



Research paper

A novel diagnostic framework for breast cancer: Combining deep learning with mammogram-DBT feature fusion

Nishu Gupta ^a, Jan Kubicek ^{b,*}, Marek Penhaker ^b, Mohammad Derawi ^b^a VTT Technical Research Centre of Finland Ltd., Oulu, Finland^b VSB–Technical University of Ostrava, FEECS, Ostrava–Poruba, Czech Republic

ARTICLE INFO

Keywords:

Breast cancer detection (BCD)
 Mammogram images (MI)
 Digital breast tomosynthesis images (DBT)
 Feature fusion
 Deep learning (DL)
 Elman neural network (ENN)

ABSTRACT

Background and motivation: Breast cancer detection remains a critical challenge in medical imaging due to the complexity of tumor features and variability in breast tissue. Conventional mammography struggles with dense tissues, leading to missed diagnoses. Digital Breast Tomosynthesis (DBT) offers improved 3D imaging but brings significant computational burdens. This study proposes a novel framework using the Fully Elman Neural Network (FENN) with feature fusion to enhance the accuracy and reliability of breast cancer diagnosis.

Materials and methods: Mammogram images from the CBIS-DDSM dataset and DBT images from the Breast-Cancer-Screening-DBT dataset were used. The preprocessing step involved Extended-Tuned Adaptive Frost Filtering (Ext-AFF) to enhance image quality by reducing noise. Feature extraction was performed using Disentangled Variational Autoencoder (D-VAE), capturing critical texture features. These features were fused using Deep Generalized Canonical Correlation Analysis (Dg-CCA) to maximize feature correlation across modalities. Finally, a Fully Elman Neural Network was employed for classification, distinguishing between benign, malignant, biopsy-proven cancer, and normal tissues.

Results: The proposed FENN-based framework achieved superior classification performance compared to existing methods. Key metrics such as accuracy, sensitivity, specificity, and Matthew's correlation coefficient (MCC) demonstrated significant improvements. The fusion of mammogram and DBT images led to enhanced discriminative power, reducing false positives and negatives across various breast cancer classes.

Discussion and conclusion: The integration of mammogram and DBT image data with advanced machine learning techniques, such as D-VAE and FENN, enhances diagnostic precision. The proposed framework shows promise for improving clinical decision-making in breast cancer screening by overcoming the limitations of traditional imaging methods. The system's ability to handle complex interdependencies in imaging data offers substantial potential for earlier and more accurate diagnosis.

Future directions: Future research will focus on real-time clinical deployment of the framework, incorporating real-time image acquisition and analysis for faster diagnoses. Additionally, scaling the system for large datasets with varying image quality will further validate its robustness and applicability in diverse clinical environments.

Abbreviations: DBT, Digital Breast Tomosynthesis; FENN, Fully Elman Neural Network; BCD, Breast Cancer Detection; MI, Mammogram Images; DL, Deep Learning; D-VAE, Disentangled Variational Autoencoder; Dg-CCA, Deep Generalized Canonical Correlation Analysis; Ext-AFF, Extended-Tuned Adaptive Frost Filtering; CBIS-DDSM, Curated Breast Imaging Subset of the Digital Database; PAA, Probabilistic Anchor Assignment; ROI, Region of interest; ML, Machine Learning; RMSProp, Root Mean Square Propagation; CBR, Case-Based Reasoning; MCC, Matthew's correlation coefficient; FPR, False positive rate; FNR, False negative rate; SGDM, Stochastic Gradient Descent with Momentum; ACA-ATRUMNet, Atrous Convolution-based Attentive and Adaptive Trans-Res-UNet; ACA-AMDN, Atrous Convolution-based Attentive and Adaptive Multi-scale DenseNet; AUC, Area Under the Curve; RN50-CNN, ResNet50-Convolutional neural network; CLAHE, Contrast Limited Adaptive Histogram Equalization; APL-POA, Adaptive Prey Location-Based Pelican Optimization Algorithm; EDC, Ensemble Deep Learning; DNN, Deep Neural Network; Bi-LSTM, Bidirectional Long Short-Term Memory; DTCN, Deep Temporal Convolution Networks; GRU, Gated Recurrent Unit; DBM, Distorted Born Iterative Method; CSI, Contrast Source Inversion; ANN, Artificial Neural Network.

* Corresponding author.

E-mail address: jan.kubicek@vsb.cz (J. Kubicek).

1. Introduction

Among the most prevalent and lethal neoplasias worldwide, breast cancer demands the greatest concern in terms of its early diagnosis [1]. Therefore, it is with great importance that medical imaging, particularly mammography and DBT, has evolved to assist in developing diagnostics for more accuracy in breast cancer [2]. These are imaging techniques that provide detailed information about the tissue of the breast and allow the radiologist to highlight any possible malignancy [3]. Yet, with all the new technologies developed so far, the task of differential diagnosis of separate benign and malignant lesions remains a challenge, and diagnostic tools must be continuously improved [4]. Mammography has been taken for quite a long period as the gold standard of breast cancer examinations; X-ray images provided, one can notice tumors, microcalcifications, and other anomalies [5]. While generally effective, mammography itself cannot yield sufficient detail to allow definitive diagnoses, especially when lesions could be obscured by dense breast tissues [6]. The shortcoming in this area has given rise to a newer generation of imaging: digital breast tomosynthesis [7]. This takes several X-ray images at different angles, resulting in a three-dimensional image of the breast that gives clearer images of tumors and reduces false positives [8]. Mammography and DBT used in combination have recently emerged as a promising approach to screening for breast cancer, capitalizing on the respective strengths of both imaging modalities [9]. This enables a more detailed evaluation of the tissues of the breast and enhances the possibility of early tumor detection [10]. However, this combined information presents certain challenges regarding data processing and analysis, as this is information from diverse sources [11].

Medical imaging, particularly mammography and DBT, plays a pivotal role in diagnosing breast cancer, providing valuable insights into breast tissue abnormalities and potential malignancies [12,13]. However, despite advances in imaging technology, accurate detection, especially in dense breast tissue, remains a significant challenge. Conventional mammography, often regarded as the gold standard in breast cancer screening, struggles to detect small tumors in dense tissues due to poor contrast and overlapping structures, leading to false positives and missed diagnoses [14]. These challenges have led to the development of DBT, which captures multiple X-ray images from different angles, creating a three-dimensional image that helps reduce false positives and improve tumor visibility. However, while DBT offers advantages over mammography, it still faces limitations, including high computational complexity and extended processing times [15,16]. The combination of mammography and DBT has shown promise in addressing some of these challenges, but the integration of these modalities presents further complexities in data processing and analysis, as the images come from different sources with varying characteristics [17,18]. Furthermore, while recent advancements in machine learning have improved image analysis, the need for more robust, accurate, and efficient systems for breast cancer detection remains. Many existing approaches still struggle with differentiating benign from malignant lesions, especially in dense breast tissue, leading to high false positive rates and delayed diagnoses. Refining these techniques and developing new approaches are what the medical community hopes to further improve diagnostic accuracy in caring for patients with risks from breast cancer [19,20]. Given these challenges, there is a growing interest in utilizing advanced machine learning techniques to enhance the diagnostic capabilities of mammography and DBT. Fully Elman Neural Networks (FENN), a type of recurrent neural network, show promise in modeling complex relationships and temporal dependencies in image data, which could significantly improve classification accuracy in breast cancer detection. By integrating features from both mammography and DBT images, ML models like FENN can leverage the complementary strengths of these modalities to provide more accurate and reliable diagnoses.

Motivation: While there has been much progress, existing methods are still linked to the diagnosis of breast cancer with several drawbacks, especially those involving mammography and DBT. Generally,

conventional mammographic approaches result in unsatisfactory rates of tumor detection due to the poor contrast of tumors amidst dense breast tissues. This is further exacerbated by the limitations of DBT, where even as the modality improves the three-dimensional imaging, it is still fraught with high computational burdens and extensive processing time. Several of these techniques very often fail to achieve good accuracy due to the inability to differentiate between benign and malignant lesions, leading to a sizeable number of false positives or missed diagnoses. It is self-evident that the processing and analysis computational resources of high-resolution images in these modalities are huge, entailing longer turnaround times and higher costs. Motivated by the above challenges, there is a growing interest in developing new diagnostic methods that are more efficient and precise compared with available diagnostic imaging. There is an increasing need to enhance breast cancer diagnostic performance, in terms of both reducing time and computational complexity. Improvement in patient outcomes and smoothing of clinical workflows depend on the way such issues are addressed. Innovative approaches need to be developed in the context of integrating and optimizing data coming from different imaging sources and conducting enhanced analytical techniques for more reliable and quicker diagnoses. A significant reduction in diagnostic error and processing time, for much better clinical decision-making and improved patient care, is foreseen with a special focus on the mentioned areas.

The key contributions of the proposed framework are depicted below:

- The Extended-Tuned Adaptive Frost Filtering (Ext-AFF) helps to reduce noise and enhance the image's quality, improving the contrast, and hence yielding clear and more accurate feature extraction in both mammogram and DBT images.
- Disentangled Variational Autoencoder (D-VAE) is a deep learning generative model used in the extraction of detailed texture features from mammograms and DBT images. By considering disentangled VAE, the model can capture subtle differences between benign and malignant tissues with improved performance.
- Deep Generalized Canonical Correlation Analysis (Dg-CCA) fuses the features from mammogram and DBT images into a unified representation that maximizes the feature correlation and improves the discriminative power of the dataset.
- Fully Elman Neural Network (FENN) approach, employing progressive training and exploiting recurrent architectures that are quintessential in modeling intricate relationships and temporal dependencies of image data to improve the system for classifying breast cancer.

The rest of the paper is organized as follows: The previous study on integrated DL for breast cancer detection was defined in [Section 2](#). [Section 3](#) considers the suggested approach. The results and discussion were covered in [Section 4](#). Future research and the conclusion were discussed in [Section 5](#).

2. Related works

Among the many studies on DL-based BCD using mammogram and DBT images, some recent works are deliberated in the section that follows.

In 2022, Jiang, *et.al.*, [21] have developed a three-phase deep learning architecture based on the PAA algorithm that is utilized in mammography to detect and categorize breast cancer. To improve the performance of breast lesion identification and classification in mammograms, a three-stage deep learning framework based on the anchor-free object detection method Probabilistic Anchor Assignment (PAA) has been suggested. Firstly, the PAA-based detector was adopted to detect suspicious regions of potential breast lesions, including masses and calcifications. Subsequently, the lesions were classified and regressed using a two-branch ROI detector to minimize false positives. It

also presented a dense breast-tissue-aware threshold adaptive post-processing algorithm. Finally, a benign/malignant ROI classifier differentiated between lesion types by fusing global image characteristics and local ROI. Although highly promising, the model had some limitations in dealing with very small-sized lesions.

In 2023, Bouzar-Benlabiod, *et.al.*, [22] have suggested an innovative architecture for BCD based on the CNN—CBR method for classifying mammograms. A framework for using mammography images to diagnose breast cancer has been suggested, having a focus on explainable classification facilitated by a Case-Based Reasoning system. Because the quality of the extracted features ultimately determines CBR performance, the suggested approach incorporates a pipeline for data augmentation and image improvement to improve feature quality and guarantee a correct diagnosis. It does RoI extraction by using the U-Net architecture-based segmentation approach from mammograms. Integration of DL with CBR helps improve the overall process of classification where DL ensures accurate segmentation, and CBR provides an interpretable and efficient diagnosis. The CBIS-DDSM dataset was subjected to the recommended methodology and showed superior performances in comparison with traditional ML and DL methods. However, this suggested framework has several limitations as the method relies upon high-resolution mammogram images; thus, it works appropriately if the input image resolution was not good enough or if there was much noise in the image, which can raise uncertainties about the robustness under realistic conditions.

In 2024, Yaqub, *et.al.*, [23] have presented mammography images for a two-stage intelligent diagnosis of breast cancer. Suggested a deep learning system using mammography images for BC screening. First, data has been collected from reliable benchmarking sources. Next, images are segmented using the Atrous Convolution-based Attentive and Adaptive Trans-Res-UNet (ACA-ATRUNet) architecture. Finally, an Atrous Convolution-based Attentive and Adaptive Multi-scale DenseNet, known as ACA-AMDN, was utilized to detect BC. In addition, the MML-EOO method was used to further refine the hyperparameters of ACA-ATRUNet and ACA-AMDN to increase accuracy. In addition, other performance metrics of the model were taken into account and compared with traditional methods. One potential weakness in the suggested framework was the need for complex hyperparameter tuning, which can be costly to compute and time-consuming, particularly with big datasets.

