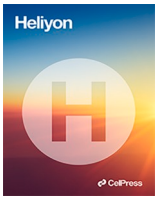
Heliyon 10 (2024) e38374

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Hybrid ensemble deep learning model for advancing breast cancer detection and classification in clinical applications

Radwan Qasrawi [a,b,*], Omar Daraghmeh [c], Ibrahem Qdaih [c], Suliman Thwib [a], Stephanny Vicuna Polo [f], Haneen Owienah [d], Diala Abu Al-Halawa [e], Siham Atari [a]

[a] Department of Computer Science, Al-Quds University, Palestine
[b] Department of Computer Engineering, Istinye University, Istanbul, Turkey
[c] Department of Medical Imaging, Al-Quds University, Jerusalem, Palestine
[d] Department of Radiology, Istishari Arab Hospital, Palestine
[e] Faculty of Medicine, Al-Quds University, Palestine
[f] Al Quds Business Center for Innovation, Technology, and Entrepreneurship, Al Quds University, Jerusalem, Palestine

ARTICLE INFO

Keywords:
Breast Cancer$_1$
Diagnostic Accuracy$_2$
Deep Learning$_3$
Mammographic Images$_4$
Image Enhancement$_5$

ABSTRACT

Being the most common type of cancer worldwide, and affecting over 2.3 million women, breast cancer poses a significant health threat. Although survival rates have improved around the world due to advances in screening, diagnosis, and treatment, early detection remains crucial for effective management. This study seeks to introduce a novel hybrid model that makes use of image-preprocessing techniques and deep-learning algorithms on mammograms to enhance the detection and classification accuracy of breast cancer lesions. The model was tested on a dataset comprising 20,000 mammograms. First, image-processing techniques, such as Contrast-Limited Adaptive Histogram Equalization, Gaussian Blur, and sharpening methods were used to optimize the images for enhanced feature extraction. In addition, the Ensemble Deep Random Vector-Functional Link Neural Network algorithm, YOLOv5, and MedSAM segmentation models were utilized for robust deep learning-based extraction, classification, and visualization of lesions. Finally, the model was clinically validated on 800 patients. The study found a notable enhancement in both accuracy and processing time for benign and malignant diagnoses using the hybrid model. The model achieves an impressive accuracy of 99.7 % and demonstrates a remarkable processing time of 0.75 s. In clinical applications, the hybrid model exhibits high proficiency, reporting 97.2 % accuracy for benign cases and 98.6 % for malignant scenarios. These results highlight the effectiveness of the hybrid model in improving diagnostic accuracy, offering a promising tool for early breast cancer detection.

1. Introduction

Breast cancer is the most common type of cancer worldwide and a leading cause of cancer-related deaths among women, accounting for over 12.5 % of overall cancer cases [1]. Early detection and diagnosis are crucial for successful treatment, improved patient outcomes, and reduced mortality rates. Currently, mammograms are the standard imaging test for breast cancer screening. However, mammograms can be difficult to interpret, particularly for inexperienced radiologists [2,3], as they present overlapping

* Corresponding author. Ramallah, Palestine.
E-mail address: radwan@staff.alquds.edu (R. Qasrawi).

https://doi.org/10.1016/j.heliyon.2024.e38374
Received 30 March 2024; Received in revised form 9 September 2024; Accepted 23 September 2024
Available online 24 September 2024

Abbreviations:

AHE	Adaptive Histogram Equalization
AUC	Area under the receiver operating characteristic curve
B-H	Benign predictions
BI-RADS	Breast-Imaging Reporting and Data System
CAD	Computer-Aided Diagnosis
CDF	Cumulative distribution function
CL	Clip limit
CLAHE	Contrast-Limited Adaptive Histogram Equalization
CNN	Convolutional Neural Networks
DBN	Deep Belief Network
DICOM	Digital Imaging and Communications in Medicine
DSC	Dice Similarity Coefficient
DT	Decision Tree
edRVFL	Ensemble deep Random Vector Functional Link Neural Network
EME	Effective measure of enhancement
GB	Gradient Boosting
JPEG	Joint Photographic Experts Group
LR	Logistic Regression
LoG	Laplacian of Gaussian
LSTM	Long short-term memory
MedSAM	Segment Anything in Medical Images
M-H	Malign predictions
ML	Machine learning
MSE	Mean squared error
PSNR	Peak signal-to-noise ratio
RF	Random Forest
RVFL	Random Vector Functional Link
SGD	Scholastic Gradient Descent
SVM	Support Vector Machine
US	Ultrasound
YOLOv5	You Only Look Once

tissues, varying densities, and variable contrast, which can hinder the differentiation between benign and malignant lesions [4]. Consequently, false positives, leading to unnecessary anxiety and interventions, or false negatives, resulting in delayed treatment, are not uncommon.

Recent advancements in image processing and analysis, as well as Artificial Intelligence (AI) technologies, such as image enhancement and deep learning, have the potential to improve the accuracy and efficiency of mammogram interpretation [5]. Image enhancement techniques, known for their ability to improve image quality, have played a pivotal role in mammogram analysis [4,6,7]. Dhungel et al. (2015) leveraged Adaptive Histogram Equalization (AHE) techniques, demonstrating that such enhancements can significantly improve the detection of breast masses [8]. By integrating image enhancement with deep learning, some studies have managed to improve the limitations of mammogram analysis. Integrating image enhancement with deep learning has further addressed limitations in mammogram analysis. For instance, Arefan et al. (2020) found that a deep learning model using mammograms achieved an Area Under the Receiver Operating Characteristic Curve (AUC) of 0.73 when analyzing the whole breast and 0.72 when focusing on dense tissue [9]. Similarly, the use of Contrast-limited Adaptive Histogram Equalization (CLAHE) combined with Gaussian filtering and Convolutional Neural Networks (CNNs) has shown promising performance metrics [10,11].

The emergence of deep learning has revolutionized mammogram analysis by enabling algorithms to recognize subtle patterns in mammograms that may be difficult for human observers to detect. Several studies have explored the use of Machine Learning (ML) and deep learning methods for breast cancer detection and classification from mammogram images, achieving high levels of accuracy, sometimes even surpassing human radiologists [1,12–15]. While traditional ML algorithms like Support Vector Machine (SVM) have shown accuracy levels similar to conventional methods [16], hybrid models that combine multiple algorithms, such as Decision Tree (DT) and Random Forest (RF), have produced even more promising results [17].

The hybrid model approach has recently gained traction as a means of bridging the gap between machine learning and deep learning [6,7,18]. For example, Zhang et al. (2019) combined SVM with CNNs, resulting in improved specificity and sensitivity in breast cancer detection [19]. There has also been significant interest in transfer learning within mammography. In 2021, Liu's study utilized a hybrid deep learning framework that integrated gene modality data with imaging modality data to predict breast cancer subtypes, achieving higher accuracy than conventional deep-learning approaches [20]. Other studies have further explored hybrid models for breast cancer detection, such as Jayandhi's integration of MobileNet and Long Short-Term Memory (LSTM) to classify

breast cancer in mammographic images [21], and Narayanan's 2022 research, which proposed a sophisticated hybrid system combining deep CNNs with RF classifiers for breast cancer detection and classification within mammogram images [22].

Recent advancements in medical data analysis increasingly emphasize the integration of diverse methodologies to enhance healthcare outcomes. For instance, a 2023 study by Yan et al. [23] highlighted the potential of combining fuzzy methods with optimization algorithms and machine learning for medical record analysis. This approach underscores the growing importance of hybrid methods in processing complex healthcare data, aligning with broader trends in the field.

Collectively, these studies emphasize the transformative potential of hybrid deep learning techniques to enhance breast cancer detection and forecasting. Despite promising results, numerous challenges remain before these methods can be fully applied in clinical settings. These challenges include dealing with unclear and low-quality images, dense breast tissue that can obscure cancer detection, and variability in image interpretation. Additional difficulties arise from imaging equipment, techniques, and patient positioning, leading to poor image quality that can confuse or distort tissue structures. This study aims to assess whether a hybrid model utilizing CLAHE image preprocessing combined with an optimized ensemble deep Random Vector-Functional Link Neural Network (edRVFL) can effectively enhance breast cancer detection and classification through mammograms in a clinical setting.

The rest of this paper is organized as follows: The literature review section provides a comprehensive discussion of existing methods and advancements in breast cancer detection using hybrid models. The materials and methods section details the dataset, data preprocessing techniques, and the use of edRVFL, including its hyperparameter optimization for mammogram classification. The development of the hybrid model is further elaborated with a focus on lesion detection and segmentation, ensemble deep-learning classification, and testing within a clinical setting. The results section presents the outcomes of the hyperparameter optimization for the edRVFL, followed by an evaluation of the classification model's performance before and after image enhancement, as well as its application in clinical settings. Additionally, the performance of the segmentation model in clinical applications is assessed, accompanied by a Grad-CAM analysis. The discussion section interprets the findings, while the conclusion section summarizes the study's contributions and suggests potential future research directions.

