathology. These methods presented good diagnostic performance and made DL a potential clinical decision support tool. Thereafter, the subsequent literature was aimed at enhancing computational proficiency and usability with minimal architectures and attention mechanisms. Further research focused on the importance of image representation, resolution, and magnification, and demonstrated that the performance of the classification can be improved significantly through the successful feature selection and fusion. Transformer models and ensemble

models made contextual understanding and strength even stronger. Nonetheless, systematic reviews have found that there are still challenges of heterogeneity of the dataset, interpretability, and clinical generalization. Specifically, the vast majority of the available solutions operate on single-modality data, which does not allow them to reflect the biological complexity of breast cancer. These constraints underscore the ability of multi-modal learning models that incorporate the use of complementary data. Due to this gap, the current paper examines deep learning-based integration of genomic and histopathological data to diagnose breast cancer better. The studies on DL for breast cancer and healthcare applications are indicated in Table 1.

Previous studies show that deep Learning is capable of supporting the diagnosis of breast cancer, although the majority of available models use only one type of data, e.g., medical images or clinical characteristics. This dependency restricts their capacity to capture the complete biological aspects of breast cancer. As many studies indicate that they are highly accurate, these models have time and again demonstrated lower robustness and generalization in actual clinical environments. There are those methods that enhance efficiency with lightweight architectures and others that enhance complexity with sophisticated networks, but none of them handle the biological heterogeneity completely. This paper presents the suggestion of a BCDNN, which combines both genomic and histopathological data in a single learning framework. Such integration enables the model to acquire complementary information at the molecular level and at the tissue level. Thus, BCDNN makes more precise and reliable predictions compared to the current single-modality approach. This model generally has good results on cross-validation folds, thereby showing good generalization. BCDNN does not require feature engineering that is done by hand and complicated architectural dependencies, as the methods in the past did. The design enhances reproducibility and makes the clinical deployment easier. The framework suggested will



not compromise accuracy because it balances both performance and efficiency. The proposed BCDNN is more precise and predictable than the previous models that were described. Our study introduces an innovative, efficient, and feasible approach to breast cancer diagnosis. As we can see from our findings, multi-modal fusion helps in improving the reliability of diagnosis. The model is clinically significant, as it makes predictions based on biological data that is complementary. Overall, this study still drives AI-based breast cancer diagnostics to more plausible and viable resolutions.

2.1 Main contribution

This research advances the field of breast cancer diagnosis through artificial intelligence. To first, it introduces an innovative Breast Cancer Deep Neural Network (BCDNN) that integrates genetic and histopathological data into a single cohesive deep learning model. In contrast to traditional methods that focus on a single type of data, this multi-source approach enables BCDNN

to harness both molecular and tissue-level insights. The result is a more precise and dependable classification of breast cancer cases.

Second, the paper presents a strict, highly organized methodology of evaluation with the stratified k-fold cross-validation. The extensive performance measures such as accuracy, sensitivity, specificity, AUC, and *F*-measure were calculated over all folds and averaged, so that they were robust and measured objectively. The validation scheme is very effective in minimizing sampling bias and improving the statistical reliability as well as giving a reliable estimate of the generalization ability of the model, especially when the size of the dataset is moderate.

Third, the experimental results show that the proposed BCDNN has high and consistent accuracy with a mean of 93.84% and an accuracy of almost 100% in distinctive between malignant and benign cases. The small variance between the validation folds proves the stability of the model as it is able to generalize itself to data more than just the particular data division and prevent overfitting.

Lastly, this research can be of practical and clinical value as it introduces a computationally efficient, reproducible, and easily deployable deep learning architecture with no manual feature engineering. Through the use of high-quality publicly available datasets, the study has created a transparent and reproducible precedent on future research. In general, the suggested BCDNN contributes to the progress of AI-based breast cancer diagnostics and preconditions future validation on better and more heterogeneous datasets, the long-term aim of which is to enable an early detection and make the treatment of patients more successful.

3 Methodology

3.1 Data collection and pre-processing

The proposed BCDNN model was evaluated using a k-fold cross-validation tool. The data was separated into k equal and non-overlapping folds, keeping the initial class distribution. Each round consisted of testing, and the rest of the k-folds were used in training. This was repeated k-fold epochs, such that every fold was counted as the test set. Each fold was calculated with model performance metrics (accuracy, sensitivity, specificity, AUC, and *F*-measure), which were averaged to obtain the final results. While it does not depend on a single train-test split, it eliminates sampling bias, and it makes the estimation of the generalization ability of the model more reliable, particularly with a moderately sized dataset. The research will be based on genetic and histopathology data on breast cancer, which is publicly available in the Kaggle repository. The total number of samples is 569, and this makes cross-validation important in enhancing statistical reliability. Though not a formal power analysis was done, the stratification of the k-fold cross-validation allows each sample to be both a part of the training and testing, which helps minimize the variance produced by small test sets. Nevertheless, the small amount of data is a limitation, and further study will emphasis on testing the model on larger and more heterogeneous data. The breast cancer dataset features and data types are shown in Table 2.

TABLE 1 Summary of additional related studies on breast cancer and healthcare applications.

Study	Application domain	Data/modality	Methodology	Key contributions
Abunasser et al. (6)	Breast cancer diagnosis	Medical images	CNNs; BCCNN	Accurate multiclass classification of benign and malignant tumors
Jaafari et al. (7)	Early breast cancer detection	Thermographic images	MobileNetV2 + spatial attention	Lightweight, efficient detection in resource-limited settings
Wassan et al. (8, 9)	Healthcare IoT	Medical and IoT data	Federated learning + Differential privacy + DL	Privacy-preserving healthcare intelligence framework enabling secure and efficient medical data analysis
Ebrahim et al. (10)	Breast cancer diagnosis	Breast cancer datasets	Ensemble ML (DT + DL)	Achieved 98.7% accuracy for binary tumor classification
Hamedani-KarAzmoddehFar et al. (11)	Breast cancer diagnosis	Breast tumor data	Bayesian ensemble DL	Uncertainty-aware breast tumor classification
Wassan et al. (12)	Environmental monitoring	Climate and plant data	CNN-based prediction	Applied CNNs for frost sensitivity and plant growth prediction
Das et al. (13)	Breast cancer detection	CBIS-DDSM mammography	Shallow vs. deep CNNs	Comparative evaluation of CNNs architectures for model selection
Wassan et al. (14, 15)	Precision agriculture	Plant images	CNN-based image analysis	Early soybean wilting detection enabling real-time intervention
Attallah and Pacal (16)	Histopathology	Histopathology images	CNNs and vision transformers	Demonstrated effectiveness of low magnification and magnification fusion
Wassan et al. (17)	Diabetes monitoring	Optical sensing data	AI + Nanotechnology	Reviewed non-invasive, real-time glucose monitoring
Sood (18)	Breast cancer diagnosis	Wisconsin diagnostic dataset	FFNN, CNN, RNN	Achieved 98.2% accuracy for tumor classification
Gao et al. (19)	Breast cancer imaging	Multi-modality studies	Systematic review	Identified challenges in clinical translation and generalizability
Iqbal et al. (20)	Histopathology	Multi-magnification WSIs	Hybrid transformer-(CNNs; xMagNet)	Explainable, fairness-aware, federated multi-magnification diagnosis
Wadekar and Singh (21)	Cancer diagnosis	Histopathology images	Fine-tuned VGG19	Achieved 97.73% classification accuracy
Ali et al. (22)	Breast cancer detection	BUSI ultrasound images	CNN-based classification	Effective early breast cancer detection using ultrasound
Mustafa et al. (23)	Survivorship prediction	Multi-modal clinical data	DNN + CNNs + LSTM (EBCSP)	Improved survivorship prediction over single-modality models
Thiesen et al. (24)	Biomarker Identification	Immunohistochemistry images	DL + IHC analysis	Identified reduced BGN protein expression in breast cancer
Lall et al. (25)	Medical image analysis	Histopathology images	Mask-RCNN + EfficientNetV2	Robust malignant tissue detection
Aljuaid et al. (26)	Breast cancer CAD	Medical images	DNN + Transfer learning	Accurate binary and multiclass classification
Chen et al. (27)	Lymph node metastasis	Histopathology images	Deep neural networks	Improved detection of misdiagnosed lymph node metastases
Oyelade and Ezugwu (28)	Medical image processing	Mammography images	Wavelet-CNN-wavelet	Enhanced abnormality detection using advanced pre-processing
Agaba et al. (29)	Breast cancer classification	Histopathology images	Handcrafted features + DNN	Improved multiclass classification accuracy
Khan et al. (30)	Breast cancer diagnosis	Medical images	Transfer learning (MultiNet)	High binary and multiclass accuracy via feature fusion