In 2022, El-Shazli, *et.al.*, [24] have developed a DL-based computer-aided model for efficient diagnosis of digital breast tomosynthesis 3D imaging. It developed an AlexNet-based computer-aided diagnostic system based on color feature mapping and DBT augmentation, with some modifications, optimized by techniques such as Adam, RMSProp, and SGDM. Two experimental scenarios were carried out: the first analyzes the performance of six DL models concerning feature extraction, while the second one assesses how well Mod_AlexNet performs in comparison to regular AlexNet for varying batch sizes. Hence, it increases the performance of the classification model concerning accuracy. However, the main disadvantage was an increase in computational complexity because of multiple optimization techniques involved, which may necessitate more computational time and other resources.

In 2023, Lee, *et.al.*, [25] have suggested a classification of breast cancer in digital breast tomosynthesis images using a transformer-based deep neural network. It has been recommended to use contextual information from adjacent sections in a deep neural network model to improve breast cancer diagnosis in DBT images. This approach is based on the transformer and was compared with two baseline models: a 3D convolution model and another with analysis by individual sections in 2D. Trained on 5174 DBT studies and validated with 1000 studies, the model was put to test on 655 studies. Therefore, the AUC, sensitivity, and specificity for the transformer model outperformed those from the 2D model considerably with only 25 % of floating-point operations compared to the 3D model. A drawback of the transformer approach was that it relies on large volumes of neighboring data. This may limit its

performance when incomplete or sparse DBT stacks are present.

In 2023, Hassan, *et.al.*, [26] have established breast cancer tumor identification in a dynamic ensemble approach based on DL for tomosynthesis images. Created a strategy that combines a powerful individual breast tumor detector with a dynamic ensemble approach to identify breast cancers in DBT images. Here, a deep neural network has been employed by the dynamic ensemble approach to determine which IBDT is optimal based on the features of an input image. Addressing the problems of scarce data and modeling multiple scenarios, the IBDTs make use of the robustness of DL architectures and introduce some methods for enhancing data, namely channel concatenation and replication. On a publicly accessible DBT dataset, the suggested approach outperformed two state-of-the-art ensemble techniques: non-maximum suppression and weighted box fusion. Dynamic selection and data augmentation employed in this method can be very tricky to implement in real applications.

In 2024, Oba, *et.al.*, [27] have presented a DBT combined with a DL algorithm to forecast breast cancer's expression of Ki-67. The work that was suggested aimed to create a DL algorithm that could forecast the degree of Ki-67 expression in DBT images. Following institutional review board permission, 499 individuals (126 of whom had been diagnosed with breast cancer) had their levels of Ki-67 expression measured. The Xception architecture of a DL model was used to classify high/low levels of Ki-67 expression. Performance was measured in terms of different metrics: accuracy, sensitivity, specificity, and AUC. It has shown a great overall performance but with important fluctuations between the various radiological findings. The low sensitivity of the model constitutes the limit of the latter's performance in predicting, with a high degree of precision, the low level of Ki-67 expression.

In 2024, Kumar et al., [28] have defined the effective heuristic-aided ensemble classifier for BCD using DBT and mammogram images. Initially, contrast-limited Adaptive Histogram Equalization (CLAHE) and histogram equalization (HE) schemes were utilized to enhance image quality. Then, the DeepLabv3 model was applied to perform image segmentation. Various features like texture features, color features, shape features, deep features, statistical features, and morphological features were extracted. Then, the optimal features were identified and their respective weights were refined using the Enhanced Adaptive Prey Location-Based Pelican Optimization Algorithm (APL-POA). Finally, several ensemble classifiers like DNN, LSTM, GRU, Bi-LSTM, and DTCN were utilized to identify breast cancers. The hyperparameters for each classifier within the ensemble were fine-tuned using the enhanced APL-POA algorithm to improve the model's performance. However, it may struggle with complex cases where the boundaries of tumors were not clearly defined or were obscured by noise or artifacts.

In 2024, Kassis et al., [29] have established the CAD system to enhance the diagnostic performance of DBT by utilizing the capabilities of vision transformers (ViT). This approach employed a neural network to extract features from the 2D slices of DBT, followed by post-processing that incorporated information from adjacent slices to classify the entire 3D scan. A transfer learning technique was applied to train and validate the CAD framework on a specialized dataset containing 3831 DBT scans, with testing conducted on 685 scans. Among the architectures evaluated, the Swin Transformer (ST) outperformed both ResNet101 and the vanilla ViT. Limitations of the study included the potential bias due to the dataset's composition, the reliance on high-quality scan data, and the need for further testing across more diverse populations and imaging conditions to assess the generalizability of the model.

In 2023, Franceschini et al., [30] have defined the DBM, ANN, and CSI techniques to detect the presence of tumors using tomographic data. The method was evaluated using a simulated database and demonstrated promising results, especially in scenarios involving very small tumor masses. In such cases, traditional reconstruction techniques struggled to detect suspicious tissues, whereas this approach accurately

identified these small masses as potentially pathological. As a result, this method showed potential for early detection, particularly when the tumor size is minimal. Limitations included the use of a simulated database, which may not fully capture the variability of real-world data, and the need for further validation with clinical datasets to assess the model's robustness in diverse patient populations and imaging conditions.

3. Materials and methods

In this section, an advanced framework for detecting breast cancer by using FENN, coupled with deep feature fusion of mammograms and DBT images is discussed. The proposed approach aims to improve accuracy and robustness by combining complementary imaging modalities. In this proposed framework, multiple processing steps are involved: image preprocessing, feature extraction, feature fusion, and classification. Mammogram images are drawn from the CBIS-DDSM dataset, whereas Breast-Cancer-Screening-DBT is the dataset from which DBT images are generated.

The block diagram of the suggested strategy for identifying breast cancer is displayed in Fig. 1. The workflow begins with image acquisition from two datasets: CBIS-DDSM and BCS-DBT. These datasets provide mammogram and breast computed tomography images, respectively. The acquired images undergo preprocessing using the Extended-Tuned Adaptive Frost Filtering method, which enhances image quality by reducing noise and preserving critical details. The pre-processed images are then used for feature extraction through a Disentangled Variational Autoencoder (D-VAE), a deep learning model designed to capture meaningful and independent features from complex data. The extracted features from both datasets are then fused using

Deep Generalized Canonical Correlation Analysis (DGCCA), which combines multimodal information effectively by finding correlated patterns in the datasets. Finally, the fused features are input into a Fully Elman Neural Network, a recurrent neural network architecture, for breast cancer detection. The model classifies images into several categories, including Benign, Malignant, Actionable, B-p Benign, B-p Cancer, and Normal. Each classification provides valuable insights for medical decision-making. This framework integrates preprocessing, feature extraction, fusion, and classification steps to ensure accurate and efficient detection of breast cancer, demonstrating its potential in improving diagnostic accuracy and assisting healthcare professionals.

3.1. Image acquisition

Initially, the breast cancer imaging data are gathered from two major sources. Mammogram images are obtained from the CBIS-DDSM Breast Cancer Image Dataset [31], a very extensive dataset consisting of 2620 cases amounting to a total of 10,480 images with annotations regarding abnormalities like masses and calcifications. These images are of standard mammogram size (1024×1024 pixels) and come with annotations that classify the cases as benign or malignant. Moreover, the DBT images were gathered from the Breast Cancer Screening DBT dataset [32], in which high-resolution 3D volumetric images are provided for more detailed information. It includes 2057 images from 1063 patients, captured using a 3D imaging technique that offers detailed views of breast tissue. This dataset also includes annotations labeling benign and malignant findings. The large sample sizes and variety in imaging modalities enhance the model's ability to generalize and improve its performance across different imaging techniques. In this regard, there are multiple class types included in the dataset, such as normal, actionable,

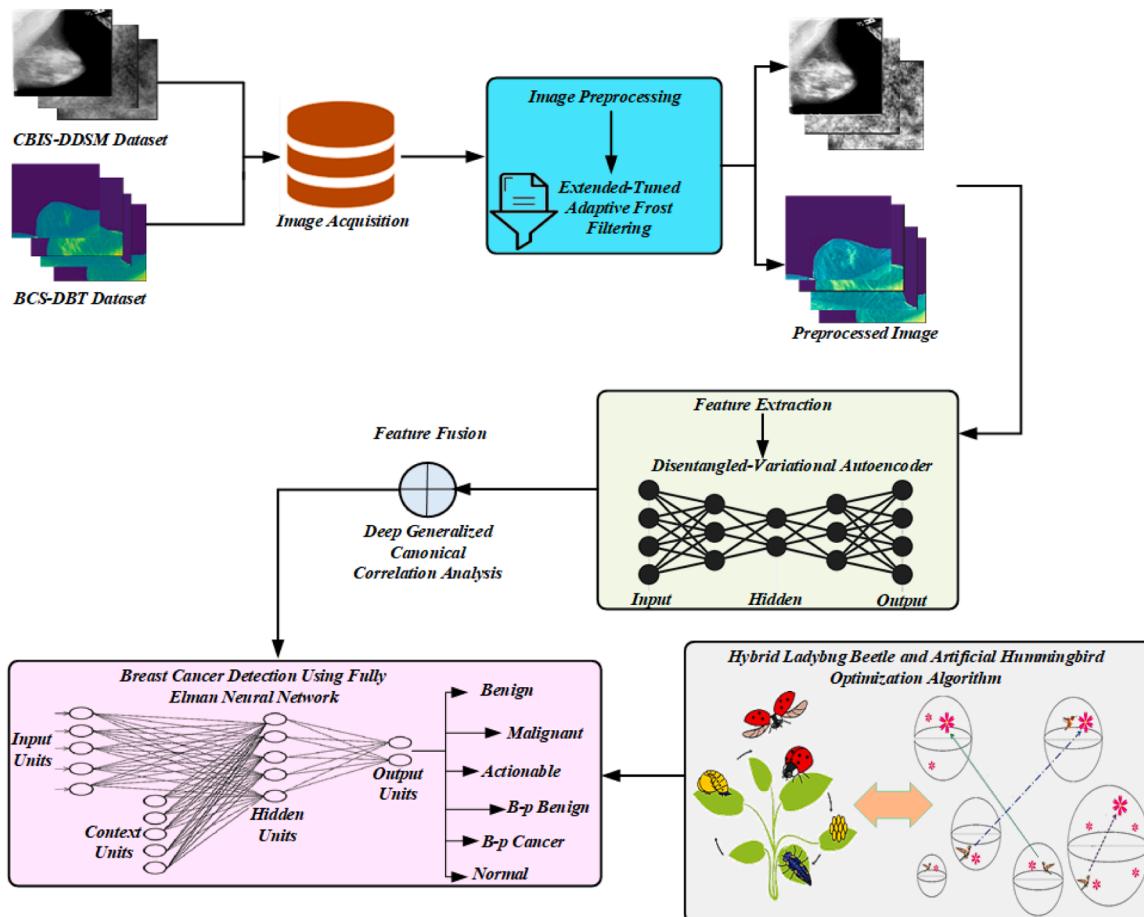


Fig. 1. Block schematic of suggested breast cancer detection methodology.

biopsy-proven benign, and biopsy-proven cancer cases. Then the collected input images are moved to the preprocessing stage due to the need for standardization and improvement of image quality, which is necessary to raise the precision and potency of the analysis that follows.

3.2. Image denoising using extended-tuned adaptive frost filtering (Ext-AFF) technique

The raw mammograms and DBT images collected from the database exhibit a high noise level, which impairs further image analysis's efficiency. Traditional frosting-filtering techniques reduce unwanted speckle noise but often compromise the structural details of the images. The proposed framework addresses this limitation by employing an Extended-Tuned Adaptive Frost Filtering (Ext-AFF) technique, where each pixel is tuned with values to enhance the quality of the image while preserving important structural information [33]. The Ext-AFF approach represents an extension of the standard frost filters within the Lee and Kuan filters through an adaptive tuning factor, which would, in effect, optimize trade-off effectiveness between edge detail retention and noise reduction. Unlike Lee and Kuan filters, the Frost filter used in Ext-AFF does not assume any simple rectilinear relationship between the real and filtered images.