2. Literature review

The field of breast cancer detection using machine learning and computer vision techniques has seen remarkable advancements over the past decade, with researchers continuously pushing the boundaries of accuracy and reliability. Early studies in the 2010s laid the foundation with traditional machine learning approaches. De Nazaré Silva et al. (2015) made significant strides by utilizing wavelet transform enhancement and Support Vector Machine (SVM) classification, demonstrating strong diagnostic performance with sensitivity, specificity, and accuracy rates of 92.31 %, 82.2 %, and 83.53 %, respectively [24,25].

This study was pivotal in showcasing the potential of combining sophisticated image enhancement techniques with machine learning classifiers, setting the stage for future innovations. As computational power increased and deep learning methodologies matured, the field witnessed a paradigm shift. Dhungel et al. (2017) introduced the use of deep belief networks (DBNs) and Convolutional Neural Networks (CNNs) for mass detection, segmentation, and classification, achieving high accuracy [26]. This marked a significant transition towards end-to-end learning approaches that could automatically extract relevant features from raw image data, reducing the reliance on hand-crafted features and potentially capturing more nuanced patterns in mammographic images.

The integration of multiple deep learning techniques ushered in a new era of performance improvements. Mugahed A Al-antari et al. (2018) proposed an integrated Computer-Aided Diagnosis (CAD) system that exemplified this trend. Their approach combined You Only Look Once (YOLO) for detection, Full Resolution Convolutional Networks (FrCN) for segmentation, and CNN for classification of breast masses, achieving impressive detection, segmentation, and classification accuracies of 98.96 %, 92.97 %, and 95.64 %, respectively [27]. This comprehensive approach demonstrated the power of leveraging specialized deep learning architectures for different stages of the breast cancer detection pipeline. Building on this momentum, Nasir Khan et al. (2019) introduced a multi-view feature fusion CNN model, achieving an Area Under the Curve (AUC) of 0.932 for mass and calcification classification and 0.84 for malignant and benign classification [28]. In the same year, Yu X et al. employed DenseNet201 for mammogram classification, reaching a diagnostic accuracy of 92.73 % [29]. These studies underscored the potential of specialized CNN architectures and multi-view approaches in improving classification performance, particularly in handling the complex and varied presentations of breast abnormalities in mammographic images.

Recent years have witnessed a trend towards ensemble methods and hybrid approaches, aiming to combine the strengths of different algorithms. Song R et al. (2020) exemplified this trend by combining deep CNN with XGBoost and SVM classifiers, achieving an overall accuracy of 92.80 % [30]. Gokhan Altan's 2020 study employed a CNN for mammogram classification, reaching impressive accuracy, sensitivity, specificity, and precision rates of 92.84 %, 95.30 %, 96.72 %, and 96.72 %, respectively [31]. Altan et al. continued to refine their approach, and in 2021, they utilized DBNs to classify regions of interest (ROI) in mammogram images, further improving their results with accuracy, specificity, sensitivity, and precision rates of 96.32 %, 96.68 %, 95.93 %, and 96.40 %, respectively [32]. This progression in Altan's work demonstrates the iterative nature of improvements in the field, with researchers continuously refining their methods to achieve incremental gains in performance. In the same year, Ragab et al. (2021) explored the combination of features from pre-trained CNNs like AlexNet and ResNet with SVMs for classification, demonstrating improved detection [33]. This approach highlighted the potential of transfer learning and feature fusion in enhancing breast cancer detection algorithms.

The most recent studies have continued to push the boundaries of accuracy. In 2022, Lakshmi Narayanan et al. combined AlexNet and Random Forest (RF) classifier, achieving a remarkable accuracy of 98.3 %, marking significant performance improvements over existing methods [34]. This study underscores the ongoing potential for innovation in the field, even as performance metrics approach

near-perfect levels.

The study by Yan et al. [35] proposes an automated methodology for breast cancer detection in mammography images using an ensemble classifier and feature weighting algorithms. Their comprehensive approach includes novel techniques for region extraction, pectoral muscle removal, and ROI segmentation. By fusing and weighting textural and morphological features, and employing an ensemble classifier model (KNN, bagging, and EigenClass), they achieved average accuracies of 93.26 % and 91.00 % on the DDSM and MIAS datasets, respectively. This study represents a holistic approach that addresses multiple aspects of the breast cancer detection pipeline, from preprocessing to final classification.

This chronological review reveals a clear trend towards more sophisticated and integrated approaches, combining multiple techniques to improve overall performance. The evolution from traditional machine learning methods to deep learning, and then to ensemble and hybrid models, has consistently pushed the boundaries of accuracy and reliability in breast cancer detection. Despite these advancements, there remains room for improvement, particularly in reducing false positive rates, enhancing the interpretability of deep learning models, and ensuring robust performance across diverse patient populations and imaging conditions. The motivation for ongoing research in this field, including the current study, stems from the need to address these challenges while building upon the successes of previous approaches. By proposing novel methods that integrate the strengths of various techniques, current and future studies aim to contribute to the ongoing efforts to improve breast cancer detection accuracy and reliability, ultimately leading to better patient outcomes and more efficient healthcare processes.

3. Materials and methods

3.1. Dataset

This study gathered data from two distinct sources. The first source was the publicly available VinDr-Mammo collection (available at www.vindr.ai/datasets/mammo) [36], which comprises mammography examinations taken between 2018 and 2020. These examinations were stored in the Picture Archiving and Communication Systems (PACS) of Hanoi Medical University Hospital (HMUH) and Hospital 108 (H108). From this pool, 5000 four-view mammography exams (equivalent to 20,000 images) were randomly sampled. The images were then de-identified, formatted in JPEG, and standardized to a resolution of 640x640 pixels, forming the foundation for training and evaluating our hybrid diagnostic model. Table 1 presents the statistics for breast-level BI-RADS assessments.

The predictive framework was deployed across two distinct sets of data (Dataset 1: Benign-Normal group; Dataset 2: Malignant-Normal group). We initiated the computational process by first splitting the original dataset into three subsets: 70 % for training, 20 % for testing, and 10 % for validation. To counter class disparity and enrich the dataset's diversity, we then applied data augmentation exclusively to the training set. This expansion involved various transformations of the original images in the training set, including rotation at multiple angles (90°, 180°, and 270°), adjustments in width and height, rescaling, shearing, zooming, horizontal flip operations, and specific filling techniques. This pivotal step expanded the training data pool, resulting in approximately 56,000 images in the training set, thereby enhancing the generalization performance of the model.

The strategic application of data augmentation to only the training set, combined with systematic 25-fold cross-validation, significantly mitigated the risk of overfitting in the learning model. The test and validation sets remained the same, preserving their integrity for accurate model evaluation. Furthermore, the incorporation of dropout layers into the model's architecture provided additional protection against overfitting, enhancing the model's robustness and generalization capabilities. This approach ensured that our model was trained on a diverse range of augmented data while being evaluated on original, non-augmented data, leading to more reliable performance assessments.

The second source was a specialized clinical dataset originating from the Dunya Women's Cancer Center in Palestine, with data obtained in 2023. Mammograms were acquired using the MAMMOMAT Revelation equipment, adopting the Craniocaudal (CC) and Mediolateral Oblique (MLO) mammographic orientations (KVP:20-40—25/26, MAS: 70–120). Expert health professionals, including medical doctors and radiologists, were responsible for identifying the lesions and determining their specific boundaries. Based on the criteria outlined by the BI-RADS [37], these images were categorized into three groups: normal, benign, and malignant. This categorization was subsequently confirmed by comparing it with histopathology findings. This dataset encompasses 800 mammography exams, equivalent to 1888 images, derived during the model's real-world application in a clinical setting. These real patient data points were involved in comparing the model's outcomes with histopathological findings, offering a precise assessment of accuracy. This comparison not only confirmed the accuracy of the model's predictions but also emphasized its applicability and relevance in a practical medical environment.

The distribution of patient ages in both datasets provides valuable insight into the demographic variability of the study population.

Table 1
Statistics of breast-level BI-RADS assessment.