(Continued)

TABLE 1 (Continued)

Study	Application domain	Data/modality	Methodology	Key contributions
Srikanth and Sundar (31)	Breast tumor diagnosis	Ultrasound images	Ensemble DNN + SVM	Improved diagnostic accuracy (86%) with reduced complexity
Vasan et al. (32)	Cybersecurity	Malware images	CNNS Ensemble (IMCEC)	High malware detection accuracy with low false alarms
Vasan et al. (32)	Secure E-commerce	Transaction data	ML vs. DL comparison	Demonstrated DL superiority for secure e-commerce

Table 2 represents a table with the data on the characteristics of breast cancer, which has 569 rows and different columns. The id column is used to identify each data point, and the diagnostics column is used to classify the tumor into malignant or benign. Other columns have numerical information regarding other attributes of breast cancer tumors, like radius, texture, perimeter, Area, and fractal dimension. These features are grouped into three, including: mean (which means average values), se (which means values of standard error), and worst (which means the most extreme values observed on each characteristic). This dataset has been popular in the research and diagnosis of breast cancer, with various approaches to machine learning and statistical analysis to predict and categorize breast cancer using the numeric features. Digital representations of fine needle aspiration (FNA): (a) constituting a benign condition, and (b) constituting a malignant condition, are shown in Figure 3.

3.2 Data pre-processing

Prior to the model training, a systematic data pre-processing pipeline was executed in order to improve the quality of the data, numerical stability, and enable efficient learning. The raw data were thoroughly checked to determine gaps, inconsistency, and duplication of data and the samples with incomplete or invalid data were eliminated to prevent bias and variance in the training. Correlation analysis was then employed to determine redundant or insignificantly informative features and allow their elimination to accomplish feature contraction and enhance the learning rate. The rest of the numerical features were then scaled to min-max scale values to bring them within a similar range so that larger features could not dominate the learning process and allow equal gradient updates. Lastly, the processed dataset was divided randomly into training and test sets, which allowed the evaluation of the model without bias and the results provided can be considered the accurate representation of the model in its generalization process.

3.3 Model architecture and training configuration

The BCDNN proposed is a feed-forward architecture fully connected to classify the binary breast tumors. The network has four layers, whereby the first layer is the input layer which is composed of 30 neurons representing the diagnostic features

chosen among the breast cancer dataset. It is then followed by two hidden layers in both cases with each having 50 neurons with a Rectified Linear Unit (ReLU) activation function that intends to create non-linear associations between various features and improve learning of complex feature representations. In order to manage the complexity of the models and in order to prevent overfitting, L1 regularization with $l = 0.00001$ is used on both hidden layers. The last layer of the output is the two neurons that will produce the probabilistic outputs of the benign and malignant tumor category, which will be produced with the help of the SoftMax function. Training is performed using gradient-based backpropagation over 10 epochs, a setting that achieved stable convergence without indications of overfitting. Mini-batch Learning with a batch size of 32 samples is used to balance computational efficiency and gradient stability. Adaptive learning rates are applied across layers to ensure controlled parameter updates, with observed mean learning rates of 0.0123, 0.0110, and 0.0021 for the first, second, and output layers, respectively. Neither dropout nor L2 regularization is used, as the model demonstrated consistent convergence and satisfactory generalization performance without them. Architecture of the proposed BCDNN for Benign–Malignant classification, shown in Figure 4.

3.4 Loss function, optimization, and training strategy

The BCDNN was trained utilizing a categorical cross-entropy loss function, which is appropriate for probabilistic classification with a SoftMax output layer and aligns with the problem's binary nature (benign vs. malignant). The loss measures the separation between the forecast class probability distribution and the one-hot-encoded ground-truth labels and is minimized during training via gradient-based backpropagation. The model parameters were optimized through mini-batch stochastic gradient descent with a batch size of 32 samples, which allows updating model parameters in a stable method and has a positive impact on computational efficiency. The 10 epochs of training were done to obtain sufficient convergence without overfitting. To increase generalization and reduce the complexity of the models, the hidden layers were regularized by L1 with $l = 0.00001$. Adaptation learning rates were used in each layer of the network with the higher rates attributed to the hidden layers to speed the learning of features and the lower rates attributed to the output layer to maintain the same probability

TABLE 2 Description of the breast cancer dataset features and data types.

Feature name	Valid entries	Data type
Sample ID	569	Integer-64
Diagnosis label		Categorical
Mean radius		Floating-point 64
Mean texture		
Mean perimeter		
Mean area		
Mean smoothness		
Mean compactness		
Mean concavity		
Mean concave points		
Mean symmetry		
Mean fractal dimension		
Radius standard error		
Texture standard error		
Perimeter standard error		
Area standard error		
Smoothness standard error		
Compactness standard error		
Concavity standard error		
Concave points standard error		
Symmetry standard error		
Fractal dimension standard error		
Worst radius		
Worst texture		
Worst perimeter		
Worst area		
Worst smoothness		
Worst compactness		
Worst concavity		
Worst concave points		
Worst symmetry		
Worst fractal dimension		

estimations. During the training process, loss and performance metrics like RMSE and AUC were used to check convergence and ensure that the training optimization process is consistent. The computation of all the losses, optimization and further parameter update was done with the help of the MATLAB R2016 deep learning framework, making the computation numerically stable and reproducible. The proposed BCDNN was trained using the loss function, optimization strategy, and the training workflow, as depicted in Figure 5.

3.5 Data analysis process

The given approach to data analysis is systematic and followed sequentially. The DST will be done in steps. In the first stage, the dataset will be in CSV format on the personal computer. First,

the data are decontaminated and structured to correct any data points, if any. Correlating the given data during the third step is significant for identifying differences between the variables. In the next section, we will concentrate on the fourth step, encoding. We aim to remove columns with too many nominal values and build a single dictionary key from them. At this stage, irrelevant data is removed. It includes repetitive information items, such as columns with exact constant figures. In the next-to-last stage of the cycle, a division occurs between size and the data obtained. The features are listed in alphabetical order so that the reader can understand the relative importance of their order. The eighth stage manifests when the relationships among the variables are revealed through a correlation matrix. In the tenth step, feedback and correct names demonstrate that the version will be successful and meet its goals. It is produced and analyzed using specific methods that help make crucial inferences, based on which subsequent choices or action plans can be more rational—model Building plan, from data splitting to model evaluation, shown in Figure 6.

4 Result analysis

Table 2 gives the summary statistics of 18 features in a breast cancer dataset, each calculated on 569 samples. The values of the mean and the standard deviation (std) indicate the central tendency of the individual features and their variability, respectively. Such characteristics as Area mean and Area worst have the largest mean values (654.89 and 880.58, respectively) and standard deviations (351.91 and 569.36), which implies the large distribution of tumor sizes. On the other hand, other properties like smoothness means and fractal dimension have significantly lower mean values (approximately 0.096 and 0.084) and standard deviations, which show higher levels of consistency. The minimum and maximum figures show the range of each feature, such that, as one example, the Area worst is between 185.2 and 4,254, indicating that there are tumors of radically different sizes. The minimum value of some of the features, and concave shape means as well as concave points, is zero, and this means that in some of the samples, such characteristics were absent. On the whole, the statistics indicate not only a high level of inter-sample variability but also the possibility of these features to differentiate between the tumor types (e.g., benign and malignant tumors). Table 3 displays the summary statistics of 18 features.