To enhance the robustness and generalization of the proposed framework across varying image qualities, detailed data preprocessing is applied to both the CBIS-DDSM and Breast-Cancer-Screening-DBT datasets. Preprocessing involved using Extended-Tuned Adaptive Frost Filtering (Ext-AFF) to improve image quality by effectively reducing noise while preserving critical edge information. This step was crucial for ensuring consistent feature extraction, especially given the variability in image resolutions and noise levels across datasets. The outcome of the frost filtering (FF) technique is mathematically modeled by the Eq. (1) as follows,

$$\bar{Z}(u, v) = \sum_a \sum_b P_{ab} m_{ab} / \sum_a \sum_b m_{ab}, \quad k_{ab} = \exp(-QC_Z^2 d_{ab}) \quad (1)$$

Here, (u, v) indicates the position of the present pixel, $\bar{Z}(u, v)$ indicates the filter outcome, P_{xy} represents the pixel values among the center window at (u, v) , $Q(Q > 0)$ indicates the tuning factor (TF), C_Z shows the variation coefficient, which is the ratio of the instance mean to the instance standard deviation (SD), and d_{xy} represents the distance between the current and center pixels.

In the proposed filtering scheme, the TF Q is highly essential for enhancing the performance of the Ext-AFF technique. When Q value is low, it overwhelms noise but fails to preserve the essential information and edge pixels. When Q value is high, the noises get suppressed but the preservation of edge pixels gets improved. Moreover, t_0 evaluation is used to represent the center pixel due to the instance mean and variance for can be calibrated through total pixels in the present window. Then, the extended-tuned factor in FF can be mathematically formulated in the equations (2–4),

$$Q(a, b) = T(t_0) \times U(a, b) \quad (2)$$

$$T(t_0) = |Z(t_0) - m(t_0)| / \lambda(t_0) \quad (3)$$

$$Q(a, b) = \frac{|Z(a, b) - Z(t_0)|}{\sum_{m=u-L}^{j+L} \sum_{n=p-L}^{v+L} |Z(m, n) - Z(t_0)|} \quad (4)$$

Here, $Q(a, b)$ indicates the extended tuned factor calibrated based on $T(t_0)$, $U(a, b)$ indicates the grayscale adjacent pixels, $Z(t_0)$ represents the center pixel value. Moreover, $m(t_0)$ and $\lambda(t_0)$ represents the center window's current mean and standard deviation of pixels at t_0 and L indicate the window size and square type window is utilized. The proposed Ext-AFF methodology is capable of producing high-quality images with preserved critical structural features that are

important for the later stages of analysis.

3.3. Feature extraction using disentangled-VAE

The Disentangled-VAE extracts statistical, gray-level, morphological, and texture features from the pre-processed mammogram and DBT images [34]. For effective encoding of label information in the representation, it is assumed that the latent space consists of two components: a disentangled variable, and a non-interpretable variable. The categorical information is captured by the disentangled variable, and it is important for classification tasks like identifying benign, malignant, and other categories of things that are breast cancer-related. Using label information aids in constraining the disentangled variable to improve the accuracy of the prediction. However, the non-interpretable variable encapsulates all other uncertain or latent information stemming from the images, which may even include noise or non-categorical aspects. For better clarity, demand the non-interpretable variable u and the disentangled variable asv . The encoder function, which maps input data x to the latent space is defined as $q_\phi(u, v|x)$ where u and v are assumed independent, conditional on x . More formally, this is expressed in equation (5),

$$q_\phi(u, v|x) = q_\phi(u|x) q_\phi(v|x) \quad (5)$$

This assumption ensures that the categorical information is captured mainly by v and that u carries minimal information about class labels. The representation u and its disentangled representation v are then encoded separately as $q_\phi(u|x)$ and $q_\phi(v|x)$ respectively. The objective function, thus modified to take on this latent space division, can be written as in equation (6),

$$L(\theta, \phi; x, \lambda) \geq E_{q_\phi(u|x), q_\phi(v|x)} [\log p_\theta(x|u, v)] + \lambda (\log p(v) + \log p(u) - \log q_\phi(u|x) - \log q_\phi(v|x)) \quad (6)$$

This results in regularized reconstruction error for the latent variables u and v , using KL -divergence terms KL_u and KL_v for each of these variables are represented in Eq. (7) as follows,

$$L(\theta, \phi) = RE(u, v) - \lambda(KL_u + KL_v) \quad (7)$$

Here, $RE(u, v)$ is reconstruction error conditioned on the latent variables and KL_u , KL_v are the Kullback-Leibler divergence between the approximate and prior distributions for u and v , respectively. This encoding scheme makes sure that the categorical information is well-captured by v and, if partial labels are available, it can regularize the disentangled variable more effectively to extract accurate features. Then, after feature extraction, a technique of feature fusion called Deep Generalized Canonical Correlation Analysis (Dg-CCA) is applied. It integrates features extracted from mammogram and DBT images in one unique form; this increases the capability of the model to differentiate between various categories of breast cancer.

3.4. Deep generalized canonical correlation analysis for multiview feature fusion

In this section, Dg-CCA is adopted for Multiview representation learning due to the efficient fusion that arises when the features extracted from both mammogram and DBT images are concerned based on the power of deep neural networks [35]. Although Deep CCA is restricted to just two views, Dg-CCA allows for a case with multiple views by maximizing the correlation between various learned representations from different modalities. For each view-mammogram and DBT images apply multiple nonlinear layers on data and take a minimum of reconstruction error across outputs of all views. The fusion of mammography and Digital Breast Tomosynthesis (DBT) images significantly enhances the discriminative power of a model by combining complementary information from both imaging techniques. Mammography, with its 2D projection, is effective for detecting calcifications and

architectural distortions, but it can be limited by tissue overlap, which may obscure smaller or deeper lesions. On the other hand, DBT, a 3D imaging technique, provides multiple slices of the breast, reducing the effects of tissue overlap and improving the visibility of lesions, particularly in denser breast tissues. By integrating these two modalities, a model can leverage the strengths of both: the high-resolution detail of mammography and the 3D structural clarity of DBT. This fusion enables the model to better distinguish between benign and malignant abnormalities, as it captures more comprehensive spatial information and fine details of tissue structures. Image modality fusion helps reduce false positives and false negatives by combining complementary information from different imaging techniques, thereby providing a more comprehensive view of the tissue. For instance, mammography is highly sensitive to microcalcifications and is good at detecting certain types of breast cancer, but it can struggle with dense breast tissue, often leading to false positives (where benign lesions are incorrectly identified as cancer) or false negatives (missed tumors due to tissue overlap). On the other hand, Digital Breast Tomosynthesis (DBT), with its 3D imaging capability, provides better separation of overlapping tissues, reducing the likelihood of missed tumors (false negatives) and improving visualization of lesions hidden in dense breast tissue. This fusion enhances the accuracy of detection, ensuring that benign conditions are not mistaken for cancer (reducing false positives), while also improving the identification of tumors that may otherwise be overlooked (reducing false negatives). Reducing false positives and false negatives in breast cancer diagnosis is crucial for several reasons, as it directly impacts both patient outcomes and clinical practice. False positives occur when a non-cancerous condition is incorrectly diagnosed as cancer, leading to unnecessary follow-up tests, biopsies, and procedures, which can cause patient anxiety, physical harm, and increased healthcare costs. On the other hand, false negatives occur when a cancerous lesion is missed, leading to delayed diagnosis, which can result in more advanced disease, poorer prognosis, and a reduced likelihood of successful treatment outcomes. In clinical practice, minimizing these errors improves the accuracy of early breast cancer detection, allowing for timely intervention and treatment, which is associated with higher survival rates and better long-term health outcomes. It also helps to optimize the

allocation of healthcare resources by reducing unnecessary procedures, making the diagnostic process more efficient and cost-effective. Fig. 2 shows a schematic of Dg-CCA for multi-modal feature fusion in mammogram and DBT image classification.

This is done via the optimization of weight matrices W_j for each layer together with the discovery of linear transformations U_j mapping network outputs on a shared representation G . Mathematically, the problem can be expressed as minimizing the error in equation (8),

$$\min_{U_j, G} \sum_{j=1}^J \| G - U_j^T f_j(X_j) \|_F^2 \quad (8)$$

Subject to $GG^T = I_r$, where $f_j(X_j)$ is the nonlinear transformation of the input data, and G is the learned shared representation. The unified feature representation that captures the inter-view correlations between mammograms and DBT gets classified. The fused features are then passed to the classification process for final diagnosis.

3.5. Fully Elman neural network for breast cancer classification

In this section, a Fully Elman Neural Network (FENN) is used to classify breast cancer into benign, malignant, actionable, biopsy-proven benign, biopsy-proven cancer, and normal. FENN is one kind of recurrent neural network suited for sequential data types, which applies to the domain of breast cancer detection as it captures temporal dependencies and relationships in data [36]. The architecture uses the context layer to store past states, hence remembering previous inputs through an improvement in the pattern being detected over time. The FENN architecture consists of four main layers: the input, the hidden, the context, and the output layer. Features from a medical image, like a mammogram, feed into the input layer. The hidden layer is responsible for processing these inputs and produces hidden states. The network could learn from earlier sequences of hidden states, maintained inside the context layer, and fed back into the hidden layer. The output layer then produces predictions regarding the classification of breast cancer. The interconnections within the context layer are fully connected and allow each of the hidden neurons to be computationally influenced by all of the previous hidden states. This configuration maximizes the sensitivity of the network to capturing complex patterns in time or subtle features in medical images. The FENN architecture is depicted in Fig. 3.

Analyzing mammogram and DBT images often necessitates contextual interpretation due to subtle interrelated features inherent in these data sets. FENN's recurrent connections facilitate more nuanced pattern recognition in texture density and structural irregularities of breast tissue by leveraging complex interactions. FENN's architecture facilitates the integration of varied feature sets like those yielded by multiple imaging modalities thus bolstering complex feature learning. FENN leverages these capabilities, yielding robust classification performance that distinguishes benign malignant biopsy-proven cancer and normal tissues with superior accuracy versus standard neural networks. Breast cancer detection tasks benefit greatly from their proficiency in handling diverse complex data sets effectively. The mathematical model of the FENN for this classifying task can now be stated as follows in equation (9) - (13),

$$b(s) = f(k_2 a(s) + k_3 w(s-1) + b_{d2}(s) + y_2) \quad (9)$$

$$a(s) = g(a_d(s) + k_1 w(s-1) + b_{d1}(s) + y_1) \quad (10)$$

$$a_d(s) = k_4 a_d(s-1) + k_5 a(s-1) \quad (11)$$

$$b_{d1}(s) = k_6 b_{d1}(s-1) + k_7 b(s-1) \quad (12)$$

$$b_{d2}(s) = k_8 b_{d2}(s-1) + k_9 b(s-1) \quad (13)$$

where the input layer to the hidden layer, the hidden layer and the output layer, and the hidden layer and the input layer have the following

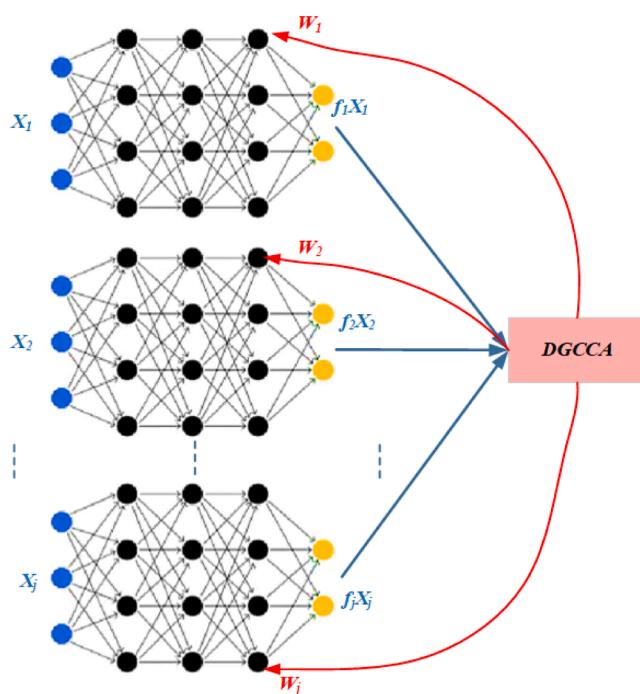


Fig. 2. A Schematic of Dg-CCA for Multi-Modal Feature Fusion in Mammogram and DBT Image Classification.