Breast BI-RADS						
	1	**2**	**3**	**4**	**5**	**Total**
Training	10,724 (67.03 %)	3742 (23.39 %)	744 (04.65 %)	610 (03.81 %)	180 (01.13 %)	16,000
Test	2682 (67.05 %)	934 (23.35 %)	186 (04.65 %)	152 (03.80 %)	46 (01.15 %)	4000
Overall	13,406 (67.03 %)	4676 (23.38 %)	930 (04.65 %)	762 (03.81 %)	226 (01.13 %)	20,000

The ages of patients ranged from 20 to 80 years, with probabilities associated with specific age groups. The most common ages among the patients were 42, 45, and 44, each with a probability of 0.046, indicating a relatively higher frequency of patients in their early forties. A histogram of this age distribution is presented in Fig. 1, providing a visual representation of the patient age spread across the datasets.

3.2. Data pre-processing

A sequence of advanced preprocessing techniques has been systematically implemented in this study to enhance mammogram images for analysis by deep learning models. Indeed, this study utilized the morphological erosion preprocessing technique to enhance mammographic images [38] as an initial step in the image enhancement process. This technique was used to reduce the complexity and the granular noise present in raw mammograms to enhance the visibility of significant structures within the breast tissue and suppress irrelevant noise. Furthermore, the technique involved the convolution of a specific structuring element across the image. This matrix, carefully defined in terms of size and shape through preliminary experiments, interacted with the local pixels of the image. When the pixel configuration did not match the structuring element, the process would adjust the central pixel, gradually decrease the boundaries, and lower the dimensionality of the structures.

Secondly, the CLAHE was employed [11]. CLAHE is an image processing technique that improves the contrast of images by redistributing the intensity values of the pixels. CLAHE is a variant of AHE, which also redistributes intensity values but can lead to over-amplification of noise in relatively homogeneous regions of an image. In this phase, the mammography image is segmented into distinct tiles with each tile undergoing individual histogram equalization [10,11]. CLAHE's primary strength is its ability to highlight local contrast, enhancing sophisticated details like micro-calcifications. An essential aspect of this method is the inclusion of a contrast threshold given that if a histogram bin exceeds this threshold, the excess is spread out to the neighboring bins. Thus, helping to avoid excessive contrast boosts that could amplify noise. CLAHE mitigates this concern by controlling the contrast boost in different image sections by limiting the histogram of every segment to a set value before determining the Cumulative Distribution Function (CDF).

The CLAHE has been estimated according to the following equation:

$$CDF(x) = \sum_{y=0}^{x} \frac{H(y)}{NT} \tag{1}$$

Where.

- $CDF(x)$ is the cumulative distribution function at intensity level x.
- $H(y)$ is the histogram value at intensity level y.
- N is the total number of pixels in the image
- T is the clipping value

The clipping value determines the balance between contrast enhancement and noise reduction. A higher value reduces contrast but cuts noise to a higher extent, while a lower one increases contrast yet reduces noise to a lower extent.

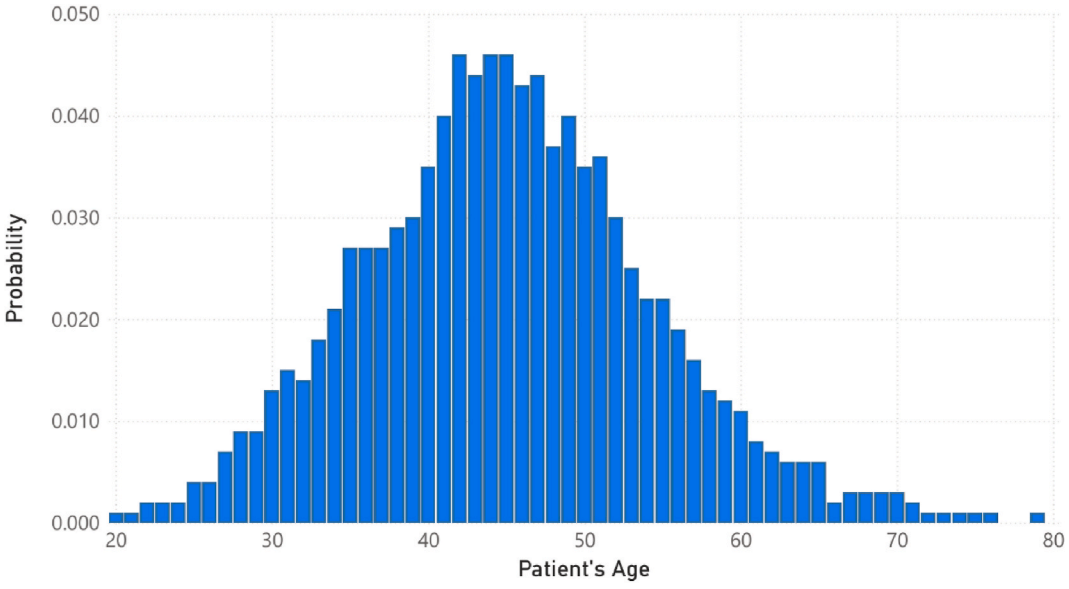


Fig. 1. Patient age distribution in study datasets.

To address noise issues in consistent areas of an image, CLAHE applies.

- **Histogram clipping**: Each region's histogram is controlled at a set value before determining the CDF, reducing excessive noise amplification.
- **Larger tile size**: The image is segmented into grid tiles, each treated separately, reducing noise's effect on enhancing contrast.
- **Bilinear interpolation**: When merging processed tiles for the final image, CLAHE employs bilinear interpolation, smoothing tile edges and lessening noise-induced artifacts.

Furthermore, the Effective Measure of Enhancement (EME) was used to quantify and measure the image enhancement obtained by splitting the image into several blocks. The following equation is utilized for *EME*:

$$EME = \frac{1}{K_1 K_2} \sum_{L=1}^{K_2} \sum_{K=1}^{K_1} 20 \log \left(\frac{I_{max}(k,l)}{I_{min}(k,l)} \right) \tag{2}$$

Where K_1, K_2 are the number of horizontal and vertical blocks in the image, $I_{max}(k,l)$ and $I_{min}(k,l)$ are the maximum and minimum pixel values in each block. The Peak Signal-to-Noise Ratio (PSNR) was used to measure the deviation of the current image from the original image with respect to the peak value of the gray level. Given a reference image f and a test image g, both of size $M \times N$, the *PSNR* between f and g is defined by:

$$PSNR(f,g) = 10 \log_{10} \left(255^2 / MSE(f,g) \right) \tag{3}$$

Where,

$$MSE(f,g) = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} \left(f_{ij} - g_{ij} \right)^2 \tag{4}$$

As the Mean Squared Error (MSE) approaches zero, the *PSNR* value tends to approach infinity. This relationship demonstrates that a higher *PSNR* value corresponds to a higher image quality. In other words, as the *PSNR* value increases, the image quality improves, indicating a closer resemblance between the original and the enhanced image.

Thirdly the CLAHE enhanced images were subjected to the Laplacian of Gaussian (LoG) edge enhancement [39]. The two-step procedure commences with the image being mildly smoothed using a Gaussian filter to reduce sensitivity to noise, thereafter, a Laplacian operator identifies areas of rapid intensity transition, indicative of edges. This integration of Gaussian and Laplacian techniques improves the delineation of pathological structures.

Finally, the Unsharp Masking technique, a digital image sharpening method, has been utilized to enhance the clarity of details in mammography images, an essential step in improving diagnostic accuracy in the study [40]. This technique primarily works by amplifying the contrast at the edges within an image, thereby augmenting the sharpness and enhancing the visibility of finer details. In the implementation of Unsharp Masking, a slightly blurred version of the original mammographic image was created by applying a Gaussian blur filter. This process effectively suppressed noise and reduced high-frequency details, delivering what is referred to as the "unsharp" image. Subsequently, this blurred version was subtracted from the original image, resulting in the creation of a high-contrast mask that represented the detailed components removed during the blurring process.

The high-contrast mask was then carefully added back to the original mammography image to ensure that the reintroduction of the details occurred with an emphasis on their clarity, enhancing the sharpness of structures and features within the breast tissue. The degree of sharpening was precisely controlled by adjusting the extent to which the high-contrast mask was combined with the original image. This adjustment was vital to avoid over-sharpening that might introduce artificial features or noise, potentially leading to misinterpretation.

Having undergone comprehensive enhancement, the images are then presented to edRVFL for analysis. Trained on such enhanced images, the model examines and classifies potential pathologies. Radiologists can then employ these clearer, enhanced images to augment their diagnostic accuracy. The structured integration of these methods ensures the comprehensive and optimal processing of each mammography image, aiming to enhance the precision of deep learning evaluations and, by extension, the final diagnostic outcomes.

3.3. Ensemble deep Random Vector Functional Link (edRVFL)

The study utilized the edRVFL methodology, an advanced technique that integrates both deep learning and ensemble learning concepts, building on the foundation of the Random Vector Functional Link (RVFL) network [41]. The RVFL was introduced by Zhang et al., in 2016 [42] as a single-layer feedforward neural network.