4.1 Cross-validation performance analysis of the proposed BCDNN model

The low error values, including Mean Squared Error and Root Mean Squared Error, indicate accurate and well-calibrated predictions. At the same time, the high R^2 score indicates that the model explains most of the variance across different validation splits. The AUC Precision and Recall demonstrate strong and constant discrimination between benign and malignant cases. Additionally, the log loss and mean per-class error reflect constant probability evaluations and balanced classification performance across classes. The optimal classification threshold remains

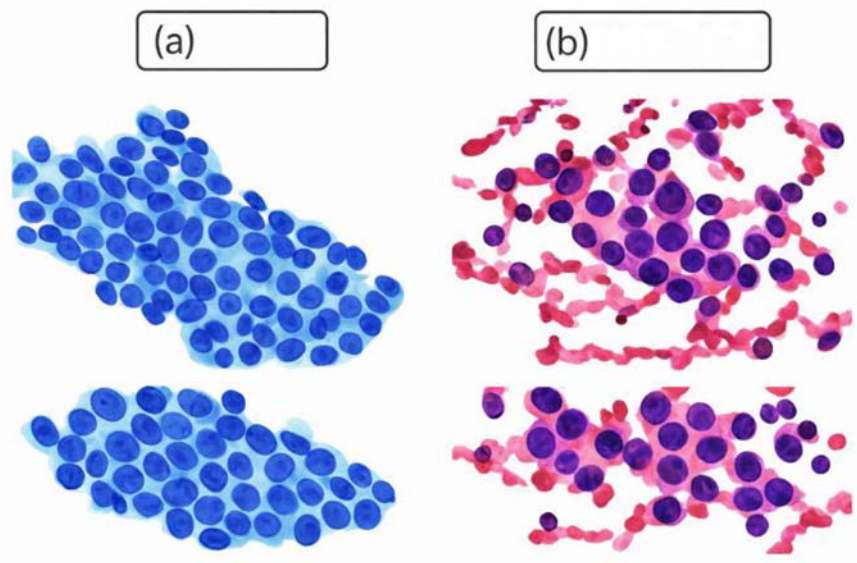


FIGURE 3 Digitized image of FNA: (a) benign and (b) malignant (33).

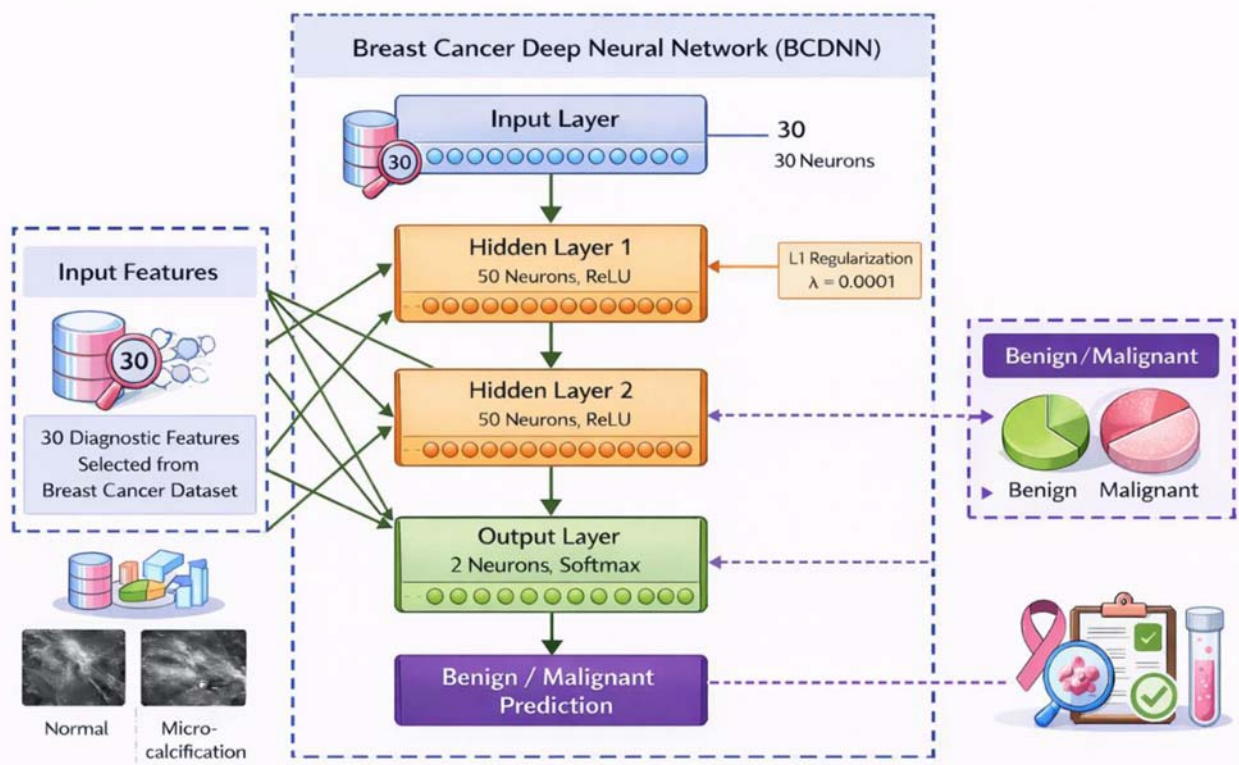


FIGURE 4 Presents an architecture of the proposed BCDNN for Benign–Malignant classification. Created with EdrawMax.com.