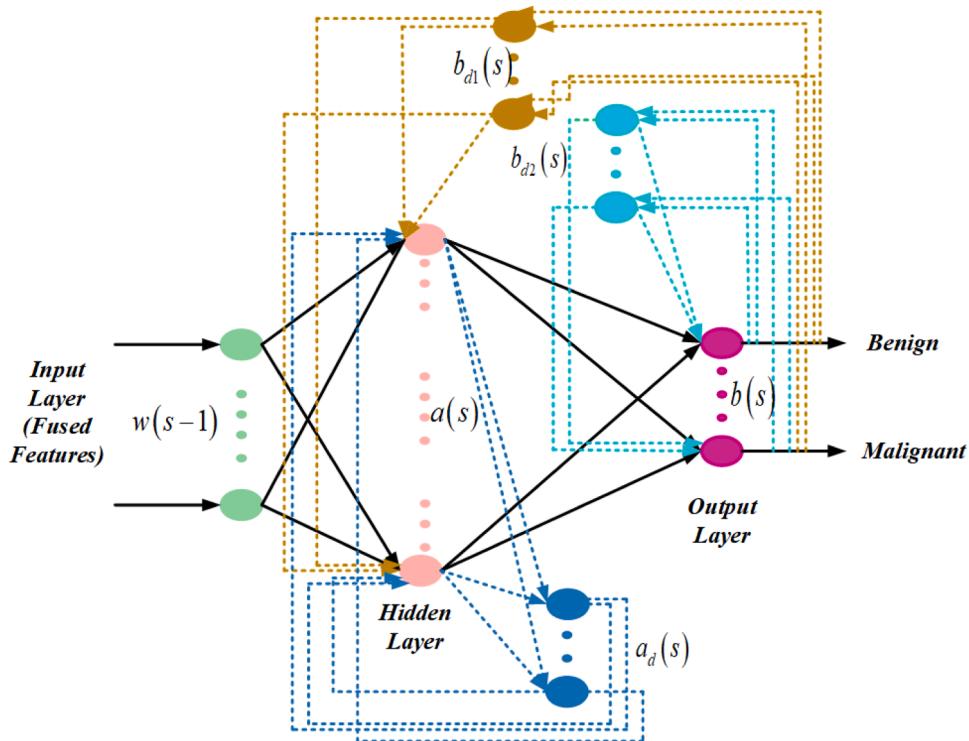


Fig. 3. The FENN model's architecture.

connection weights, respectively, k_2 , k_3 , and k_1 .

The output context layers 1 and 2, as well as the context layer for the hidden layer at sare indicated by the symbols $b_{d1}(s)$, $b_{d2}(s)$, $a_d(s)$, which specify the length of the cloudlet, the total amount of data processed, and the computation time, in that order. The bias of the hidden layer and the output layer is y_2 and y_1 . Weights of the recurring relationship between the current sand $s - 1$ of a_d , b_{d1} and b_{d2} are denoted by k_4 , k_6 , and k_8 accordingly. The connection weights of $a(s - 1)$ to $a_d(s)$, $b(s - 1)$ to $b_{d1}(s)$, and $b(s - 1)$ to $b_{d2}(s)$ are indicated by k_5 , k_7 , and k_9 . The vectors of the hidden layer and output layer at sare denoted by $b(s)$ and $a(s)$, which expresses the task input and output size respectively, and the input layer's vector at $s - 1$ is represented by $w(s - 1)$. The functions that activate the output layer (represented by f) and hidden layer (represented by g) are the hyperbolic tangent sigmoid and SoftMax, respectively. This approach enables the model to remember the temporal dependency between medical image sequences in the context layers. This leads to a more accurate classification of breast cancer.

3.6. Computational complexity analysis

The computational complexity of the FENN and d-VAE techniques can be analyzed based on their respective architectures and operations during training and inference. The FENN is a type of RNN with feedback loops in its hidden layer. Its complexity depends on the size of input data, the number of hidden neurons, and the number of layers. The overall training and inference computation complexity obtained by the FENN framework is interpreted as, $O(T.N_h(N_x + N_h) + N_h N_y)$ and $O(T.(N_h N_x + N_h^2))$ respectively. Here, T , N_h , N_x , N_h , and N_y indicates the time steps, hidden neurons, and Size of input and output vector respectively. The d-VAE has two main components: an encoder-decoder architecture and the variational inference mechanism. Its complexity depends on the size of input data, latent dimensions, and network architecture. The overall training and inference computation complexity obtained by the d-VAE framework is interpreted as, $O(L.N_l N_x + L.N_l N_x + N_x)$ and $O(L.N_l N_x + L.N_l N_z)$ respectively. Here, N_x , N_z , L , and N_l deliberates the input data size, Latent space size (dimension

of a latent vector), number of layers in the encoder/decoder, and a number of neurons per layer respectively. Fig. 4 deliberates the flowchart of the proposed framework.

4. Result and discussion

This section highlights breast cancer detection simulation results using FENN and a feature fusion of mammogram and digital breast tomosynthesis images. The simulation process is held in the MATLAB platform and two publicly available databases are considered in this framework. The experimentation is performed by taking 80 % of the training data, 10 % of the testing data, and 10 % of the validation data. The developed FENN framework was employed to mark and validate tumor boundaries in images from both datasets. For images without tumors, clear criteria were used to exclude ambiguous cases, ensuring high data reliability. These calibrated categories were used for model training and evaluation, providing a robust dataset split for tumor and non-tumor images. The suggested method's effectiveness is compared to current techniques, such as a three-stage DL architecture built on the PAA algorithm used for the identification and classification of breast cancer in mammograms (PAA) [21], an innovative architecture for breast cancer detection based on the CNN-CBR method for classifying mammograms (U-Net-DL) [22], mammography images for a two-stage intelligent diagnosis of breast cancer (ACA-AMDN) [23] for CBIS-DDSM dataset. Additionally, intelligent computer-aided model based on DL for efficient diagnostic 3D imaging of digital breast tomosynthesis (Mod_AlexNet) [24], Classifying breast cancer in digital breast tomosynthesis images using a transformer-based deep neural network (T-DNN) [25], identifying breast cancer tumors in tomosynthesis images with a dynamic ensemble approach based on DL (RN50-CNN) [26] for Breast-Cancer-Screening-DBT dataset. Table 1 tabulates the hyperparameter values of the proposed FENN and d-VAE framework. The output image result of the suggested method is displayed in Table 2.

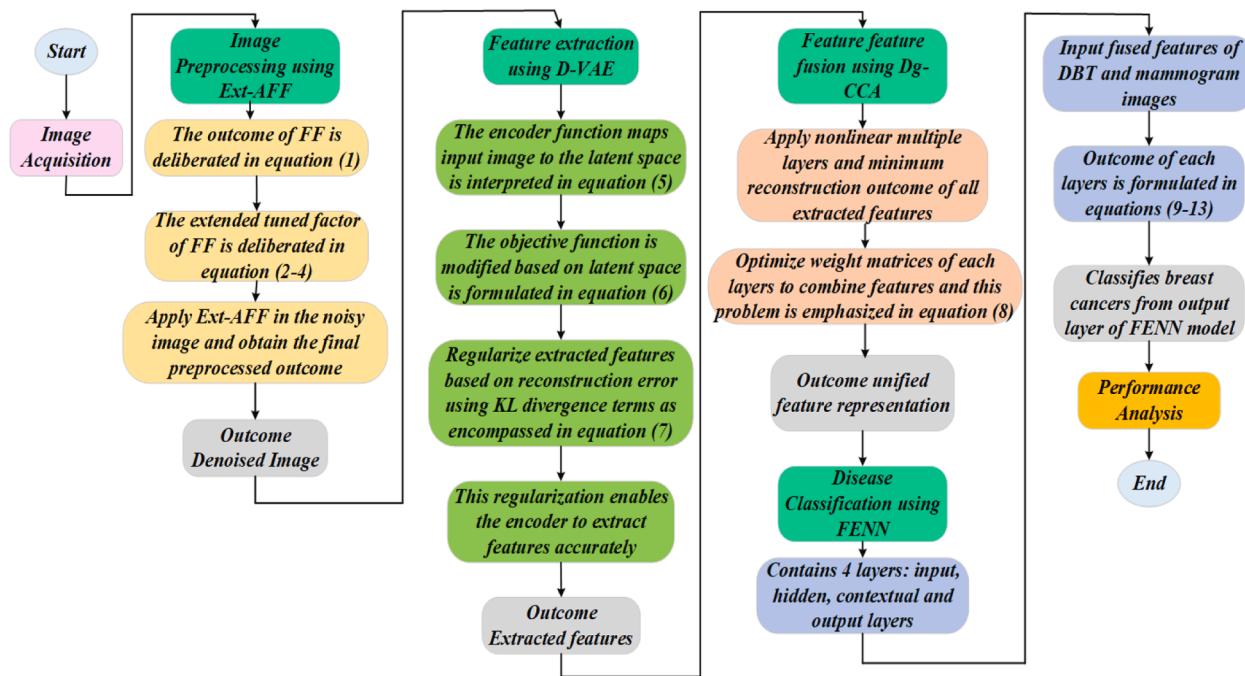


Fig. 4. Flowchart of the proposed framework.

Table 1
Hyperparameter values of the proposed FENN and d-VAE framework.

Hyperparameters	FENN	D-VAE
Learning Rate	0.001	0.0005
Hidden Units	128	64
Dropout Rate	0.3	0.2
Batch Size	32	16
KL-Divergence Weight (Beta)	–	0.01

4.1. Performance measures

To get further insight into the suggested framework, performance metrics like recall, F1-score, FNR, FPR, precision, specificity, accuracy, MCC, ROC, and sensitivity are assessed. These metrics are vital because they address different aspects of the model's performance. They ensure a balance between identifying true disease cases (sensitivity) and avoiding over-diagnosis (specificity). Precision and recall trade-offs are particularly significant in healthcare, where false negatives (missed diagnoses) and false positives (unnecessary follow-ups) have critical implications. Using multiple metrics provides a holistic view of the model's reliability and robustness, which is essential in medical applications where patient outcomes are directly impacted. Here, employ the metrics of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) to better understand the proposed approach.

True Positives (TP): The number of cases when breast cancer was accurately predicted by the model.

True Negatives (TN): The number of cases where the model accurately indicated there would be no breast cancer.

False Positives (FP): The number of cases where the model predicted its existence incorrectly.

False Negatives (FN): The number of cases when the model correctly predicted that breast cancer would not occur when it is present.

4.1.1. Accuracy

The ratio of accurately anticipated occurrences to the total number of instances is known as accuracy. Its definition is given in Eq. (14).

$$\text{Accuracy} = \left(\frac{TP + TN}{TP + TN + FP + FN} \right) \quad (14)$$

4.1.2. Sensitivity/recall

The actual number of positives that the suggested model accurately detects is measured by sensitivity. This is defined as in equation (15),

$$\text{Sensitivity} / \text{Recall} = \left(\frac{TP + TN}{FN} \right) \quad (15)$$

4.1.3. Precision

The percentage of accurate positive forecasts is known as precision. This is defined as in equation (16),

$$\text{Precision} = \left(\frac{TP}{TP + FP} \right) \quad (16)$$

4.1.4. F1-Score

The harmonic mean of Precision and Recall is known as the F1-Score. As a result, it provides a certain level of memory and accuracy balance. It is particularly helpful in cases of unequal class distribution. The definition is given in Eq. (17).

$$\text{F1-score} = 2 * \left(\frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \right) \quad (17)$$

4.1.5. Specificity

Specificity is measured by Eq. (18), which indicates how well the recommended approach detects undesired events.

$$\text{Specificity} = \left(\frac{TN}{TN + FP} \right) \quad (18)$$

4.1.6. Matthews correlation coefficient (MCC)

All four variables (TP , TN , FP and FN) are considered by MCC, which produces a single number that captures the model's effectiveness while taking both over- and under-prediction into consideration. Eq. (19) is used to compute it.

Table 2

Output image result of proposed approach.

<i>Datasets</i>	<i>Input Image</i>	<i>Preprocessed Image</i>	<i>Cancer Type</i>
<i>CBIS-DDSM</i>			<i>Benign</i>
			<i>Malignant</i>
			<i>Normal</i>
<i>BCS-DBT</i>			<i>Actionable</i>
			<i>Biopsy-Proven Benign</i>
			<i>Biopsy-Proven Cancer</i>
			<i>Normal</i>

$$MCC = \frac{(TP * TN - FP * FN)}{\sqrt{((TP + FP) * (TP + FN) * (TN + FP) * (TN + FN))}} \quad (19)$$

4.2. Simulation result of proposed model compared with existing techniques for CBIS-DDSM dataset

The results of the simulation using the recommended technique are shown in Fig. 4–9. Here, the efficacy of the suggested approach is contrasted with that of the state-of-the-art methods, such as ACA-AMDN, U-Net-DL, and PAA.