Due to the RVFLs distinct and streamlined architecture, it has gained recognition in the academic community for its power in areas such as machine learning, image classification, pattern recognition, and signal processing.

The essential structure of the RVFL can be elaborated as follows.

- **Input layer:** This primary layer is adept at receiving and processing the incoming data stream.

- **Hidden layer**: Quite unique in its functionality, this layer is embedded with weights that are initialized randomly. Its design intent is to ensure that the processing of data remains unbiased and uninfluenced by any projected or anticipated outputs.
- **Output layer:** Culminating the entire processing, this layer is responsible for generating and forwarding the final expected output.

In regard to training the RVFL, detailed attention is primarily directed towards the optimization of the weights associated with the output layer. This sophisticated process typically necessitates the resolution of a linear least squares problem. Within this context, employing techniques like the conjugate gradient method has been proven to significantly enhance the efficiency of the entire operation.

As illustrated in Fig. 2, at each higher-level model, the features from the preceding model are combined with the original input data features after undergoing non-linear transformations. The computational essence of RVFL can be encapsulated in the mathematical representation:

$$y = W_o \tanh(W_i x + b_i) + b_o \tag{5}$$

Where.

- y stands for the predicted output.
- W_o and W_i designate the weights associated with the output and hidden layers, respectively.
- b_o and b_i are the bias terms corresponding to the output and hidden layers.
- x is representative of the input data.

The edRVFL methodology depends on the foundational RVFL and boosts its capabilities. This is achieved by developing multiple RVFL networks, with each network being accurately trained on unique and distinct subsets of the overarching dataset. In the grand scheme, the outputs derived from each of these individual networks are cohesively integrated. This strategic organization results in the formation of the final output of the edRVFL, creating a holistic model with minimized overfitting risks and enhanced generalizability across varied datasets. Mathematically, the edRVFL can be delineated as:

$$edRVFL(x) = F(XW_1 + b_1)W_2 + b_2 \tag{6}$$

Where.

- X symbolizes the input data.
- W_1 and W_2 represent weight matrices corresponding to the sequential RVFL networks.
- b_1 and b_2 denote bias vectors that are aligned with the respective RVFL networks.
- F is indicative of a nonlinear activation function.

The procedural training of the edRVFL can be enumerated in these steps.

1) Initiate by randomly setting the weight matrices and bias vectors for the RVFL networks.
2) Engage in individual training cycles for each of the RVFL networks, utilizing their specific data subsets for optimal outcomes.

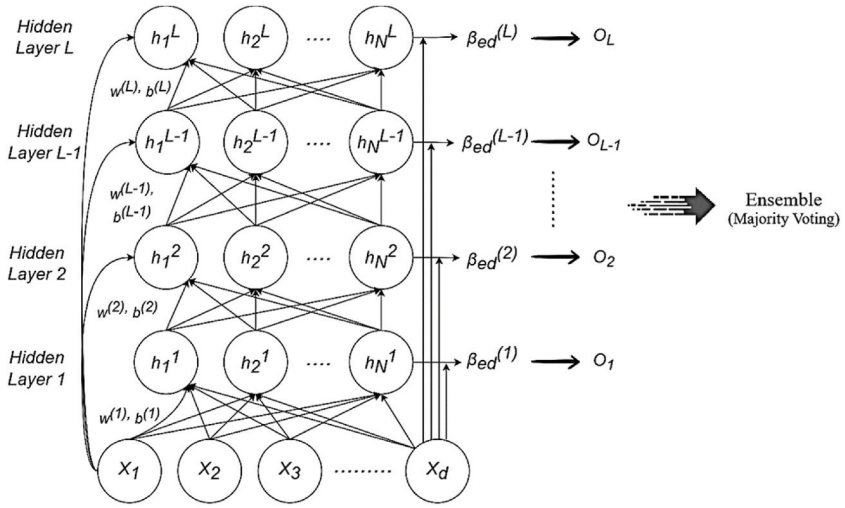


Fig. 2. The ensemble deep RVFL network.

3) Aggregate and amalgamate the outputs originating from each of the RVFL networks to synthesize the comprehensive edRVFL output.
4) Ensure iterative recalibrations by cycling through steps 2 and 3 until the model achieves a state of convergence.

The edRVFL technique brings to the forefront a multitude of benefits. Notably, it stands out for its ease of implementation and training, as well as its remarkable interpretability, particularly when contrasted with many deep learning algorithms. Further, it brings about efficient computational time and memory usage, alongside a robust defense against data noise and overfitting. Despite being in its early stages, the edRVFL's profound potential for a spectrum of applications is evident, as pointed out by this study.

3.3.1. Optimizing edRVFL for mammogram classification

In our study, we focused on optimizing hyperparameters for the edRVFL algorithm, which significantly influences its classification capability. We precisely adjusted three key hyperparameters: enhancement nodes, the regularization parameter, and hidden layers. These were optimized concurrently to ensure optimal algorithm performance and efficient resource use. By varying the number of enhancement nodes within the edRVFL architecture, we aimed to determine the ideal configuration for achieving superior classification results. This grid-search approach allows us to find the right balance between model complexity and resource consumption.

The regularization parameter plays a pivotal role in controlling overfitting and generalization. We investigated different values for this parameter to prevent the model from fitting noise, thereby enhancing its predictive accuracy. Additionally, the regularization parameter contributes to memory efficiency and better resource utilization. The number of hidden layers is a critical factor affecting the model's ability to capture complex features. We systematically explored the optimal number of hidden layers that would yield the best classification results. This step impacts not only classification accuracy but also execution time, as deeper networks tend to require more computational resources.

To test our approach, we established a concurrent execution pool, harnessing the power of 24 cores from the Intel Core i9 14900K processor. We then evaluated the effectiveness of various hyperparameter combinations by measuring different performance metrics, including accuracy, execution time, and resource utilization. A sample dataset consisting of 500 benign images and 500 normal images was utilized. Considering the search space, the total number of trials would be 250,000 trials. Since the process involving a cross-validation strategy with 25 folds for each distinct hyperparameter combination, the real number of trials would be 6,250,000 trials, which is approximately 3 days of continuous computation. This thorough approach helps us find the best hyperparameters, making sure our binary classification model is finely tuned for the best performance.

3.4. Hybrid model

3.4.1. Lesion detection and segmentation model

Image segmentation in the medical domain allows for detailed analysis and diagnosis. Traditional techniques often necessitate manual labeling for every unique anatomical feature or pathology, making them less adaptable. Recently, foundation models have emerged as potential tools for broader medical image segmentation. Meta's Segment Anything Model (SAM) by Alexander Kirillov et al. (2023) was an initial attempt at this [43]. SAM could segment diverse anatomical structures across varied imaging techniques with a textual prompt. Recently, Segment Anything in Medical Images (MedSAM) was introduced by Bo Wang et al. (2023), with an advanced improvement in automatic image segmentation [44]. MedSAM was pre-trained on over a million diverse medical images. Tests indicated that MedSAM outperformed many task-specific models and was adept at generalization. Consequently, it emerged as a powerful tool for medical image segmentation.

This study introduced a hybrid method for the automatic segmentation of mammogram breast lesions by merging supervised lesion detection with unsupervised segmentation utilizing foundation models, including the YOLOv5 object detection and MedSAM segmentation models [44,45]. This aimed to utilize the power of pre-trained models while minimizing data labeling efforts needed for wholly supervised models.

The study team utilized the publicly accessible VinDr-Mammo dataset, comprised of mammogram images with diverse lesion characteristics. Expert radiologists marked these images by enclosing lesions within bounding boxes. For model training consistency, all images were resized to 640x640 pixels, and normalized pixel intensity values to account for scanning variations. Next, the YOLOv5 object detection model was trained on this dataset for lesion detection. Chosen for its leading performance in real-time detection, YOLOv5 was trained for 250 epochs using the Stochastic Gradient Descent (SGD) optimizer and an initial learning rate of 0.001. Standard data modification methods like flipping and rotating were applied. Post-training, the YOLOv5 model achieved 95.46 % accuracy in detecting lesion areas on mammogram images. Furthermore, the MedSAM was utilized for lesion segmentation based on the detected bounding regions by the YOLOv5 model instead of processing the segmentation on the whole image. Thus, channeling MedSAM's segmentation capabilities towards the identified lesion areas. By integrating YOLOv5's detection precision with MedSAM's broad segmentation skills in this two-tier method, the proposed model has achieved higher segmentation accuracy than by only utilizing MedSAM.

The Dice Similarity Coefficient (DSC) was estimated to measure the accuracy of the hybrid model for both approaches: the MedSAM model, and the hybrid model integrating YOLOv5 and MedSAM [46]. The DSC is a popular metric in assessing the quality of segmentation as it calculates the overlap between two images – the predicted segmentation and the true segmentation. DSC values range from 0 to 1. A score of 1 represents a perfect overlap between the two images, as evidenced below:

$$DSC = \frac{2\,|A \cap B|}{|A| + |B|} \tag{6}$$

Where.