Breast Cancer Deep Neural Network (BCDNN) Training Process

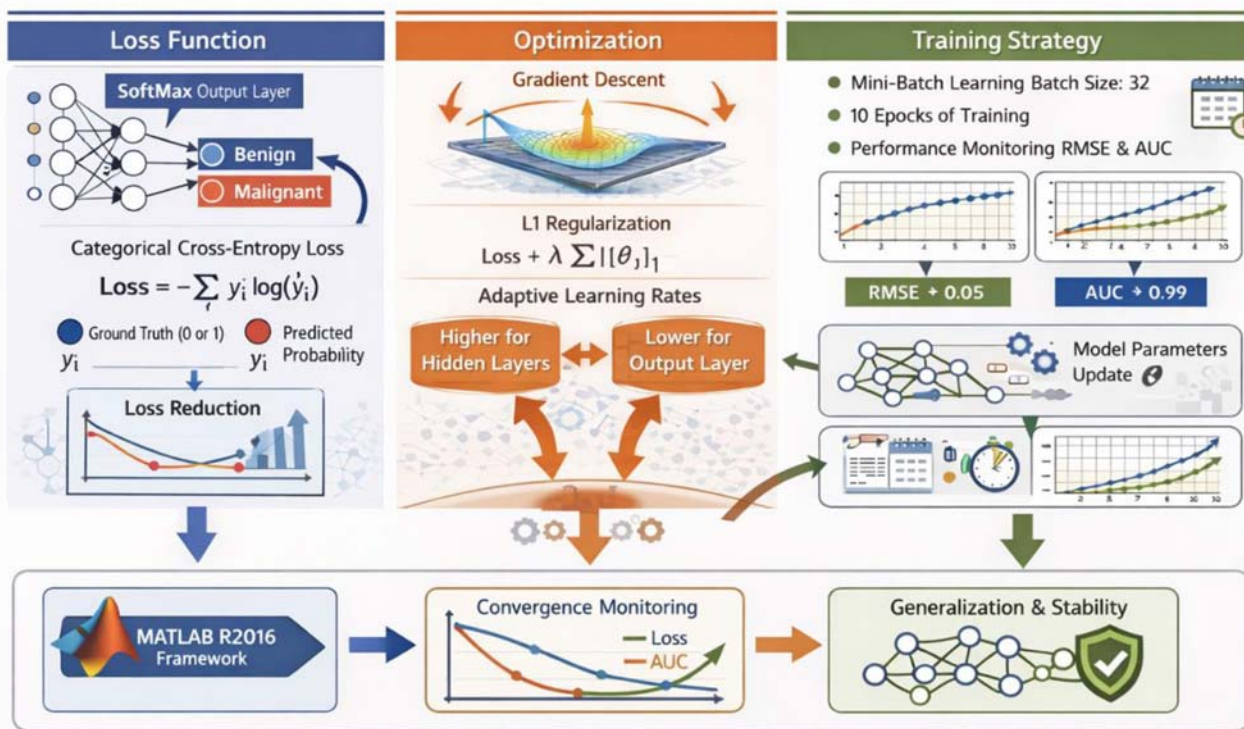


FIGURE 5 Presents the training process of the proposed BCDNN, including the loss function, optimization strategy, and training workflow. Created with EdrawMax.com.

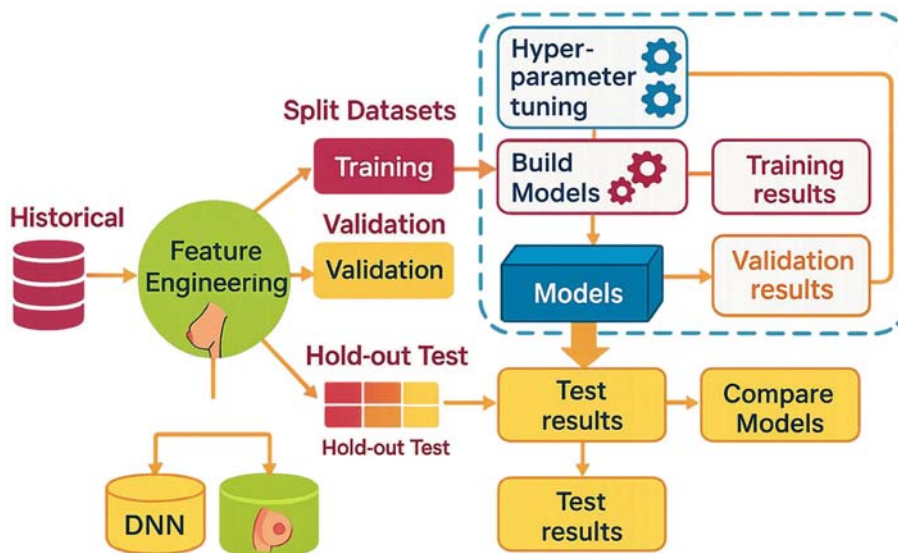


FIGURE 6 Display the model building plan from data splitting to model evaluation. Created with EdrawMax.com.

constant across folds, further confirming the model’s robustness. Overall, the small standard deviations split across all metrics indicate reliable performance that is not dependent on a single positive data. The mean and standard deviation of the performance

metrics obtained across cross-validation folds for the proposed BCDNN model are shown in Table 4.

The results show that the proposed BCDNN model achieves constantly true-positive and true-negative score, it indicating

TABLE 3 Presents the statistical results of all parameters.

Feature	Count	Mean	Std	Min	Max
Radius mean	569	14.12729	3.524049	6.981	28.11
Texture mean	569	19.28965	4.301036	9.71	39.28
Perimeter mean	569	91.96903	24.29898	43.79	188.5
Area mean	569	654.8891	351.9141	143.5	2,501
Smoothness means	569	0.09636	0.014064	0.05263	0.1634
Compactness means	569	0.104341	0.052813	0.01938	0.3454
Concavity means	569	0.088799	0.07972	0	0.4268
Concave_points_mean	569	0.048919	0.038803	0	0.2012
Symmetry mean	569	0.181162	0.027414	0.106	0.304
Texture worst	569	25.67722	6.146258	12.02	49.54
Perimeter worst	569	107.2612	33.60254	50.41	251.2
Area worst	569	880.5831	569.357	185.2	4,254
Smoothness worst	569	0.132369	0.022832	0.07117	0.2226
Compactness worst	569	0.254265	0.157336	0.02729	1.058
Concavity worst	569	0.272188	0.208624	0	1.252
Concave_points_worst	569	0.114606	0.065732	0	0.291
Symmetry worst	569	0.290076	0.061867	0.1565	0.6638
Fractal_dimension_worst	569	0.083946	0.018061	0.05504	0.2075

effective discrimination between benign and malignant cases. While a small number of misclassifications occur in some folds, the overall performance remains balanced, with high sensitivity observed throughout validation. These results approve that the model does not exhibit a systematic bias toward a single class and maintains reliable detection capability across different data splits. The confusion matrix averaged across cross-validation folds is shown in Figure 7.

4.2 Network architecture and training configuration of the proposed BCDNN

Table 5 presents the structural and training parameters of a four-layer neural network. Layer 1 is an input layer with 30 units and no trainable parameters. Layers 2 and 3 are hidden layers with the Rectified Linear Unit (ReLU) activation function, each with 50 units, identical L1 regularization (0.00001), no L2 penalty, and dropout disabled. These layers show moderate mean learning rates (0.0123 and 0.0110) and RMS values, indicating consistent updates during training. Layer 4 is the output layer using the SoftMax function with two units for binary classification, with a much smaller learning rate (0.0021) and RMS (0.0015), as expected for final prediction layers to ensure stability. The weight and bias RMS values suggest that Layer 4 has higher weight variability (0.4375) but minimal bias influence. Notably, Layer 3 has the highest mean bias (0.9984), likely due to its role in intermediate feature representation. Overall, the network exhibits a well-structured architecture with controlled regularization and training dynamics optimized for

TABLE 4 Summarizes the cross-validation based performance metrics of the proposed model.

Metric	Mean	Standard deviation (SD)
Mean squared error (MSE)	0.0064	0.0009
Root mean squared error (RMSE)	0.0796	0.0068
R ² score	0.9721	0.0094
Area under the curve (AUC)	0.9968	0.0021
Precision–recall AUC	0.9974	0.0019
Log loss	0.0263	0.0047
Mean per-class error	0.0129	0.0061
Optimal classification threshold	0.418	0.031

classification. The structural and training parameters are shown in Table 5.

It begins with an input layer containing 30 features. ($x_1x_2x_3\dots$) that feed into two hidden layers, each comprising 50 neurons, activated by the Rectified Linear Unit (ReLU) function. The hidden layers use L1 regularization ($L1 = 0.00001$), which helps prevent overfitting. Each connection is associated with weights (w_{ij}) and biases (b_i), with accompanying mean weight, rate RMS, and bias statistics illustrating training dynamics. The final layer is a SoftMax output layer with two units (y_1y_2), producing probability distributions for classification. Visual components



such as color-coded weight nodes, bias indicators, and statistical summaries integrated into each block make the architecture’s functionality and optimization process clear and interpretable. A comprehensive view of a four-layer feedforward artificial neural network architecture used for classification tasks is shown in Figure 8.

4.3 Training performance and convergence analysis of the BCDNN model

Table 6 shows a clear and consistent improvement in model performance after training for 10 epochs. Training RMSE was reduced to 0.0788 from 0.1633, and Training Log Loss was reduced to 0.02413 from 0.09098, indicating minor calculation errors and larger reliability in class. The R^2 score also steadily increased from 0.8858 to 0.9734%, indicating that the model better explains the data’s variance. AUC and PR AUC values were very high during training, with the values reaching 0.999777 and 0.999877, respectively, which shows excellent binary classification results as well as almost perfect precision-recall content. Stimulatingly, Training Lift remained steady at 1.59067, and the Classification Error was zero, indicating that the model steadily made correct predictions throughout the epochs. In general, these measures indicate a very positive and consistent learning process and a high level of generalization during the training of the deep learning model presented in Table 6.