Fig. 5 represents the confusion matrix performance of an applied classification model on three classes: Benign, Malignant, and Normal. The matrix shows that for the Benign class, it had misclassified 3 instances into the Malignant category and none into the Normal category, while correctly predicting 527 out of 530 instances. In the Malignant category, the model correctly predicted 495 out of 498 instances, misclassifying 1 instance as Benign and 2 instances as Normal. In the Normal category, classification is almost perfect with 542 instances correctly classified among 544, and with just two misclassifications as Malignant. The matrix shows very high accuracy and minimal misclassifications on the whole, thus supporting the fact that the model classifies well among the three categories.

Figs. 6(a) and 6(b) display the performance metrics of several techniques for classifying breast cancer using the CBIS-DDSM dataset, with a focus on accuracy and F1-score for the various classes. The Proposed model is always on the top of all classes and the highest accuracy achieved was 98.91 % for Benign, 98.63 % for Malignant, and 98.00 % for Normal. Other method's accuracy is drastically low. Still, PAA shows the

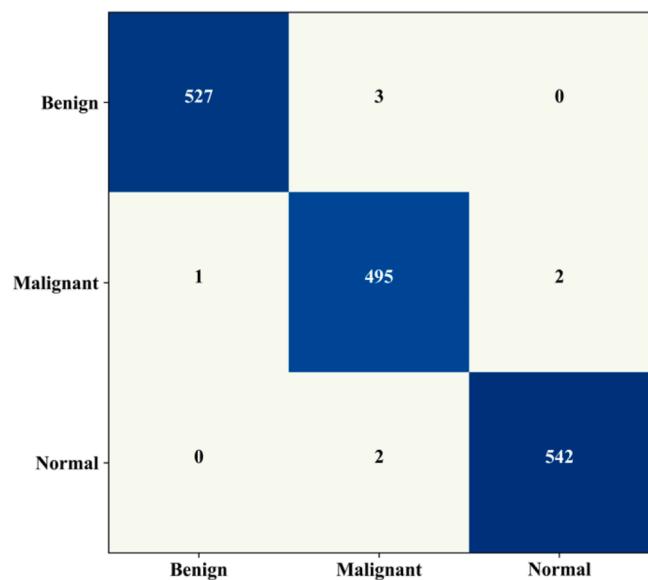


Fig. 5. Confusion matrix for breast cancer classification on the CBIS-DDSM dataset.

highest accuracy, which was 90.30 % for Benign, 90.30 % for Malignant, and 78.90 % for Normal, U-Net-DL displays 93.70 % for Benign, 96.75 % for Malignant and 84.50 % and ACA-AMDN shows 96.96 % for Benign,

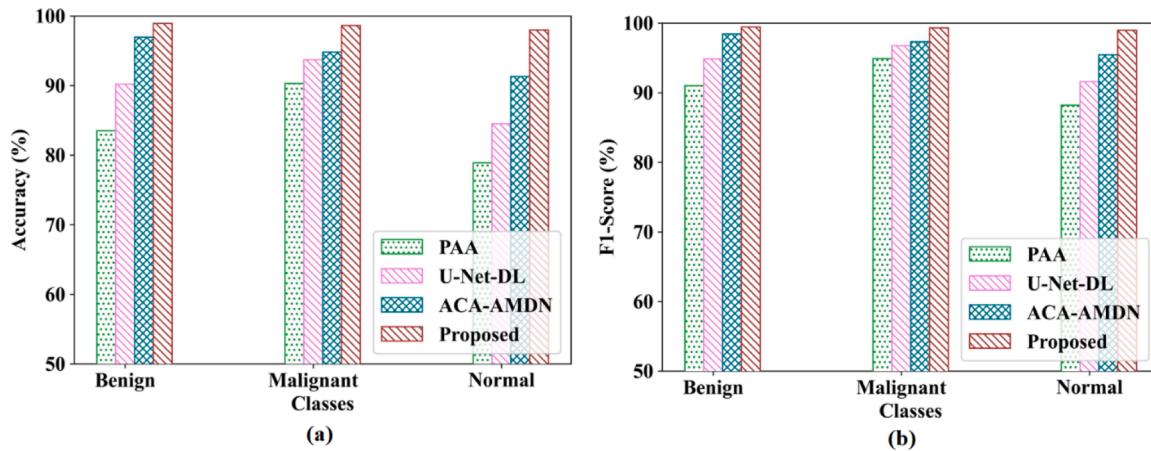


Fig. 6. Performance of (a) accuracy (b) F1-score analysis for CBIS-DDSM dataset.

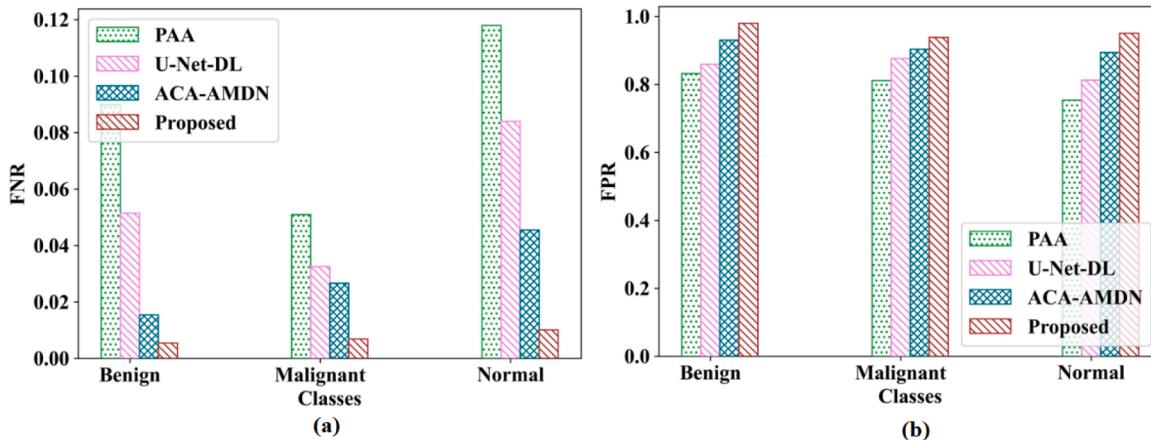


Fig. 7. Performance of (a) FNR (b) FPR analysis for CBIS-DDSM dataset.

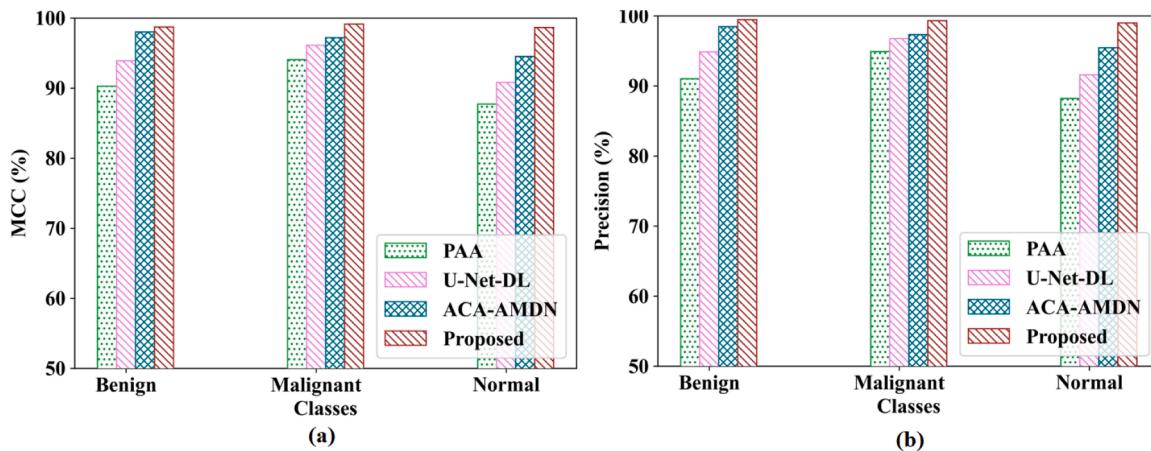


Fig. 8. Performance of (a) MCC (b) precision analysis for CBIS-DDSM dataset.

97.33 % for Malignant and 91.30 % for Normal. This shows that the proposed model is not only accurate in distinguishing between benign and malignant classes but also performs well in classifying the Normal category, which is often more challenging due to its similarity to benign cases. The F1 score captures the trade-off of the precision and recall for each class. Here also, the proposed model again highlights the highest F1-scores: 99.45 for Benign, 99.31 for Malignant, and 98.99 for Normal. The other methods show quite low F1 scores with maximum values by

PAA, U-Net-DL, and ACA-AMDN to be 91.01 %, 96.75 %, and 98.46 % for Benign, 94.90 %, 96.75 %, and 97.33 % for Malignant and 88.20 %, 91.60 %, and 95.45 % for Normal. This is important because in breast cancer diagnosis, FPs and FNs can have significant consequences, and the F1-score helps mitigate these concerns.

The overall assessment of FNR and FPR for the breast cancer classification methods in the analysis of the CBIS-DDSM dataset is shown in Figs. 7(a) and 7(b) respectively. For PAA, the false negative rates for

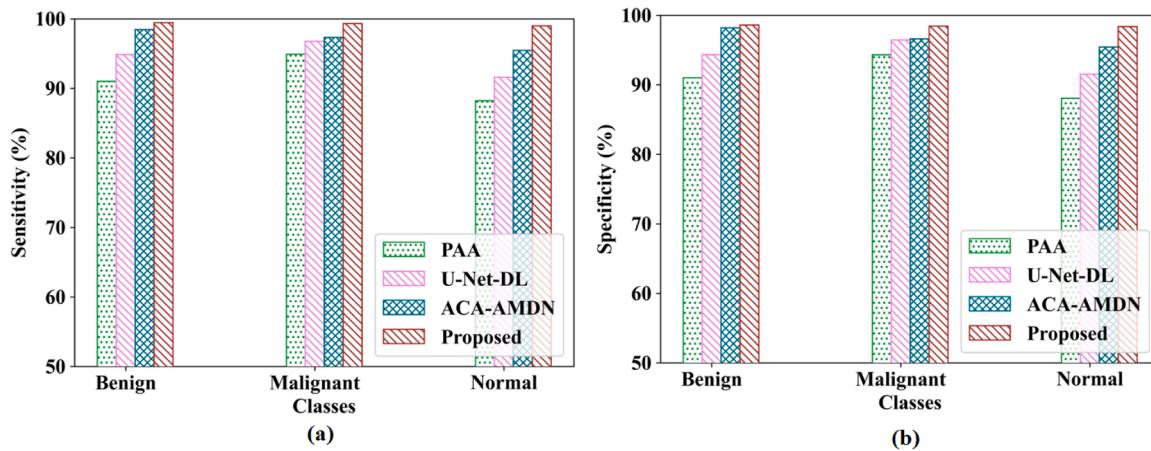


Fig. 9. Performance of (a) sensitivity (b) specificity analysis for CBIS-DDSM dataset.

Benign, Malignant, and Normal classes are 8.99 %, 5.10 %, and 11.80 %. For U-Net-DL, it is 5.15 %, 3.25 %, and 8.40 %. It is 1.54 %, 2.67 %, and 4.55 % for ACA-AMDN and 0.55 %, 0.69 %, and 1.01 % for the Proposed model. False Positive Rates are at 83.20 %, 81.14 %, 75.38 % for PAA, 85.92 %, 87.57 %, 81.27 % for U-Net-DL, 92.99 %, 90.31 %, 89.41 % for ACA-AMDN, and 97.99 %, 93.80 %, 95.05 % for the proposed model. Although the proposed model reveals the higher values of FPRs, it provides a lower value of FNR across all classes, demonstrating an improved performance in terms of a decrease in false negatives and, thus, raising the accuracy of disease detection. In clinical practice, FPs can be followed up with additional testing (e.g., biopsies or advanced imaging) to verify the diagnosis. On the other hand, missing a malignant case (FN) can have severe consequences for the patient's health. A lower FNR is highly desirable in breast cancer diagnosis, as it minimizes the risk of missed diagnoses, meaning fewer malignant tumors go undetected.