- A is the predicted segmentation
- B is the ground-truth segmentation

3.4.2. Ensemble deep-learning classification model

The architecture of the proposed hybrid model is illustrated in Fig. 3. The model started with data acquisition from Digital Imaging and Communications in Medicine (DICOM) mammography images ensuring a balanced representation of both benign and malignant cases. To standardize the input for subsequent stages, all images were resized to uniform dimensions suitable for subsequent deep-learning architectures.

Image enhancement played a pivotal role in improving the clarity and quality of the images. The first step in the enhancement phase was the application of a sequence of image enhancement algorithms, including morphological erosion, CLAHE, Gaussian Blur, LoG, and unsharp mask techniques. The CLAHE boosted image contrast without over-amplifying it, ensuring a uniformly enhanced image. Subsequently, Gaussian Blur was implemented to suppress noise and preserve significant structures within the mammograms. This iterative method enhanced the image by reducing high-frequency noise. Lastly, to highlight ROI, such as edges or potential anomalies within the breast tissue, the LoG and unsharp mask techniques were adopted to improve the lesion of interest visibility.

Following image enhancement, the edRVFL was employed to process the images. In addition, two deep CNN architectures, VGG16 and DenseNet121, were also employed. VGG16, recognized for its depth with 16 weighted layers, was tailored for our dataset by fine-tuning its latter layers [47]. On the other hand, DenseNet121, renowned for its densely connected layers ensuring effective gradient flow and feature reuse, was similarly adapted for our mammography images [48]. Both architectures utilized transfer learning, initializing their weights from models pre-trained on ImageNet.

After training, a thorough comparative analysis between VGG16 and DenseNet121 was conducted, recording performance metrics such as accuracy, sensitivity, specificity, and F1-score to identify which architecture exhibited superior efficacy in predicting breast cancer from the mammography images. The most proficient model, as per the analysis, was integrated with the outputs from the edRVFL, thereby forming the final hybrid deep learning model. This integrated model was subsequently evaluated on a distinct test set of mammography images, and its performance metrics were thoroughly documented.

3.4.3. Testing in clinical setting

The hybrid model underwent training on annotated mammography images from the VinDr-Mammo dataset, categorized by the BI-RADS standards [49], clustering the data into training, testing, and validation subsets. As the model progressed through its training phase, it continually fine-tuned its parameters. By employing conventional optimization methods, the aim was to reduce misclassifications, thus increasing its predictive precision. In this research, a diverse set of machines learning models, including RF, Gradient Boosting (GB), SVM, Logistic Regression (LR), and the edRVFL, were evaluated against each other to measure their efficacy in breast cancer predictions.

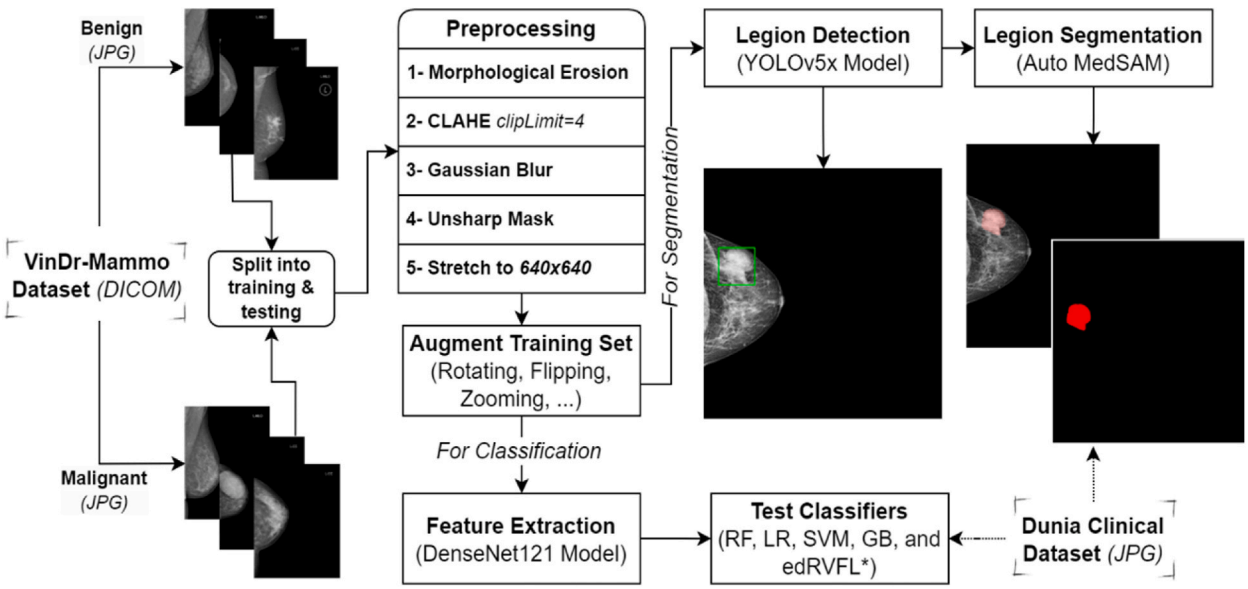


Fig. 3. Overall architecture of the suggested hybrid clinical model.

The models' efficacy was measured by using criteria such as accuracy, precision, recall, F1-score, and AUC. The goal was to determine the model demonstrating exceptional predictive ability and a solid capacity to generalize. As part of our thorough evaluation, the data was meticulously partitioned, cross-validated, and statistically analyzed to ascertain an unbiased assessment of the models in implementation. In a more practical possibility, the hybrid model was also field-tested in a clinical environment, leveraging data from 1888 cases sourced from the Dunya Women's Cancer Center in Palestine, dated August 2023. After exhaustive training on the primary dataset, the model was then applied to the clinical data.

The model's predictions, based on features gathered from the mammography images, were compared with histopathological findings. A comprehensive tuning of model hyper-parameters was undertaken as posting the models' outcomes on the validation data were scrutinized using metrics like accuracy, precision, recall, F1-score, and AUC. Analyzing the outcomes from both methods was essential to validate the models' consistency and adaptability. Beyond that, by comparing various machine learning models, our aim was to recognize the most effective algorithm for breast cancer detection, correlating with BI-RADS categorization and pathological assessments.

4. Results

The results of the study are presented in five key phases. First, we discuss the hyperparameter optimization results for the edRVFL algorithm, which set the foundation for subsequent analyses. Second, we assess the hybrid model's performance using unenhanced data, establishing a baseline for comparison. Third, we apply image enhancement techniques, particularly CLAHE (Contrast Limited Adaptive Histogram Equalization), to investigate performance improvements. Fourth, we evaluate the model's deployment in a real-world clinical environment, bridging the gap between theoretical robustness and practical application. Finally, we present a Grad-CAM analysis to enhance the interpretability of the DenseNet121 model in detecting and classifying breast cancer from mammograms.

Throughout all phases, we focused on three key components to evaluate the most suitable classifier for clinical settings: 1) Accuracy, 2) Precision, and 3) Processing Time. Striking a fair balance between accuracy and processing time within the chosen model proved imperative for its effectiveness and efficiency in time-sensitive clinical settings. Additionally, we compared our results with related works to demonstrate the advancements achieved by our proposed method.

4.1. Hyperparameters optimization results for edRVFL

The performance of the edRVFL algorithm improved as we systematically explored and tuned three crucial hyperparameters to obtain the best combination. As illustrated in Table 2, the optimal hyperparameters were found to be 10 enhancement nodes, a regularization parameter of 100, and 10 hidden layers.

This configuration achieved an impressive accuracy of 0.995 on the sample dataset, compared to 0.981 with the default hyperparameters. The improved performance is attributed to the higher regularization parameter, which effectively controlled overfitting, leading to a simpler final hypothesis and better generalization. Additionally, this combination resulted in more efficient CPU usage, reducing it from 0.16 percent to 0.1 percent. However, the execution time increased slightly from 0.55 to 0.75 s due to the increased number of nodes, adding computational complexity. Overall, these adjustments demonstrate a significant improvement in the model's performance and efficiency.

4.2. Performance of the classification model before image enhancement

The performance evaluation of various machine learning classifiers, integrated with two deep learning models (VGG16 and DenseNet121), was precisely carried out to understand the efficacy of each approach in classifying benign and malignant lesions based on the original, unenhanced data. Results in Table 3 illustrate the model's performance analysis in the detection of benign and

Table 2
Performance of edRVFL corresponding to systematically explore and tune three crucial hyperparameters to obtain the best combination.