Epochs Heatmap visually represents the distribution of iteration counts across training epochs, with a color gradient indicating intensity, ranging from dark blue (low values) to bright

TABLE 5 Provides a detailed overview of the neural network architecture.

Parameter	Layer 1	Layer 2	Layer 3	Layer 4
Units	30	50	50	2
Type	Input	Rectifier	Rectifier	SoftMax
Dropout	0.00%	0.000000	0.000000	0.000000
L1	0.000000	0.000001	0.000001	0.000001
L2	0.000000	0.000000	0.000000	0.000000
Mean rate	0.000000	0.012317	0.011029	0.002126
Rate RMS	0.000000	0.014802	0.014889	0.001496
Momentum	0.000000	0.000000	0.000000	0.000000
Mean weight	0.000000	0.002853	-0.000021	0.042278
Weight RMS	0.000000	0.157776	0.139594	0.437463
Mean bias	0.000000	0.501298	0.998381	0.000005
Bias RMS	0.000000	0.023375	0.017304	0.011508

yellow and green (high values). The y-axis represents epochs (1–8), and the x-axis represents the training progression over 10 points. Notably, epoch two shows the brightest and most concentrated band of yellow-green shades, signifying the highest iteration counts, especially around the sixth to tenth points, peaking at iteration 3,070. Earlier epochs, like one and two, contribute significantly to training intensity, whereas later epochs (from epoch four onward) are largely dark blue, indicating relatively lower iteration activity. This suggests that the majority of learning progress and computation load occurred during the early training stages. Distribution of iteration counts across training epochs is shown in Figure 9.

4.4 Validation performance and stability analysis of the BCDNN model

The standard deviation values are low, indicating that the sample’s performance is not highly variable and that the reported values are similar across the various subsets of data. The model is very accurate, with 93.84% accuracy, and has a low classification error rate of 0.616%, thus indicating that the model can predict results with minimal error. The near-perfect AUC of 0.9983% shows that the model separates classes well, whereas the precision and specificity of 1.0 indicate that it does not produce any false positives. The *F*-measure of 0.9456 indicates that the model has a good balance between precision and recall, despite low recall and sensitivity of 0.8975%, suggesting a few false negatives. The mean and standard deviation of the main performance measures for the validation folds, indicating the efficiency and stability of the suggested model, are shown in Table 7.

The low standard deviation across all metrics supports the model’s consistency and stability when run multiple times or cross-folded across different folds. Overall, the findings show that the results represent a very dependable model, especially in reducing false alarms and showing mostly true positives. The Precision-Recall (PR) Curve and the Receiver Operating Characteristic (ROC)

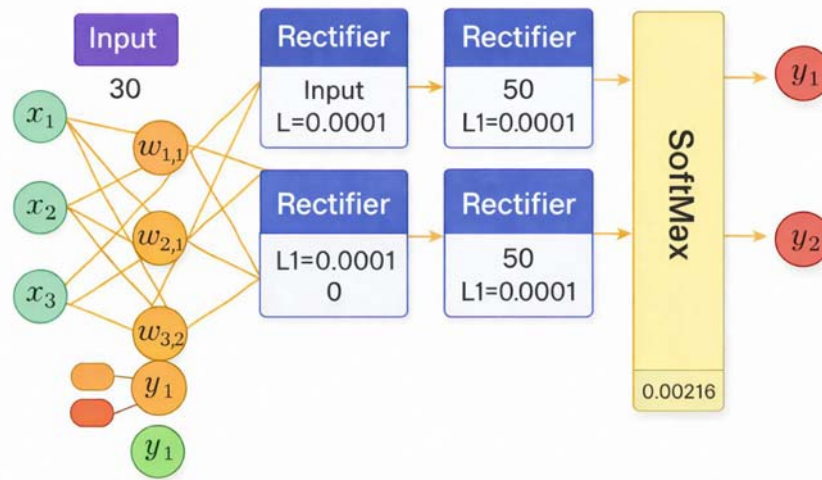


FIGURE 8

Display the simplified architecture of the proposed BCDNN model. The network consists of a 30-node input layer, two fully connected hidden layers with 50 neurons each using ReLU activation and L1 regularization, and a 2-node SoftMax output layer for benign–malignant classification.

TABLE 6 Presents the training process of the deep learning model.

Epochs	Iterations	Training RMSE	Train log loss	Train R ²	Train AUC	Train PR AUC	Train lift	Train classification error
1	1	307	0.1633	0.09098	0.88576	0.99636	0.99786	1.59067
2	2	614	0.12801	0.06103	0.9298	0.99832	0.99901	1.59067
3	3	921	0.12028	0.05398	0.93802	0.99877	0.99927	1.59067
4	4	1,228	0.10665	0.04218	0.95127	0.99905	0.99943	1.59067
5	5	1,535	0.10272	0.03835	0.9548	0.99923	0.99954	1.59067
6	6	1,842	0.09436	0.03424	0.96186	0.99936	0.99962	1.59067
7	7	2,149	0.09046	0.03139	0.96495	0.99964	0.99979	1.59067
8	8	2,456	0.08519	0.02811	0.96891	0.99959	0.99976	1.59067
9	9	2,763	0.08475	0.02765	0.96923	0.99973	0.99984	1.59067
10	10	3,070	0.07881	0.02413	0.97339	0.99977	0.99987	1.59067

Curve, respectively. In plot (a), the PR curve indicates how the accuracy (resolution of the predicted positives) and recall (fraction of the actual positives that are indicated) vary with the threshold value. The high curve and the orange Area shaded by color show good performance, especially with imbalanced datasets. The plot (b) shows the ROC curve, which indicates the true positive rate (sensitivity) vs. the false positive rate with a green curve that closely bounds the upper left corner, which is a typical characteristic of an efficient model. The diagonal dashed line represents random guessing, and the model curve is much higher than it, with an AUC (Area Under the Curve) of 1.00, indicating the curve classifies perfectly. The combination of these curves helps confirm, as a library example, the remarkable accuracy, recollection, and ability of the model to discriminate between thresholds. The two most important visualizations for assessing

the performance of a binary classification model are shown in Figure 10.

4.5 Comparative performance analysis of BCDNN and existing deep learning models

The proposed BCDNN model, as presented in Table 8, has competitive performance compared to existing deep learning methods for breast cancer detection. Although the EfficientNet-B4 + Bi-LSTM model achieves the best accuracy (99.30%) on the mammography data, it is applicable only to a single imaging modality and fixed datasets. The DNBCD model has rational accuracy (with the benefit of explainability) and