Fig. 8(a) and 8(b) illustrates the MCC and Precision performance of various models in breast cancer categorization respectively. For MCC, the scores for the Benign, Malignant, and Normal classes were 90.26 %, 94.07 %, and 87.74 % for the PAA model; 93.88 %, 96.13 %, and 90.81 % for the U-Net-DL model; 98.02 %, 97.20 %, and 94.53 % for the ACA-AMDN model; and 98.71 %, 99.15 %, and 98.66 % for the Proposed model. Similarly, in terms of Precision, the PAA model achieved 91.01 %, 94.90 %, and 88.20 % for the Benign, Malignant, and Normal classes, respectively, while the U-Net-DL model achieved 94.85 %, 96.75 %, and 91.60 % for the same classes. Meanwhile, the ACA-AMDN model recorded precisions of 98.46 %, 97.33 %, and 95.45 % for the respective classes. The Proposed model, however, demonstrated the highest Precision, achieving 99.45 %, 99.31 %, and 98.99 % for the Benign, Malignant, and Normal classes, respectively. This performance highlights the enhanced classification accuracy and reliability of the Proposed model across different breast cancer stages.

Figs. 9(a) and 9(b) shows the Sensitivity and Specificity metrics which represent how well the Proposed model is towards the breast cancer classification respectively. For Sensitivity, the scores are 99.45 %, 99.31 %, and 98.99 % respectively for the classes Benign, Malignant, and Normal, which are higher as compared to PAA where the values achieved are 91.01 %, 94.90 %, and 88.20 % respectively, U-Net-DL where the values achieved are 94.85 %, 96.75 %, and 91.60 % respectively, and by ACA-AMDN as 98.46 %, 97.33 %, and 95.45 % in respective classes of Benign, Malignant, and Normal. In clinical practice, the high sensitivity of the proposed Model ensures that malignant cases are less likely to be overlooked, leading to earlier interventions and potentially improving patient outcomes. In the case of Specificity, the Proposed model headed the chart with 98.58 %, 98.43 %, and 98.36 % for the Benign, Malignant, and Normal classes, respectively. PAA was at 90.99 %, 94.31 %, and 88.07 % while U-Net-DL was at 94.33 %, 96.44

%, and 91.53 %. ACA-AMDN was at 98.19 %, 96.61 %, and 95.43 %. These results indicate that the Proposed model outperforms in both True Positive detection as well as correct rejection for all classes of negatives. The high specificity of the Proposed Model reduces the risk of unnecessary follow-up tests, biopsies, or treatments for patients who do not have cancer. In a clinical setting, reducing FPs is important for minimizing patient stress, medical costs, and resource utilization.

Fig. 10 displays the analysis of the ROC. In comparison to existing techniques, such as PAA, U-Net-DL, and ACA-AMDN respectively, the proposed method offers a 1.25 %, 1.67 %, and 1.55 % higher AUC. The Proposed Model having higher AUC values means it is better at distinguishing between malignant and non-malignant cases across all possible thresholds. This suggests that the model performs more consistently and accurately, even when the decision threshold is adjusted. PAA, U-Net-DL, and ACA-AMDN, while still offering some level of discrimination, are slightly less effective at separating benign, malignant, and normal cases across the entire range of thresholds. The smaller AUCs indicate that these models might have a higher rate of misclassification at certain thresholds, leading to a less robust performance overall.

4.3. Simulation result of proposed model compared with existing techniques for BCS-DBT dataset

The results of the simulation using the recommended technique are shown in **Fig. 10–15**. Here, the efficacy of the suggested approach is contrasted with that of the state-of-the-art methods, such as RN50-CNN, T-DNN, and Mod_AlexNet.

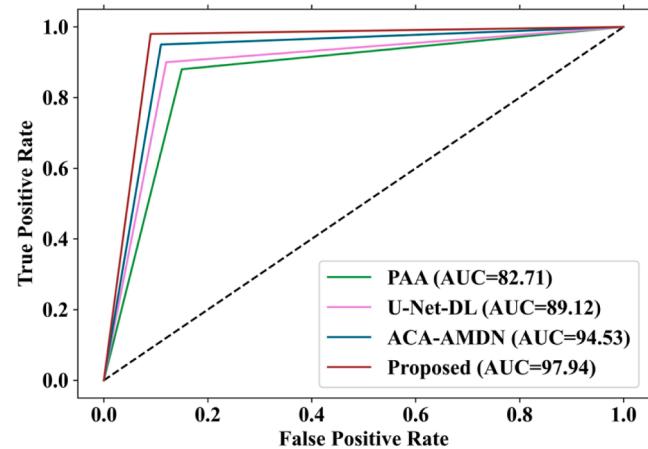


Fig. 10. ROC analysis for CBIS-DDSM dataset.

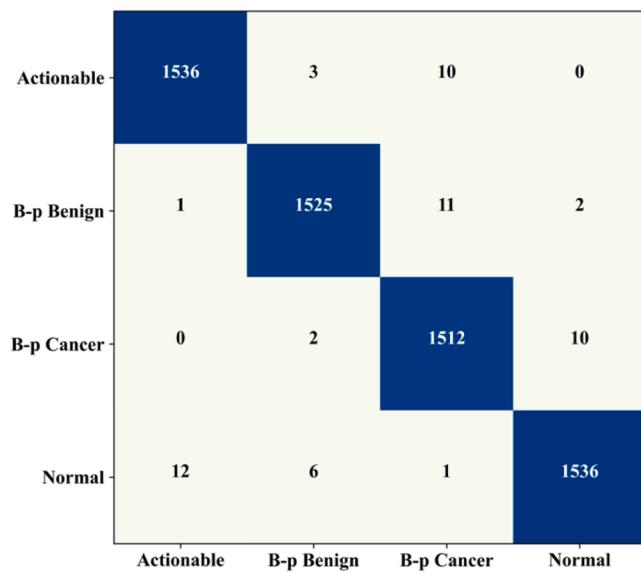


Fig. 11. Confusion matrix for classifying breast cancer using the BCS-DBT dataset.

Fig. 11 represents the confusion matrix for classifying breast cancer using the BCS-DBT dataset. This confusion matrix represents the result of a classification model trained on four classes, namely: Actionable,

Biopsy-proven benign (B-p Benign), Biopsy-proven cancer (B-p Cancer), and Normal. The matrix indicates very high accuracy by mostly falling with the instances in their respective categories to which they should belong. It depicts that 1536 cases are correctly classified as Actionable, whereas smaller misclassification occurs in the remaining three classes. By similar reasoning, the number of 1525 B-p Benign was very well recognized with very minor confusion with the rest. The B-p Cancer class is also very strong with 1512 correct recognitions and few errors. The Normal class obtained 1536 accurate classifications with some minor misclassifications as Actionable or other classes. This matrix represents that accuracy and recall values across all classes are very good except for slight errors to classes with overlapping features as well as a massive number of errors.

Figs. 12(a) and 12(b) show how each model performed in terms of accuracy and F1-Score for the four classes that were classified for the BCS-DBT Dataset respectively. The Proposed model is better than others in both two criteria. For Accuracy, the Proposed model scores 97.94 %, 98.25 %, 98.99 %, and 97.95 % across the four classes, while RN50-CNN trails with 93.05 %, 92.82 %, 94.24 %, and 92.47 %, and Mod_AlexNet and T-DNN have even lower scores. Under F1-Score, once again, the Proposed model leads with 98.96 %, 99.12 %, 99.50 %, and 98.96 %, significantly outperforming RN50-CNN, Mod_AlexNet, and T-DNN, which have relatively lower F1-Scores across all classes. These results indicate the better classification performance of the proposed model in detecting breast cancer. With consistent and reliable classification across the four classes, the Proposed Model can serve as an effective decision-support tool for clinicians, helping them prioritize cases that require urgent attention while not misclassifying benign or normal cases

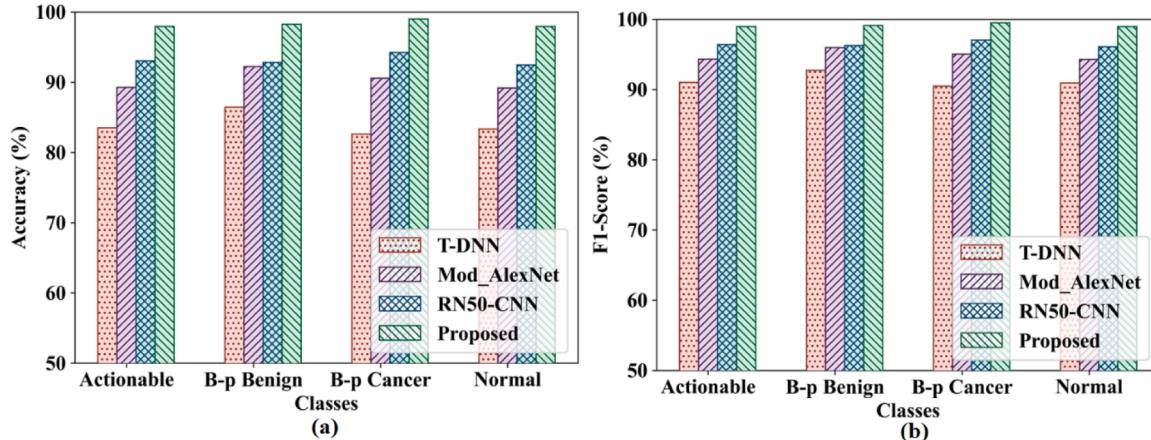


Fig. 12. Performance of (a) accuracy (b) F1-score analysis for BCS-DBT dataset.

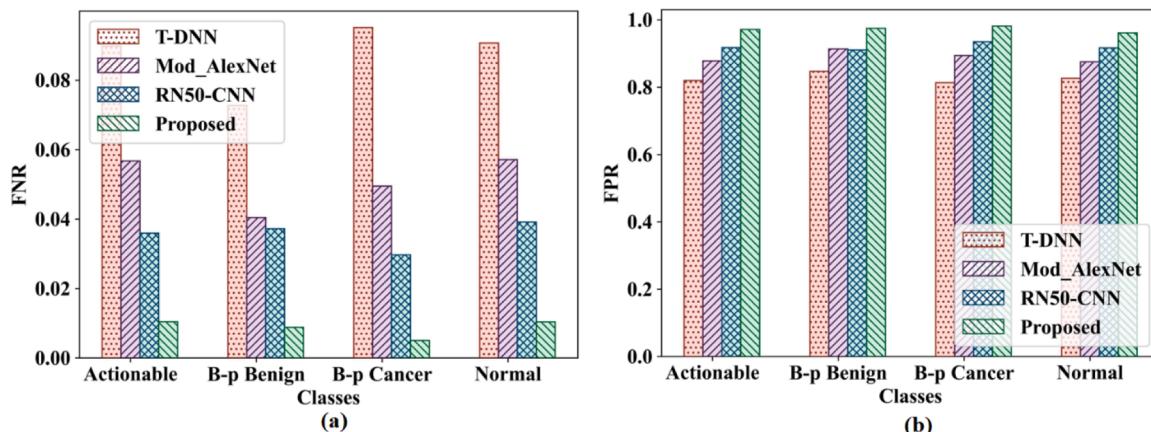


Fig. 13. Performance of (a) FNR (b) FPR analysis for BCS-DBT dataset.

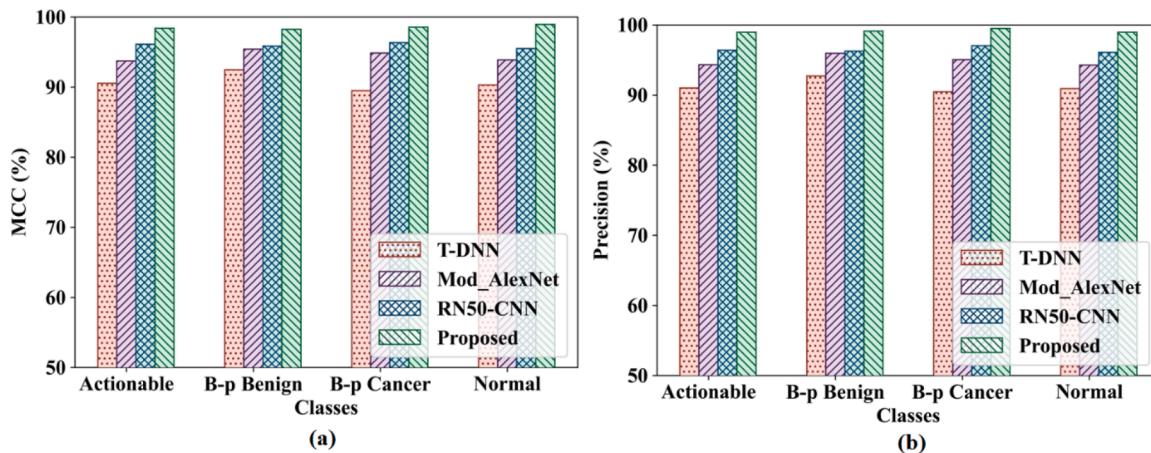


Fig. 14. Performance of (a) MCC (b) precision analysis for BCS-DBT dataset.