Default	Number of Nodes	**2**
	Regularization Parameter	1
	Number of Layers	2
	Execution Time (s)	0.55
	Accuracy	0.981
	CPU Usage	0.16
Best Combination	**Number of Nodes**	10
	Regularization Parameter	100
	Number of Layers	10
	Search Space	Num. of Nodes (1 − 20), Regularization Para. (1 − 200), Num. of Layers (1 − 20)
	Execution Time (s)	0.75
	Accuracy	0.995
	CPU Usage	0.1

Table 3
Performance evaluation of the unenhanced data on two deep-learning models concerning several machine-learning classifiers.

Model	Classifier	Accuracy	Precision	Recall	F1	AUC	Time(S)
Benign Prediction							
VGG16	Random Forest	0.846	0.610	0.315	0.460	0.806	36.21
	Gradient Boosting	0.865	0.584	0.338	0.432	0.841	261.52
	SVM	0.829	0.642	0.319	0.418	0.834	52.29
	Logistic Regression	0.848	0.511	0.582	0.524	0.814	0.8
	edRVFL	0.865	0.844	0.867	0.846	0.852	2.02
DenseNet121	Random Forest	0.853	0.656	0.342	0.513	0.836	22.46
	Gradient Boosting	0.869	0.631	0.381	0.508	0.858	187.15
	SVM	0.847	0.646	0.332	0.517	0.848	14.97
	Logistic Regression	0.856	0.523	0.543	0.540	0.837	1.14
	edRVFL	0.879	0.859	0.859	0.851	0.882	0.84
Malignant Prediction							
VGG16	Random Forest	0.850	0.652	0.318	0.471	0.827	10.56
	Gradient Boosting	0.856	0.646	0.398	0.480	0.822	78.38
	SVM	0.846	0.552	0.392	0.465	0.836	13.42
	Logistic Regression	0.852	0.560	0.460	0.531	0.834	0.2
	edRVFL	0.887	0.845	0.866	0.847	0.849	0.29
DenseNet121	Random Forest	0.867	0.710	0.369	0.553	0.842	15.58
	Gradient Boosting	0.871	0.625	0.412	0.544	0.841	122.78
	SVM	0.855	0.671	0.486	0.583	0.855	11.71
	Logistic Regression	0.872	0.579	0.564	0.560	0.843	0.1
	edRVFL	0.889	0.889	0.887	0.888	0.894	0.6

F1: F1-Score; AUC: Area under the Curve; SVM: Support Vector Machine.

malignant cases.

The result of benign prediction shows that the edRVFL integrated with VGG16 demonstrated high accuracy at 0.865, and an impressive AUC of 0.852, taking just 2.02 s. In contrast, the GB algorithm, while holding a comparable accuracy of 0.865, required significantly more time (261.5 s). However, the integration of edRVFL with DenseNet121 provided better performance metrics,

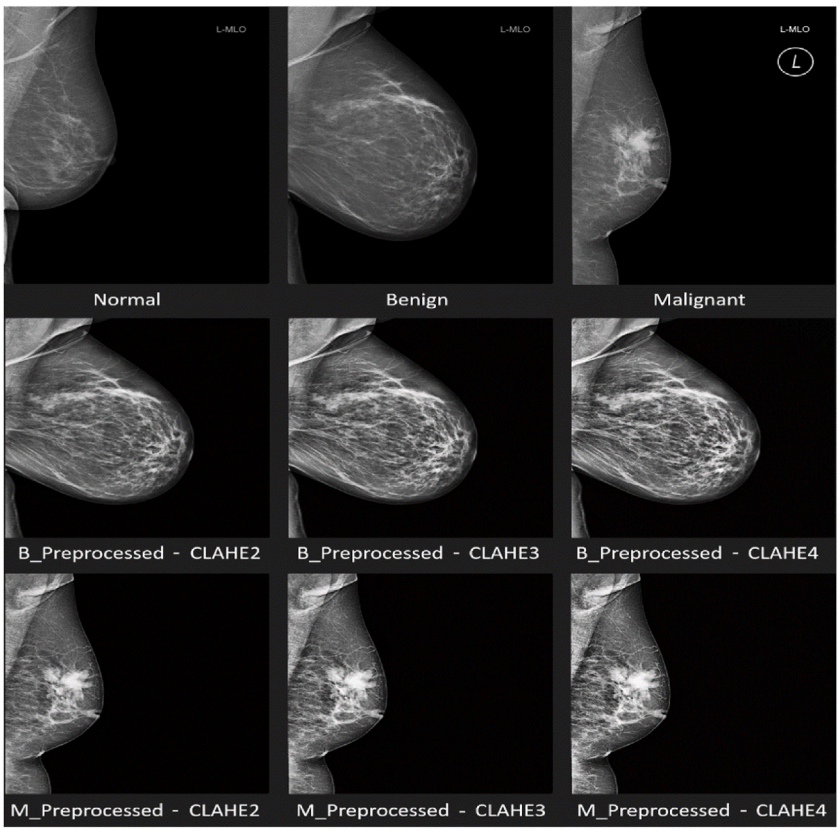


Fig. 4. The effectiveness of CLAHE algorithm in enhancing benign and malignant mammogram images at different clip limits (2, 3, and 4).

yielding an accuracy of 0.879 and an AUC of 0.882, within just 0.84 s, showcasing the potential for this approach in efficient real-time applications.

For malignant classifications, the edRVFL model again emerged with the best performance. Within the VGG16 framework, edRVFL achieved an accuracy of 0.887, compared to other classifiers, with LR presenting a viable alternative at a lower 0.852 accuracy but with an efficient processing time of 0.29 s. The combination of edRVFL and DenseNet121 provided an accuracy of 0.889 and reported the highest AUC of 0.894. The efficiency was also noteworthy here, with results obtained in just 0.6 s, emphasizing the model's suitability for time-sensitive clinical environments.

4.3. Performance of the classification model with image enhancement

The evaluation of the enhanced data using pre-processing tools, mainly the CLAHE algorithm, compared with various machine learning classifiers, reveals significant improvements in diagnostic precision, particularly in terms of accuracy and processing time. Fig. 4 illustrates the effectiveness of CLAHE in enhancing mammography images, showcasing the improvements in image contrast and quality achieved through this image processing technique.

As illustrated in Table 4, the ML classifiers demonstrated evident accuracy improvement post-CLAHE application within the benign cases, with VGG16 and DenseNet121 models showing substantial enhancements. Interestingly, LR under VGG16 achieved an impressive accuracy of 0.985 with an almost instantaneous processing time of 0.13 s, indicating fast diagnostic capability without compromising on precision. Equally compelling, the edRVFL classifier paired with DenseNet121 reported the highest accuracy of 0.997, a minimal 0.75-s processing time emphasizing the efficiency of enhanced images in reducing computational demands.

In predicting malignant cases, the precision of classifiers further emphasized the benefits of image enhancement. Under the VGG16 model, LR and the edRVFL reported high accuracy rates of 0.983 and 0.993, maintaining rapid computation at 0.08 and 0.34 s respectively. The edRVFL's performance increased with DenseNet121, reporting a near-perfect accuracy of 0.998, alongside a fast 0.75-s processing time, highlighting the model's proficiency in recognizing enhanced features for the accurate detection of malignant lesions.

The GB classifier, while showing accuracy equivalent to other models, particularly a notable 0.982 in malignant lesion detection with DenseNet121, required considerably more processing time, peaking at 219.23 s. Furthermore, the enhancement of mammograms using CLAHE significantly strengthened the accuracy of all employed classifiers. Nonetheless, the standout performance of edRVFL and LR, given their combination of exceptional accuracy and reduced processing times, emphasizes their potential for reliable, real-time clinical decision-making.

4.4. Performance of the classification model in clinical setting

The performance evaluation of the proposed hybrid model on a real-world clinical dataset, particularly referring to confirmed histopathology results, has provided critical insights into the model's practical efficiency and reliability. Results in Table 5 show the model's performance analysis based on patients' clinical data.

As indicated in Table 5, the model was applied on the unenhanced dataset, where the classifiers, when employed with the VGG16

Table 4
Performance evaluation of the enhanced data on two deep-learning models using the CLAHE algorithm and several machine-learning classifiers.