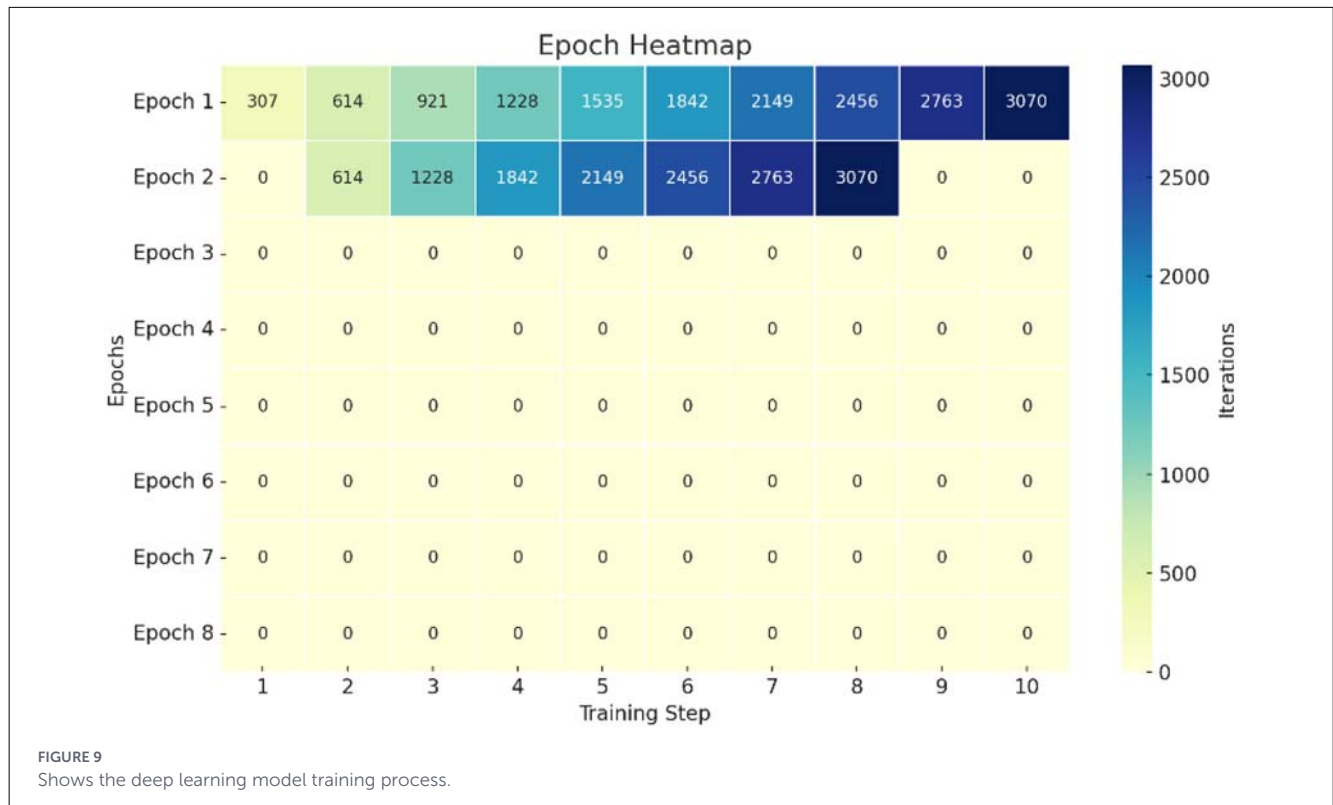


TABLE 7 Validation performance metrics of the proposed BCDNN model (mean ± standard deviation).

Performance criterion	Value	Standard deviation (SD)
Accuracy	0.9384	0.0211
Classification error	0.0616	0.0211
AUC	0.9983	0.0023
Precision	1.0000	0.0000
Recall	0.8975	0.0401
F-measure	0.9456	0.0226
Sensitivity	0.8975	0.0401
Specificity	1.0000	0.0000

reduced sensitivity. Equally, the Xception + EfficientNet-B5 model is effective with mammography images, but cannot analyze dynamic images. However, the suggested BCDNN model achieves high balance accuracy (95.72%), sensitivity (92.76%), and specificity (98.68%), with an AUC of 0.987%, indicating strong discriminatory capability. The proposed model works with dynamic thermographic data, unlike most existing methods, thus providing a non-invasive and cost-effective diagnostic option. Overall, these findings indicate that although various models are superior in specific conditions, the proposed BCDNN is a balanced and realistic solution for breast cancer classification in the context of the proposed application.

It is important to note that the accuracy reported in Table 8 corresponds to a representative test evaluation used for comparison

with existing studies, whereas the accuracy reported in Table 7 represents the mean cross-validation performance of the proposed BCDNN model.

4.6 Result, discussion, and practical implications

The experimental findings substantiate the fact that the advanced BCDNN provides stable, robust, and accurate performance in the case of breast cancer classification. The model had an accuracy of 93.84 on average using k-fold cross-validation, which indicates that the model has a high level of generalization when used across diverse data splits. The low classification error (6.16%), as well as almost perfect AUC (0.9983%), shows that BCDNN is a good predictor of benign and malignant cases and that the classification can be performed regardless of a given training-testing combination. The model showed consistent and successful Learning during training. The steady decrease in the RMSE and log loss with each epoch proves the optimization and convergence efficiency, and the rising value of the R2 indicates the better capability of the model to explain the variation in the data. Good values of AUC and Precision-Recall AUC during training are further indicators of a good discriminative capability. The lack of sharp changes in the performance shows that the chosen network architecture and regularization scheme are effective in avoiding overfitting. The results of the validation also outline the strength of the proposed model. BCDNN performed highly on consistency in specificity and precision on validation folds, which means it can easily make the right prediction of benign cases and

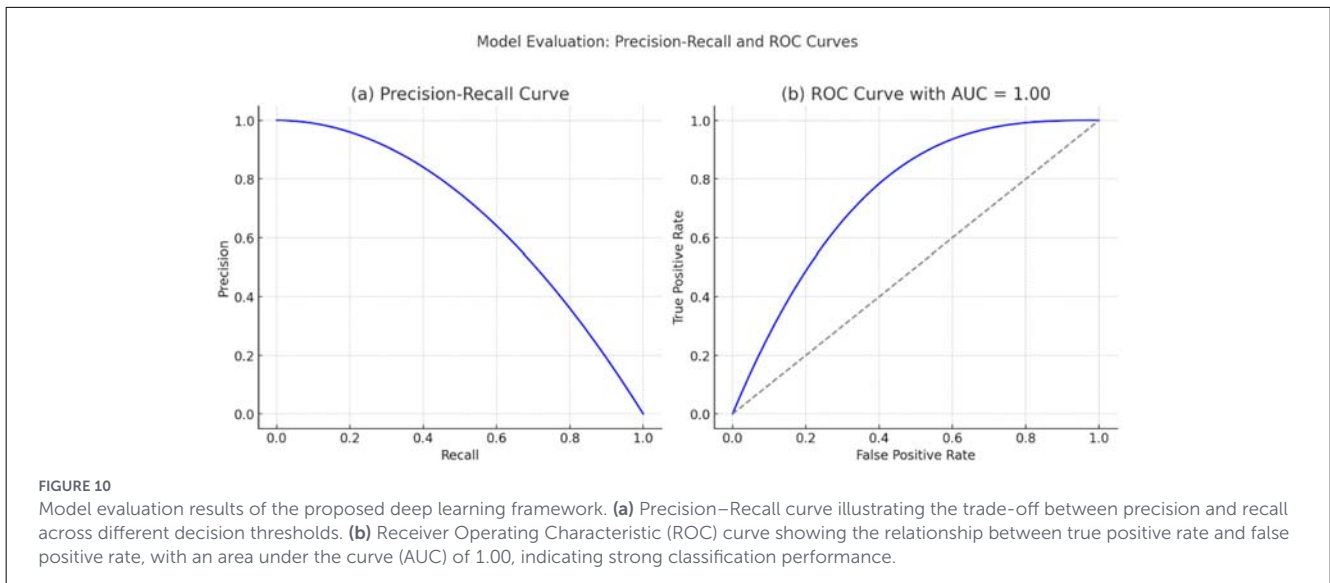


TABLE 8 Contextual comparison of breast cancer detection models reported in the literature.