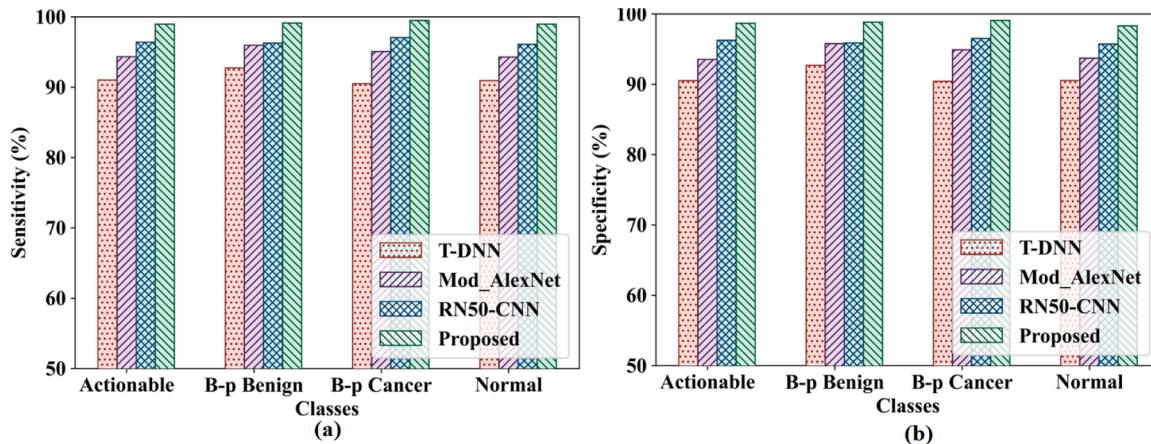


Fig. 15. Performance of (a) sensitivity (b) specificity analysis for BCS-DBT dataset.

as malignant

Figs. 13(a)13(b) show the performance of FNR and FPR analysis for the BCS-DBT dataset respectively. The FNR and FPR of classification models offer insight into error rates during the detection of breast cancers. The proposed model has the lowest FNR for all classes with values of 0.0104 for Actionable and 0.0050 for Biopsy-proven cancer, which indicates a better capability to correctly classify true positives. The more missed detection gets higher values in FNR with the FNR for RN50-CNN, Mod_AlexNet, T-DNN, and Proposed. The lowest error rates in FPR across the classes are also represented by the Proposed model with 0.9713 for Actionable and 0.9611 for Normal, meaning fewer false alarms. Other models, especially RN50-CNN and Mod_AlexNet, possess higher FPR values which show more wrong positive predictions. This would mean that the Proposed model does not only minimize false negatives but also reduces FPs; therefore, it classifies breast cancer more correctly. The proposed Model's low FNR ensures that fewer malignant tumors go undetected, which could potentially save lives by enabling earlier intervention. In contrast, its low FPR means that false alarms are reduced, preventing unnecessary biopsies or treatments and reducing the associated costs and risks.

Figs. 14(a) and 14(b) show the performance of MCC and precision analysis for the BCS-DBT dataset. It has achieved the highest MCC values for all classes: 98.38 % for Actionable, 98.24 % for Biopsy-proven benign, 98.53 % for Biopsy-proven cancer, and 98.94 % for Normal classes in the proposed model. The RN50-CNN model ranked next shows lower MCC values: 96.11 % for Actionable, 95.85 % for Biopsy-proven benign, 96.33 % for Biopsy-proven cancer, and 95.50 % for Normal.

Mod_AlexNet and T-DNN have lesser MCC values that denote a worse overall performance in accuracy. The proposed model achieves the highest precision values of 98.96 % for Actionable, 99.12 % for Biopsy-proven benign, 99.50 % for Biopsy-proven cancer, and 98.96 % for Normal. RN50-CNN takes the second position at 96.40 % for Actionable, 96.28 for Biopsy-proven benign, 97.03 % for Biopsy-proven cancer, and 96.09 % for Normal. Mod_AlexNet and T-DNN have low values for precision, meaning that misclassification happens more often. In comparison, the other models, though effective, show lower performance in both precision and MCC, which could lead to misclassifications, and ultimately affect clinical decision-making. Therefore, the proposed Model demonstrates a clear advantage in ensuring more accurate, reliable, and clinically relevant breast cancer detection.

Sensitivity and Specificity metrics performance for different models of classification for breast cancer using the BCS-DBT Dataset is shown in Figs. 15(a) and 15(b) respectively. The proposed model does its best to surpass all other models with values of 98.96 % for Actionable, 99.12 % for Biopsy-proven benign, 99.50 % for Biopsy-proven cancer, and 98.96 % for Normal. It is found superior in the effective identification of the correct positives in all classes. The sensitivity values for RN50-CNN are also good, that is, 96.40 % for Actionable, 96.28 % for Biopsy-proven benign, 97.03 % for Biopsy-proven cancer, and 96.09 % for Normal. Mod_AlexNet and T-DNN have very low sensitivity values, indicating a poor detection of positive cases compared to the proposed model. The proposed model also leads on specificity at 98.69 % for Actionable, 98.82 % for Biopsy-proven benign, 99.08 % for Biopsy-proven cancer, and 98.29 % for Normal. This reflects the model's high precision in

correctly classifying negative cases. RN50-CNN follows at 96.24 % for Actionable, 95.87 % for Biopsy-proven benign, 96.50 % for Biopsy-proven cancer, and 95.72 % for Normal. Mod_AlexNet and T-DNN models are less specific, meaning it has reduced precision in case identification of non-cancerous cases. The proposed model is highly sensitive and specific and presents better performance compared to the existing methods in the detection of both positive and negative cases of breast cancer.

Fig. 16 displays the analysis of the ROC. In comparison to existing techniques, such as Mod_AlexNet, T-DNN, and RN50-CNN respectively, the proposed method offers a 3.91 %, 2.15 %, and 1.11 % higher AUC. Higher AUC directly contributes to better clinical decision-making because it implies that the model is more likely to accurately identify cases that need attention and those that do not. A better AUC means fewer false negatives (missed cancer cases) and fewer FPs, which can lead to reduced patient anxiety and lower healthcare costs. Table 3 and 4 indicates the different performance Analysis under various conventional state-of-the-art techniques. The experimental outcomes prove that the developed method performs better while associated with other schemes. The proposed framework demonstrated notable improvements across key performance metrics compared to previous methods. Accuracy increased significantly, indicating a more reliable classification of both benign and malignant cases. Sensitivity (recall) improved, reflecting a higher true positive rate, which is critical for identifying malignant tumors accurately. Specificity also saw a marked enhancement, reducing false positives and ensuring that non-cancerous cases are correctly identified. Additionally, the Matthews Correlation Coefficient (MCC) showed substantial improvement, highlighting the balanced performance of the framework even with imbalanced datasets. The MCC is a robust metric that provides a balanced evaluation of the model's classification performance, particularly in cases with imbalanced datasets. It takes into account true positives, true negatives, false positives, and false negatives, offering a comprehensive single-value assessment of prediction quality. An MCC value close to +1 indicates a highly reliable model with strong agreement between predicted and actual classifications, while a value near 0 reflects random performance. In the context of medical classification, a high MCC demonstrates the model's ability to correctly identify specimens across all classes, ensuring both sensitivity and specificity are well-balanced. This makes MCC particularly valuable for evaluating the reliability of models in critical applications like tumor detection, where errors can have significant consequences.

4.4. Discussion

Detecting breast cancer presents a significant challenge in medical imaging, primarily due to the intricate characteristics of tumors and the diverse nature of breast tissue. Traditional mammography often faces

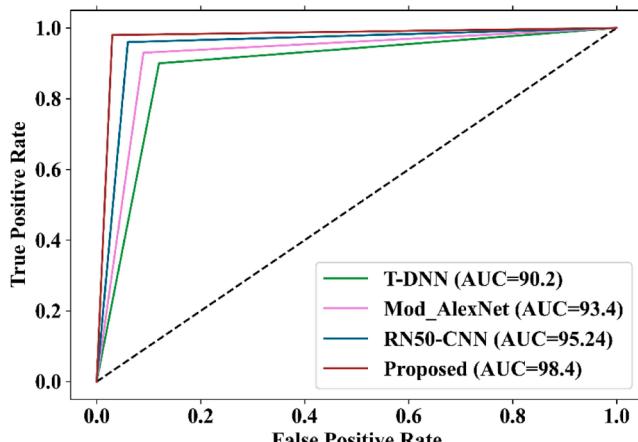


Fig. 16. Performance of ROC.

Table 3

Accuracy analysis under various conventional state-of-the-art techniques.

Author and References	Techniques Used	Image Considered	Accuracy Obtained
Kumar et.al., (2022) [19]	CNN	DBT	97.20 %
Jiang, et.al., (2022) [21]	PAA	Mammogram	98.32 %
Bouzar-Benlabiod, et.al., (2023) [22]	SE-ResNet-CBR	Mammogram	86.71 %
El-Shazli, et.al., (2022) [24]	Mod_AlexNet	DBT	97.67 %
Kumar et.al., (2024) [28]	EDC-APL-POA	DBT + Mammogram	96 %
proposed	FENN	DBT + Mammogram	98.64 %

Table 4

Performance analysis under various conventional state-of-the-art techniques.

Techniques Used	Obtained Performance	Time Consumed (Training and Testing)
VGG-16 [37]	Precision-0.760	250ms
ResNet [38]	Accuracy-97.50 %	102s
MobileNet [39]	Accuracy-98 %	180s
Google inception network [40]	Recall-96.48 %	55ms
Proposed (FENN)	Accuracy-98.64 %, specificity-98.59 %, sensitivity-99.19 %	32ms

difficulties in identifying abnormalities within dense breast tissues, which can result in undetected cases. Hence, mammogram images from the CBIS-DDSM dataset and DBT images from the BCS-DBT dataset were used. The preprocessing step involved Ext-AFF to enhance image quality by reducing noise. Feature extraction was performed using d-VAE, capturing critical texture features. These features were fused using Dg-CCA to maximize feature correlation across modalities. Finally, a Fully Elman Neural Network was employed for classification, distinguishing between benign, malignant, biopsy-proven cancer, and normal tissues. The use of d-VAE for feature extraction and Dg-CCA for feature fusion is novel, with an emphasis on robustness in capturing the complex interdependencies from multi-modal imaging. d-VAE's capability to disentangle important features while keeping irrelevant noise to a minimum aligns well with the necessities of breast cancer diagnosis where subtle differences between benign and malignant tissues are important. Similarly, Dg-CCA's ability to harmonize and maximize feature correlations from mammogram and DBT images provides a cohesive representation, reducing data dimensionality while preserving critical information. The superior performance of the model in breast cancer diagnosis translates to several key benefits for both patients and clinicians. For patients, it means more accurate, earlier detection of cancer, leading to better treatment outcomes and reduced anxiety from unnecessary procedures. Clinicians benefit from improved diagnostic confidence, enabling faster, more reliable decision-making. This enhances workflow efficiency, reduces the need for follow-up tests, and ultimately leads to cost savings, while ensuring higher quality care and reduced patient harm. This framework, combining mammography and DBT, has the potential to significantly reduce misdiagnosis in dense breast tissues. Dense breasts can obscure tumors in traditional 2D mammograms, leading to missed cancers or false negatives. By integrating DBT's 3D imaging, the framework provides clearer visualization of lesions, improving detection accuracy. This reduction in misdiagnosis could lead to earlier cancer identification, better patient outcomes, and less need for unnecessary biopsies or follow-up procedures. Ultimately, it enhances diagnostic reliability and minimizes the risks associated with dense breast tissue. Translating this system to a clinical setting may face challenges such as the integration of advanced imaging techniques into existing workflows, increased computational demands, and the need for

clinician training. Overcoming these challenges requires seamless integration of the system with current radiology platforms, optimizing algorithms for faster processing, and providing educational resources to ensure clinicians can interpret the enhanced images effectively. Additionally, standardizing the system across different healthcare settings and ensuring its cost-effectiveness will be crucial for widespread adoption and use.