Model	Classifier	Accuracy	Precision	Recall	F1	AUC	Time(S)
Benign Prediction based on Enhanced images at $CLAHE(Clip\ Limit = 4)$							
VGG16	Random Forest	0.952	0.985	0.803	0.814	0.976	4.83
	Gradient Boosting	0.973	0.978	0.867	0.924	0.987	58.25
	SVM	0.975	0.981	0.959	0.841	0.986	4.45
	Logistic Regression	0.985	0.986	0.979	0.987	0.987	0.13
	edRVFL	0.991	0.985	0.986	0.986	0.989	1.93
DenseNet121	Random Forest	0.969	0.902	0.881	0.836	0.994	15.43
	Gradient Boosting	0.957	0.985	0.872	0.939	0.992	141.05
	SVM	0.986	0.991	0.946	0.987	0.999	8.08
	Logistic Regression	0.991	0.996	0.978	0.989	0.999	0.15
	edRVFL	0.997	0.998	0.997	0.997	0.999	0.75
Malignant Prediction based on Enhanced images at $CLAHE(Clip\ Limit = 4)$							
VGG16	Random Forest	0.959	0.988	0.825	0.893	0.992	5.33
	Gradient Boosting	0.976	0.977	0.874	0.933	0.991	43.94
	SVM	0.980	0.989	0.961	0.848	0.994	3.74
	Logistic Regression	0.983	0.960	0.950	0.955	0.996	0.08
	edRVFL	0.993	0.988	0.986	0.987	0.997	0.34
DenseNet121	Random Forest	0.973	0.988	0.895	0.950	0.996	25.49
	Gradient Boosting	0.982	0.979	0.892	0.947	0.996	219.23
	SVM	0.992	0.992	0.966	0.982	0.998	19.03
	Logistic Regression	0.995	0.995	0.995	0.995	0.999	0.22
	edRVFL	0.998	0.998	0.997	0.998	0.999	0.81

F1: F1-Score; AUC: Area under the Curve; SVM: Support Vector Machine.

Table 5
Performance evaluation of the proposed hybrid model on a real-world clinical dataset before enhancement, against confirmed histopathology results.

Model	Type		Pred(%)	Accuracy	Precision	Recall	F1	AUC	Time
VGG16	Benign	Random Forest	0.857	0.833	0.928	0.406	0.565	0.798	9.53
		Gradient Boosting	0.798	0.792	0.431	0.218	0.359	0.796	68.82
		SVM	0.809	0.791	0.431	0.219	0.359	0.799	13.76
		Logistic Regression	0.797	0.767	0.536	0.5	0.517	0.772	0.21
		edRVFL	0.821	0.783	0.778	0.783	0.781	0.817	1.32
	Malign	Random Forest	0.833	0.775	0.352	0.294	0.324	0.717	4.40
		Gradient Boosting	0.764	0.766	0.692	0.29	0.409	0.794	32.66
		SVM	0.778	0.757	0.284	0.256	0.27	0.736	5.59
		Logistic Regression	0.764	0.748	0.591	0.406	0.481	0.749	0.082
		edRVFL	0.792	0.784	0.772	0.784	0.773	0.774	0.12
DenseNet 121	Benign	Random Forest	0.893	0.842	0.681	0.406	0.58	0.838	5.91
		Gradient Boosting	0.845	0.825	0.722	0.448	0.553	0.72	49.25
		SVM	0.856	0.833	0.604	0.31	0.474	0.882	3.94
		Logistic Regression	0.881	0.858	0.767	0.697	0.73	0.886	0.30
		edRVFL	0.905	0.85	0.847	0.85	0.848	0.901	0.22
	Malign	Random Forest	0.903	0.793	0.656	0.303	0.465	0.798	6.49
		Gradient Boosting	0.819	0.774	0.733	0.403	0.444	0.774	51.16
		SVM	0.833	0.784	0.562	0.273	0.429	0.814	4.88
		Logistic Regression	0.861	0.829	0.815	0.611	0.698	0.871	0.04
		edRVFL	0.916	0.847	0.857	0.846	0.838	0.874	0.25

F1: F1-Score; AUC: Area under the Curve; SVM: Support Vector Machine.

and DenseNet 121 models, demonstrated an acceptable prediction accuracy compared to the histopathology reports. As evidenced, edRVFL showed the best performance results compared to other classifiers, particularly in benign (B) case predictions with VGG16, reporting an accuracy of 0.821, and in malignant (M) cases with DenseNet 121, reaching up to 0.916. However, these figures may indicate issues experienced by the classifiers in the extraction of features presented in unenhanced medical images, such as challenges in feature extraction due to a lack the necessary contrast, clarity, and detail.

Furthermore, the results in Table 6 evidence the model's improvement of breast cancer prediction based on the enhanced dataset, mainly with the application of the CLAHE enhancement, particularly with a clip limit of 4.

The classifiers across both VGG16 and DenseNet121 models exhibited a remarkable improvement in their predictive accuracies. For benign cases with VGG16, the accuracies significantly increased, with edRVFL achieving a 0.952 score. The trend becomes even more efficient when applying the DenseNet 121 model, where the edRVFL reached a near-perfect accuracy of 0.976. ‘

For malignant predictions, the enhanced data fed into the VGG16 model shows similar improvements, with edRVFL leading at a 0.972 accuracy level. The performance increased in the DenseNet 121 application, where edRVFL's accuracy reached a high accuracy rate of 0.986, almost perfectly aligning with confirmed histopathological findings.

The achievements indicate that the proposal outperforms the state-of-the-art not only using DenseNet architectures, but across a range of different approaches and models. A complete comparison with the state-of-the-art is presented in Table 7.

Table 6
Performance evaluation of the enhanced hybrid model's clinical validation in reference to confirmed histopathology.

Model	Type	Classifier	Pred. (%)	Accuracy	Precision	Recall	F1	AUC	Time
VGG16	Benign	Random Forest	0.904	0.883	0.958	0.638	0.766	0.93	1.93
		Gradient Boosting	0.917	0.925	0.965	0.777	0.862	0.947	23.30
		SVM	0.917	0.916	0.722	0.893	0.838	0.988	1.78
		Logistic Regression	0.94	0.942	0.872	0.944	0.907	0.986	0.05
		edRVFL	0.952	0.967	0.967	0.966	0.967	0.99	0.77
	Malign	Random Forest	0.931	0.923	0.816	0.8	0.811	0.948	4.10
		Gradient Boosting	0.917	0.921	0.968	0.75	0.845	0.974	33.80
		SVM	0.917	0.901	0.706	0.852	0.81	0.991	2.88
		Logistic Regression	0.944	0.937	0.902	0.925	0.914	0.994	0.06
		edRVFL	0.972	0.964	0.964	0.963	0.963	0.996	0.26
DenseNet 121	Benign	Random Forest	0.952	0.95	0.933	0.824	0.903	0.984	6.17
		Gradient Boosting	0.929	0.933	0.964	0.794	0.871	0.973	56.42
		SVM	0.94	0.933	0.814	0.765	0.866	0.98	3.23
		Logistic Regression	0.952	0.967	0.886	0.975	0.939	0.999	0.06
		edRVFL	0.976	0.983	0.984	0.983	0.984	1	0.30
	Malign	Random Forest	0.931	0.919	0.894	0.729	0.836	0.982	19.61
		Gradient Boosting	0.931	0.928	0.955	0.789	0.862	0.976	168.64
		SVM	0.944	0.928	0.916	0.79	0.857	0.99	14.64
		Logistic Regression	0.958	0.955	0.921	0.946	0.933	0.998	0.17
		edRVFL	0.986	0.982	0.983	0.982	0.982	1	0.75

Pred%: Percentage of model compared to histopathology reports. F1: F1-Score; AUC: Area under the Curve; SVM: Support Vector Machine.

Table 7
Comparison of breast cancer classification methods across related works.

Related works	Pre-processing	Cases	Classifier	Architecture	Best Accuracy
Kebede et al. [50]	–	1281	ImageNet	EfcientNet-B3	89.0
Yu et al. [29]	Morphological opening, and thresholding	330	Customized CNN	DenseNet201	92.73
Pardamean et al. [51]	–	2605	ChexNet	DenseNet121	90.38
Hekal et al. [52]	Otsu thresholding	2800	SVM	AlexNet	91.00
Altan G [31]	Cropping	2620	Customized CNN	–	92.84
Altan G [32]	–	16,200	DBN	–	96.32
This Study	Morphological erosion, CLAHE, Gaussian blur, and unsharp mask	1888 *(Biopsy validated)*	edRVFL with optimized hyperparameters	DenseNet121	**0.986**

Our proposed method achieves the highest accuracy of 0.986, surpassing all previous works by a significant margin. This represents a substantial improvement over the next best result of 96.32 % achieved by Altan using a DBN classifier [32].

4.5. Performance of the segmentation model in clinical setting

Fig. 5 illustrates the segmentation results obtained using the proposed hybrid model. On the left side, the segmentation outcome prior to enhancement is displayed, with the hybrid model yielding an average DSC score of 0.942 relative to the reference lesion, while the MedSAM score registers at 0.824.