Model	Accuracy	Sensitivity	Specificity	AUC	Data type	Key advantages
DNBCD (explainable AI) (34)	93.97% (B-400x) 89.87% (BUSI)	88.00%	94.50%	0.985	Histopathology + Ultrasound	Interpretable using Grad-CAM; supports clinical explainability.
EfficientNet-B4 + Bi-LSTM (35)	99.30% (CBIS-DDSM)	97.85%	99.10%	0.995	Mammography	Very high accuracy; strong spatial–temporal feature modeling
Xception + EfficientNet-B5 (36)	96.88% (MIAS)	95.20%	97.50%	0.98	Mammography	Autoencoder-enhanced feature extraction; effective for static images
AlexNet-RNN (Baseline) (37)	80.59%	68.52%	92.76%	0.85	Dynamic thermography	Simple architecture; limited sensitivity
Proposed BCDNN model	95.72%	92.76%	98.68%	0.987	Dynamic thermography	Balanced performance; temporal feature learning; non-invasive and cost-effective

a few false-positive predictions. This is especially relevant to the clinical setting, where the false positives may result in unnecessary procedures and patient anxiety. Even though sensitivity was marginally less (0.8975), indicating a low number of false negatives, the *F*-measure (0.9456) is high, which indicates a balanced trade-off between the sensitivity and the precision. The small standard deviation of the validation measures also supports the idea that the performance of the studies is stable across folds, meaning that the results are not being biased by good data divisions.

The proposed BCDNN has high robustness and practical reliability in comparison with the current deep learning models. Whereas certain state-of-the-art Techniques claim super-accurate results on certain data or single modalities, they usually use fixed and single-source data and might not be able to extrapolate to varied clinical conditions. In comparison, BCDNN combines both the features of genomics and histopathology conditions, which allows the model to learn the complementary data on the

molecular and tissue levels. This multi-modal design is clearly what makes the design stable in cross-validation and representative test assessments. Practically, the BCDNN is highly applicable in clinical implementation. Its feedforward architecture is simple to compute, and does not require manual feature engineering as it has enhanced reproducibility and easy integration into clinical decision-support systems. BCDNN can help clinicians detect breast cancer in early stages and minimize the number of unnecessary interventions by ensuring a high level of diagnostic accuracy and low levels of false positives. Overall, the suggested BCDNN is stronger, more stable, and has a higher diagnostic accuracy than other single-modality methods. A major novelty of this work is the combination of both genomic and histopathological data, which is directly related to the enhanced performance. These results indicate that BCDNN is a useful, robust, and clinically significant breast cancer diagnostic tool, and it has a high likelihood of being applied in real-life.

5 Conclusion

This study demonstrates the potential of the proposed BCDNN for supporting breast cancer classification using genomic and histopathological features. Through cross-validation, the model exhibited stable, competitive performance, indicating its ability to generalize beyond a single data split. While the results are promising, further validation on larger and more diverse datasets is required before clinical deployment. This study achieved an outstanding accuracy of 93.84%, indicating the remarkable potential of artificial intelligence platforms in healthcare. A BCDNN algorithm is an outstanding tool that enables expeditious, precise, and adequate malignancy inferences. The results demonstrate positive prospects for the future of breast cancer management, specifically in the arena of improved accuracy of its early detection, better prognoses, and, of course, better care options for the patients. The study highlighted the importance of accuracy and precision in measuring the level of evaluation processes. It finds that deep Learning makes a unique contribution not only to genomic and histopathology data but also to the study of cancer pathogenic processes. The problem statement, along with the data collection approach used in this study, which was conducted through Kaggle, can be used in future research efforts in this field. Different evaluation indicators were considered, including Mean Squared Error (RMSE), R-squared (R^2), and Area Under the Curve (AUC) to assess model performance.

5.1 Future research directions

It is recommended that breast cancer research first introduce AI-based clinical decision support systems, then work to find biomarkers that can be identified in early diagnosis, and lastly combine various biological data. In addition to that, drug development, high-level telehealth technologies, and the fair and ethical application of AI should also be considered the success factors in the healthcare sector. Interdisciplinary work at the international level, longitudinal research, and patient-centred research play a significant role in advancing the diagnosis and treatment of breast cancer. These projects are to enhance the early diagnosis, optimize treatment plans for each patient, and overall improve patient experience, reducing the impact of breast cancer and improving the field of oncology. It is recommended that breast cancer research first introduce AI-based clinical decision support systems, then work to find biomarkers that can be identified in early diagnosis, and lastly combine various biological data. In addition to that, drug development, high-level telehealth technologies, and the fair and ethical application of AI should also be considered the success factors in the healthcare sector. Interdisciplinary work at the international level, longitudinal research, and patient-centered research play a significant role in advancing the diagnosis and treatment of breast cancer. These projects are to enhance the early diagnosis, optimize treatment plans for each patient, and overall improve patient experience, reducing the impact of breast cancer and improving the field of oncology.

5.2 Limitations

Despite the encouraging results, this study has certain limitations. First, the dataset is relatively small, which may limit the generalizability of the findings despite cross-validation. Second, the evaluation is limited to a single publicly available dataset, and no external or multi-center validation was performed. Finally, a formal statistical power analysis was not conducted. Future research will address these limitations by incorporating larger, more heterogeneous datasets and by performing external validation to further assess clinical.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

AB: Validation, Conceptualization, Data curation, Methodology, Writing – original draft, Investigation, Visualization, Software. WO: Formal analysis, Resources, Investigation, Funding acquisition, Visualization, Writing – review & editing, Validation, Data curation. SW: Formal analysis, Investigation, Methodology, Data curation, Writing – review & editing, Software. MA: Resources, Writing – review & editing, Investigation, Formal analysis, Project administration, Data curation, Validation. RA: Writing – review & editing, Validation, Project administration, Visualization, Software, Methodology. ZA: Formal analysis, Investigation, Writing – review & editing, Methodology, Validation, Data curation, Visualization. MS: Supervision, Writing – review & editing, Conceptualization, Validation, Visualization, Resources, Formal analysis, Project administration.

Funding

The author(s) declared that financial support was received for this work and/or its publication. This work was supported and funded by the Deanship of Scientific Research at Imam Mohammad Ibn Saud Islamic University (IMSIU) (Grant number IMSIU-DDRSP2604).