5. Conclusion

In this manuscript, we present a novel framework for breast cancer detection that integrates deep feature fusion techniques with the architecture of a Fully Elman Neural Network (FENN). The proposed system utilizes mammogram images from the CBIS-DDSM dataset and DBT images from the Breast-Cancer-Screening-DBT dataset to enhance diagnostic accuracy through a multi-step approach. The Ext-AFF improves image quality, while the D-VAE optimizes feature extraction from both texture and intensity data. The Dg-CCA further amplifies the system's discriminative power by fusing these features. The FENN-based model successfully classifies images such as benign, malignant, biopsy-proven cancer, and normal, demonstrating superior handling of complex data relationships. Overall performance metrics, including accuracy, sensitivity, and specificity show that the proposed framework outperforms existing state-of-the-art methods, with an accuracy of 98.64 %, specificity of 98.59 %, and a sensitivity of 99.19 %. These results highlight the potential of the FENN-based framework as a reliable tool for early breast cancer detection, supporting clinical decision-making, and improving patient outcomes. The confusion matrix reveals the efficiency of the model with very less misclassifications and its effectiveness on the three classes. Secondly, the high F1-scores of the proposed system testifies the accuracy in distinguishing between benign and malignant classes. Thirdly, a lower FNR value across all classes demonstrates an improved performance in terms of a decrease in false negatives. Fourthly, the MCC and precision analysis highlights the reliability of the proposed model. We also notice the high sensitivity and specific of the model that leads to correct rejection of negatives, as well as reduces the unnecessary headache of follow-up tests and treatments. Finally, the ROC analysis suggests that the model is better in distinguishing between malignant and non-malignant cases across all possible thresholds because of its higher AUC value. Future work will be placed on the real-time in-clinic deployment of the proposed framework. The plan is to make a complete system that can process mammography and DBT images with the least latency to enable instant diagnoses and feedback in the clinical space. This will involve computational requirements in real-world settings by optimizing for resource-limited settings and lightweight neural network architectures to incorporate hardware acceleration techniques. More diversely, there will be a focus on incorporating images from various sizes of breasts, dense tumor breasts, and multi-tumor cases to enhance generalization. It will be very helpful to apply hybrid or multi-scale methods of advanced feature extraction techniques for extracting finer details in mammograms and DBT images. Alternative architectures and possibly addressing dataset imbalance could become important to be as accurate and robust in such a model. So advanced data augmentation and also possibly class weighting might emerge from this study to enhance such architectures to make the framework applicable, reliable, and scalable for wide-scale clinical implementations which would support early detection toward improved patient outcomes.

CRediT authorship contribution statement

Nishu Gupta: Resources, Investigation, Conceptualization. **Jan Kubicek:** Writing – original draft, Visualization, Validation. **Marek Penhaker:** Writing – review & editing, Validation, Project administration. **Mohammad Derawi:** Writing – review & editing, Supervision, Data curation.

Declaration of competing interest

All the authors of the manuscript declare no conflict of interest.

Acknowledgments

This article has been produced with the financial support of the European Union under the LERCO project number CZ.10.03.01/00/22_003/0000003 via the Operational Programme Just Transition. The work and the contributions were supported by the project SP2024/071 'Biomedical Engineering systems XX'.

Data availability

Data will be made available on request.

References

- [1] U. Sajid, R.A. Khan, S.M. Shah, S. Arif, Breast cancer classification using deep learned features boosted with handcrafted features, *Biomed. Signal Process. Control* 86 (2023) 105353.
- [2] H.S. Das, A. Das, A. Neog, S. Mallik, K. Bora, Z. Zhao, Breast cancer detection: shallow convolutional neural network against deep convolutional neural networks based approach, *Front. Genet.* 13 (2023) 1097207.
- [3] M. Alkhaleefah, T.H. Tan, C.H. Chang, T.C. Wang, S.C. Ma, L. Chang, Y.L. Chang, Connected-segNets: a deep learning model for breast tumor segmentation from X-ray images, *Cancers* 14 (16) (2022) 4030.
- [4] A.A. Hekal, H.E.D. Moustafa, A. Elnakib, Ensemble deep learning system for early breast cancer detection, *Evol. Intell.* 16 (3) (2023) 1045–1054.
- [5] G.I. Quintana, Z. Li, L. Vancambert, M. Mougeot, A. Desolneux, S. Muller, Exploiting patch sizes and resolutions for multi-scale deep learning in mammogram image classification, *Bioengineering* 10 (5) (2023) 534.
- [6] G.M. El-Banby, N.S. Salem, E.A. Tafweek, E.N.A. El-Azziz, Automated abnormalities detection in mammography using deep learning, *Complex Intell. Syst.* (2024) 1–17.
- [7] S. Zahoor, U. Shoaib, I.U. Lali, Breast cancer mammograms classification using deep neural network and entropy-controlled whale optimization algorithm, *Diagnostics* 12 (2) (2022) 557.
- [8] Z. Jafari, E. Karami, Breast cancer detection in mammography images: a CNN-based approach with feature selection, *Information* 14 (7) (2023) 410.
- [9] M.S. Nazir, et al., A novel CNN-Inception-V4-based hybrid approach for classification of breast cancer in mammogram images, *Wireless Commun. Mobile Comput.* 2022 (2022) 5089078.
- [10] M.A.K. Raiaan, N.M. Fahad, M.S.H. Mukta, S. Shatabda, Mammo-Light: a lightweight convolutional neural network for diagnosing breast cancer from mammography images, *Biomed. Signal Process. Control* 94 (2024) 106279.
- [11] S. Maqsood, R. Damaševicius, R. Maskeliūnas, TTCNN: a breast cancer detection and classification towards computer-aided diagnosis using digital mammography in early stages, *Appl. Sci.* 12 (7) (2022) 3273.
- [12] G. Toz, A simple method for obtaining artificial 3D forms of 2D mammograms in diagnosis of breast cancer, *Imaging Sci. J.* (2023) 1–16.
- [13] N. Mobark, S. Hamad, S.Z. Rida, CoroNet: deep neural network-based end-to-end training for breast cancer diagnosis, *Appl. Sci.* 12 (14) (2022) 7080.
- [14] Y. Zhou, J. Wei, D. Wu, Y. Zhang, Generating full-field digital mammogram from digitized screen-film mammogram for breast cancer screening with high-resolution generative adversarial network, *Front. Oncol.* 12 (2022) 868257.
- [15] K.L. dos Santos, M.P. dos Santos Silva, Deep cross-training: an approach to improve deep neural network classification on mammographic images, *Expert Syst. Appl.* 238 (2024) 122142.
- [16] R. Agarwal, M.H. Yap, M.K. Hasan, R. Zwiggelaar, R. Martí, Deep learning in mammography breast cancer detection, *Artificial Intelligence in Medicine*, Springer International Publishing, Cham, 2022, pp. 1287–1300.
- [17] V. Priyadarshni, S.K. Sharma, M.K.I. Rahmani, B. Kaushik, R. Almajlid, Machine learning techniques using deep instinctive encoder-based feature extraction for optimized breast cancer detection, *Comput. Mater. Continua* 78 (2) (2024).
- [18] Y. Jiménez-Gaona, M.J. Rodríguez-Álvarez, D. Carrión-Figueroa, D. Castillo-Malla, V. Lakshminarayanan, Breast mass regions classification from mammograms using convolutional neural networks and transfer learning, *J. Mod. Opt.* 70 (10) (2023) 645–660.
- [19] P. Kumar, S. Srivastava, R.K. Mishra, Y.P. Sai, End-to-end improved convolutional neural network model for breast cancer detection using mammographic data, *J. Defense Model. Simul.* 19 (3) (2022) 375–384.
- [20] M.A.K. Raiaan, N.M. Fahad, M.S.H. Mukta, S. Shatabda, Mammo-light: a lightweight convolutional neural network for diagnosing breast cancer from mammography images, *Biomed. Signal Process. Control* 94 (2024) 106279.
- [21] J. Jiang, J. Peng, C. Hu, W. Jian, X. Wang, W. Liu, Breast cancer detection and classification in mammogram using a three-stage deep learning framework based on PAA algorithm, *Artif. Intell. Med.* 134 (2022) 102419.
- [22] L. Bouzar-Benlaioud, K. Harrar, L. Yamoun, M.Y. Khodja, M.A. Akhloufi, A novel breast cancer detection architecture based on a CNN-CBR system for mammogram classification, *Comput. Biol. Med.* 163 (2023) 107133.

- [23] M. Yaqub, et al., Intelligent breast cancer diagnosis with two-stage using mammogram images, *Sci. Rep.* 14 (1) (2024) 16672.
- [24] A.M.A. El-Shazli, S.M. Youssef, A.H. Soliman, Intelligent computer-aided model for efficient diagnosis of digital breast tomosynthesis 3D imaging using deep learning, *Appl. Sci.* 12 (11) (2022) 5736.
- [25] W. Lee, et al., Transformer-based deep neural network for breast cancer classification on digital breast tomosynthesis images, *Radiol. Artif. Intell.* 5 (3) (2023) e220159.
- [26] L. Hassan, A. Saleh, V.K. Singh, D. Puig, M. Abdel-Nasser, Detecting breast tumors in tomosynthesis images utilizing deep learning-based dynamic ensemble approach, *Comput* 12 (11) (2023) 220.
- [27] K. Oba, et al., Deep learning model to predict Ki-67 expression of breast cancer using digital breast tomosynthesis, *Breast Cancer* (2024) 1–7.
- [28] M.N.V.S.S. Kumar, D. Eswara Chaitanya, L. Ganesh, T.Sri Sudha, An automated breast cancer detection by heuristic-based ensemble deep classifier using mammogram and tomosynthesis images, *Biomed. Eng.: Appl., Basis Commun.* 36 (02) (2024) 2450001.
- [29] I. Kassis, D. Lederman, G. Ben-Arie, M.G. Rosenthal, I. Shelef, Y. Zigel, Detection of breast cancer in digital breast tomosynthesis with vision transformers, *Sci Rep* 14 (1) (2024) 22149.
- [30] S. Franceschini, M.M. Autorino, M. Ambrosanio, V. Pascazio, F. Baselice, A deep learning approach for diagnosis support in breast cancer microwave tomography, *Diagnostics* 13 (10) (2023) 1693.
- [31] <https://www.kaggle.com/datasets/awsa49/cbis-ddsm-breast-cancer-image-dataset>.
- [32] <https://www.cancerimagingarchive.net/collection/breast-cancer-screening-dbt/>.
- [33] A. Gupta, S. Kumar Yadav, C.A.D. Durai, V. Kuamr, M.H. Alsharif, P. Uthansakul, M. Uthansakul, Enhanced breast tumor localization with DRA antenna backscattering and GPR algorithm in microwave imaging, *Results Eng.* (2024) 24.
- [34] A. Staffini, T. Svensson, U.I. Chung, A.K. Svensson, A disentangled VAE-BiLSTM model for heart rate anomaly detection, *Bioengineering* 10 (6) (Jun. 2023) 683, <https://doi.org/10.1016/j.rineng.2024.103044>.
- [35] H. Yan, L. Cheng, Q. Ye, D.J. Yu, Y. Qi, Robust generalized canonical correlation analysis, *Appl. Intell.* 53 (18) (2023) 21140–21155.
- [36] M. Fetanat, M. Stevens, P. Jain, C. Hayward, E. Meijering, N.H. Lovell, Fully Elman neural network: a novel deep recurrent neural network optimized by an improved Harris Hawks algorithm for classification of pulmonary arterial wedge pressure, *IEEE Trans. Biomed. Eng.* 69 (5) (May 2021) 1733–1744.
- [37] A. Sahu, P.K. Das, S. Meher, Recent advancements in machine learning and deep learning-based breast cancer detection using mammograms, *Phys. Med.* 114 (2023) 103138.
- [38] K.L. dos Santos, M.P. dos Santos Silva, Deep cross-training: an approach to improve deep neural network classification on mammographic images, *Expert. Syst. Appl.* 238 (2024) 122142.
- [39] I. Nissar, S. Alam, S. Masood, M. Kashif, MOB-CBAM: a dual-channel attention-based deep learning generalizable model for breast cancer molecular subtypes prediction using mammograms, *Comput. Methods Programs Biomed.* 248 (2024) 108121.
- [40] W.S. Admass, Y.Y. Munaye, A.O. Salau, Integration of feature enhancement technique in Google inception network for breast cancer detection and classification, *J. Big Data* 11 (1) (2024) 78.