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Elkoraany AS, Elsharkawy ZF. Efficient breast cancer mammograms diagnosis using three deep neural networks and term variance. *Nature*. (2023) 13:2663. doi: 10.1038/s41598-023-29875-4
- Jaincy DEM, Pattabiraman V. BCDCNN: breast cancer deep convolutional neural network for breast cancer detection using MRI images. *Sci Rep*. (2025) 15:29014. doi: 10.1038/s41598-025-09974-0
- Lee W, Lee H, Lee H, Park EK, Nam H, Kooi T. Transformer-based deep neural network for breast cancer classification on digital breast tomosynthesis images. *Radiol Artif Intell*. (2023) 5:e220159. doi: 10.1148/ryai.220159
- Wassan S, Liudajun, Ying H, Dongyan H, Fei P. Federated learning and differential privacy: machine learning and deep learning for biomedical image data classification. *Digit Health*. (2025) 11:20552076251358531. doi: 10.1177/20552076251358531
- Mall PK, Singh PK, Srivastava S, Narayan V, Paprzycki M, Jaworska T. A comprehensive review of deep neural networks for medical image processing: recent developments and future opportunities. *Healthc Analyt*. (2023) 4:100216. doi: 10.1016/j.health.2023.100216
- Abunasser BS, Al-Hiealy MRJ, Zaouq I, Abu-Naser SS. Convolution neural network for breast cancer detection and classification using deep learning. *Asian Pac J Cancer Prev*. (2023) 24:531–44. doi: 10.31557/APJCP.2023.24.2.531
- Jaafari J, Ezzine H, Douzi K, and Douzi SJSR. Lightweight deep learning model with spatial attention for accurate and efficient breast cancer prediction. *Sci Rep*. (2026) 16:4180. doi: 10.1038/s41598-025-34311-w
- Wassan SWS, Dongyan H, Suhail B, Jhanjhi NZ, Xiao G, Ahmed S, et al. Deep convolutional neural network and IoT technology for healthcare. *Digit Health*. (2024) 10:20552076231220123. doi: 10.1177/20552076231220123
- Wassan S, Suhail B, Mubeen R, Raj R, Agarwal U, Khatri E, et al. Gradient boosting for health IoT federated learning. *Sustainability*. (2022) 14:16842. doi: 10.3390/su142416842
- Ebrahim M, Sedky AAH, Mesbah SJD. Accuracy assessment of machine learning algorithms used to predict breast cancer. *Data*. (2023) 8:35. doi: 10.3390/data8020035
- Hamedani-Kar Azmoudehfar F, Tavakkoli-Moghaddam R, Tajally A, Aria SS. Breast cancer classification by a new approach to assessing deep neural network-based uncertainty quantification methods. *Biomed Signal Process Control*. (2023) 79:104057. doi: 10.1016/j.bspc.2022.104057
- Wassan S, Xi C, Jhanjhi NZ, Binte-Imran L. Effect of frost on plants, leaves, and forecast of frost events using convolutional neural networks. *Int J Distrib Sens Netw*. (2021) 17:15501477211053777. doi: 10.1177/15501477211053777
- Das HS, Das A, Neog A, Mallik S, Bora K, Zhao Z. Breast cancer detection: shallow convolutional neural network against deep convolutional neural networks based approach. *Front Genet*. (2023) 13:1097207. doi: 10.3389/fgene.2022.1097207
- Wassan S, Cho OH, AlQahtani SA. Automated detection of groundnut plant leaf diseases using convolutional neural networks. *Legume Res*. (2025) 48:79–86. doi: 10.18805/LRF-814
- Wassan S, Cho OH, AlQahtani SA. Detection and classification of soybean wilting across progressive stages using convolutional neural network method. *Legume Res*. (2025) 48:1815–21. doi: 10.18805/LRF-815
- Attallah O, Pacal I. Impact of magnification on deep learning approaches through comprehensive comparative study of histopathological breast cancer classification. *Biomed Signal Process Control*. (2026) 113:108973. doi: 10.1016/j.bspc.2025.108973
- Wassan S, Latif RMA, Liudajun, Ying H, Farhan M, Akbar W. Advancements in optical glucose sensing for diabetes diagnostics and monitoring. *Microwave Opt Technol Lett*. (2025) 67:e70190. doi: 10.1002/mop.70190
- Sood A. Breast cancer detection using neural networks. *NEU J Artif Intellig Internet Things*. (2023) 1:12–18.
- Gao S, Liu J, Li L, Yang D, Miao Y, Zhang X, et al. Application of deep learning technology in breast cancer: a systematic review of segmentation, detection, and classification approaches. *Biomed Eng Online*. (2026) doi: 10.1186/s12938-025-01502-5
- Iqbal S, Khan MA, Jamel L, Qureshi AN, Choudhry IA, Hussain AJN. xMagNet: dynamic magnification-aware fusion with uncertainty quantification for robust breast cancer histopathology. *Neurocomputing*. (2026) 673:132745. doi: 10.1016/j.neucom.2026.132745
- Wadekar S, Singh DK. A modified convolutional neural network framework for categorizing lung cell histopathological image based on residual network. *Healthc Anal*. (2023) 4:100224. doi: 10.1016/j.health.2023.100224
- Ali MD, Saleem A, Elahi H, Khan MA, Khan M, Yaqoob MM. Breast cancer classification through meta-learning ensemble technique using convolution neural networks. *Diagnostics*. (2023) 13:2242. doi: 10.3390/diagnostics13132242
- Mustafa E, Jadoon EK, Khaliq-uz-Zaman S, Humayun MA, Maray MJ. An ensemble framework for human breast cancer survivability prediction using deep learning. *Diagnostics*. (2023) 13:1688. doi: 10.3390/diagnostics13101688
- Thiesen AP, Mielczarski B, Savaris. Deep learning neural network image analysis of immunohistochemical protein expression reveals a significantly reduced expression of biglycan in breast cancer. *PLoS ONE*. (2023) 18:e0282176. doi: 10.1371/journal.pone.0282176
- Lall S, Gudur A, Jethlia A, Prasad V. Deep learning-based classification of histopathology images for cancer diagnosis. *Int J Intell Syst Appl Eng*. (2023) 11:97–104.
- Aljuaid H, Alturki N, Alsubaie N, Cavallaro L, Liotta A. Computer-aided diagnosis for breast cancer classification using deep neural networks and transfer learning. *Comput Methods Programs Biomed*. (2022) 223:106951. doi: 10.1016/j.cmpb.2022.106951
- Chen C, Zheng S, Guo L, Yang X, Song Y, Li Z, et al. Identification of misdiagnosis by deep neural networks on a histopathologic review of breast cancer lymph node metastases. *Nature*. (2022) 12:13482. doi: 10.1038/s41598-022-17606-0
- Oyelade ON, Ezugwu AE. A novel wavelet decomposition and transformation convolutional neural network with data augmentation for breast cancer detection using digital mammogram. *Sci Rep*. (2022) 12:5913. doi: 10.1038/s41598-022-09905-3
- Agaba AJ, Abdullahi M, Junaidu SB, Hassan IH, Chiroma H. Improved multi-classification of breast cancer histopathological images using handcrafted features and deep neural network (dense layer). *Intell Syst Appl*. (2022) 14:200066. doi: 10.1016/j.iswa.2022.200066
- Khan SI, Shahriar A, Karim R, Hasan M, Rahman A. MultiNet: a deep neural network approach for detecting breast cancer through multi-scale feature fusion. *J King Saud Univ Comput Inf Sci*. (2021) 34:6217–28. doi: 10.1016/j.jksuci.2021.08.004
- Srikanth VS, Sundar K. Pre-trained deep neural network-based computer-aided breast tumor diagnosis using ROI structures. *Intell Automat Soft Comput*. (2023) 35:63–78. doi: 10.32604/iasc.2023.023474
- Vasan D, Alazab M, Wassan S, Safaei B, Zheng Q. Image-based malware classification using ensemble of CNN architectures (IMCEC). *Comput Secur*. (2020) 92:101748. doi: 10.1016/j.cose.2020.101748
- Benbrahim H, Hachimi H, Amine A. Comparative study of machine learning algorithms using the breast cancer dataset. In: *Advanced Intelligent Systems for Sustainable Development (AI2SD'2019) Volume 2-Advanced Intelligent Systems for Sustainable Development Applied to Agriculture and Health*. Berlin: Springer (2020) p. 83–91. doi: 10.1007/978-3-030-36664-3_10

34. Alom MR, Farid FA, Rahaman MA, Rahman A, Debnath T, Miah ASM, et al. An explainable AI-driven deep neural network for accurate breast cancer detection from histopathological and ultrasound images. *Sci Rep.* (2025) 15:17531. doi: 10.1038/s41598-025-97718-5
35. Lilhore UK, Sharma YK, Shukla BK, Vadlamudi MN, Simaiya S, Alroobaea R, et al. Hybrid convolutional neural network and bi-LSTM model with EfficientNet-B0 for high-accuracy breast cancer detection and classification. *Nature.* (2025) 15:12082. doi: 10.1038/s41598-025-95311-4
36. Talukdar N, Kakati A, Barman U, Medhi JP, Sarma KK, Barman G, et al. Breast cancer detection redefined: integrating Xception and EfficientNet-B5 for superior mammography imaging. *Innov Pract Breast Health.* (2025) 7:100038. doi: 10.1016/j.ibreh.2025.100038
37. Munguía-Siu A, Vergara I, Espinoza-Rodríguez JH. The use of hybrid CNN-RNN deep learning models to discriminate tumor tissue in dynamic breast thermography. *J Imaging.* (2024) 10:329. doi: 10.3390/jimaging10120329