On the other hand, the right side showcases the results post-CLAHE image enhancement, indicating an improvement in DSC with a score of 0.987, and a MedSAM score of 0.935. The hybrid model for clinical purposes combines multiple techniques: image enhancement using the CLAHE algorithm, prediction, and classification via the optimized edRVFL algorithm, complemented by an integrated segmentation strategy deploying both YOLOv5 and MedSAM models. These advancements suggest a promising tool for enhancing diagnostic accuracy and efficiency in clinical settings.

4.6. Grad-CAM analysis

Grad-CAM (Gradient-weighted Class Activation Mapping) was employed in this study to enhance the interpretability of the DenseNet121 model in detecting and classifying breast cancer from mammograms. This advanced visualization technique generates heatmaps that highlight the most influential regions within input images, offering crucial insights into the model's decision-making process. By visualizing which areas of the mammogram contribute most significantly to the classification, Grad-CAM helps bridge

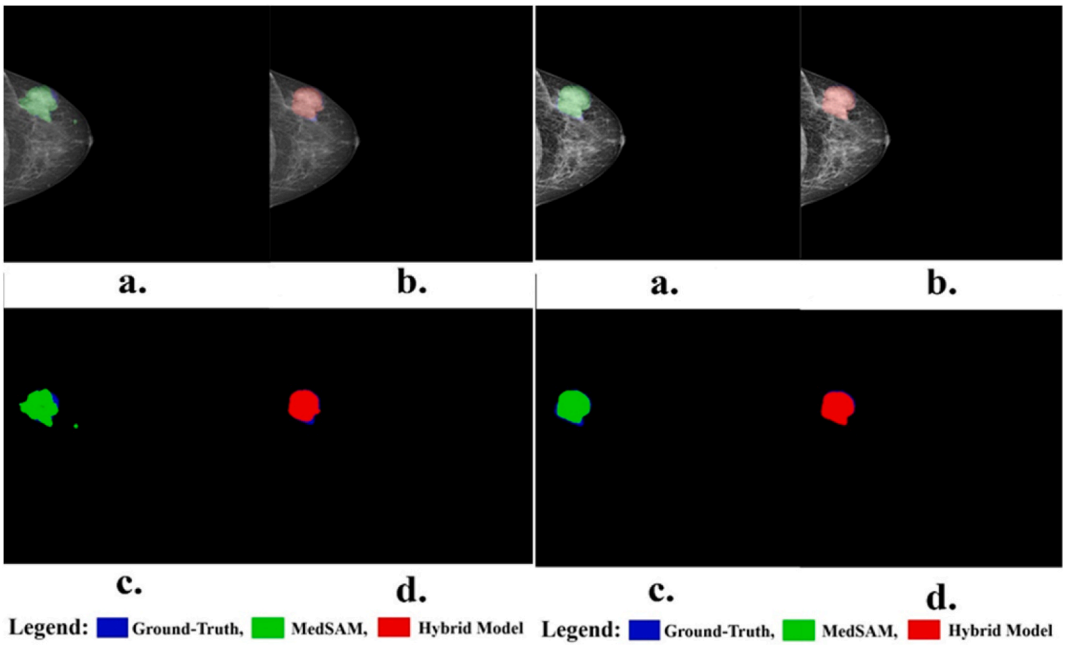


Fig. 5. Hybrid segmentation model results based on original and enhanced images.

the gap between the model's performance and clinical interpretability. This approach is particularly valuable in medical imaging, where understanding the basis for AI-driven decisions is critical for clinical trust and adoption.

The study focused on mammograms enhanced using the CLAHE algorithm, which improves image contrast and detail. CLAHE was specifically chosen for its ability to enhance local contrasts without over-amplifying noise, a common issue in medical imaging. This preprocessing step proved crucial in highlighting subtle tissue differences that might indicate the presence of cancerous lesions. By adaptively equalizing the histogram of local image regions, CLAHE enhances the visibility of important structures while preserving the overall image characteristics. Fig. 6 shows Grad-CAM heatmaps for a confirmed malignant case, comparing original and CLAHE-enhanced mammograms. The CLAHE-enhanced heatmap show stronger activation in areas with irregular shapes and higher density, typical of tumors.

This highlights how CLAHE enhancement helps the model identify important features associated with breast cancer, potentially improving the accuracy and reliability of the AI-assisted diagnosis. The clearer image allows for more precise detection of potential malignancies, demonstrating the value of this preprocessing technique in AI-assisted mammogram analysis.

The Grad-CAM analysis provided valuable visual explanations for the DenseNet121 model's predictions, aligning closely with clinical observations. By comparing original and enhanced heatmaps, the study demonstrates how CLAHE enhances the model's ability to detect subtle but crucial features indicative of breast cancer. This transparency allows healthcare professionals to better validate and understand the model's decisions, facilitating its integration into clinical practice for improved diagnosis and patient care. Moreover, the Grad-CAM visualizations serve as an educational tool, helping radiologists and clinicians understand how AI models interpret mammographic features, potentially leading to improved human-AI collaboration in breast cancer screening and diagnosis.

5. Discussion

This study presents a hybrid ensemble deep-learning model designed for breast cancer detection and classification, emphasizing enhanced mammograms for clinical applications. Our findings demonstrate that integrating deep learning models with image enhancement techniques significantly improves the diagnostic accuracy of breast cancer detection algorithms.

The image enhancement techniques applied in our model, including Morphological Erosion, CLAHE, Gaussian Blur, and unsharp masking, markedly improved the diagnostic quality of mammograms. Specifically, the CLAHE algorithm preserved essential diagnostic details without oversaturating the image, Gaussian Blur reduced noise, and the LoG highlighted critical features. These enhancements ensured clearer and more accurate mammogram analysis, consistent with studies by Dhungel et al. (2017) and Janan and Brady (2021) that emphasized the importance of image enhancement in diagnostic performance [53,54].

The incorporation of the edRVFL algorithm provided rapid convergence and computational efficiency. This, combined with the use of advanced CNNs like VGG16 and DenseNet121, further enhanced our model's performance. VGG16's structural depth allowed for optimized internal feature representations through precise fine-tuning processes tailored to our dataset's specific characteristics. DenseNet121's architecture promoted feature reuse and enhanced gradient flow, improving pattern recognition within mammograms. This approach aligns with research by Zhang et al., Liu et al., and Narayanan et al., who also demonstrated the effectiveness of hybrid deep learning methodologies [55–59].

Our results are compelling: the hybrid model achieved accuracy rates of 99.7 % for benign cases and 99.8 % for malignant cases on enhanced images, with processing times as low as 0.75 s. Clinical validation against histopathological findings showed accuracy rates of 97.6 % for benign cases and 98.6 % for malignant cases. These findings indicate the model's robustness and potential for real-time clinical applications. The integration of YOLOv5 and MedSAM segmentation models further improved the model's ability to accurately extract, classify, and visualize breast lesions.

Comparatively, our hybrid model addresses several challenges that impede the clinical application of AI in breast cancer detection. Prior studies, such as those by Janan and Brady (2021) and Yoon and Kim (2020), validated their models on standard datasets, but our study goes further by verifying the model's performance against actual histopathology reports [53,60]. This innovative approach is crucial for clinical applicability. While results on the unenhanced dataset were not as promising, the enhanced dataset provided confidence in the model's clinical utility. Notably, the model trained on enhanced data closely aligns with actual clinical findings, reinforcing its potential for real-time use.

Our model not only ensures more reliable results by cross-verifying findings but also reduces the likelihood of false results due to human error and minimizes image interpretation variability, leading to more consistent and accurate diagnoses. These advantages address the limitations identified in previous models and highlight the importance of an ensemble approach for reliable clinical outcomes. In comparison with other significant works, De Nazare Silva et al. (2015) achieved strong diagnostic performance using wavelet transform enhancement and SVM classification [24]. De Sampaio et al. (2015) reported high sensitivity and low false positive rates with an adaptive algorithm for breast density classification and micro-genetic algorithms [25]. Shen et al. (2019) developed an end-to-end CNN method with excellent performance on multiple mammography databases [61]. Gokhan Altan's 2020 study achieved accuracy, sensitivity, specificity, and precision rates of over 92 % using a CNN [31]. In 2021, Altan's study with DBNs further improved these metrics [32]. Lakshmi Narayanan (2022) combined AlexNet and RF classifier, achieving an accuracy of 98.3 % [34].

In this manner, although substantial advancements were achieved through this study, the complexity of mammogram interpretation presents limitations to the model. Several factors could arise in practical applications, such as varying imaging devices, and limitations in the training dataset might impact the model's performance. The training dataset must be diverse to ensure the model's universal applicability. Indeed, future efforts could potentially involve further refining the model, expanding the dataset for more comprehensive training, and exploring additional hybrid techniques to ensure that breast cancer detection is as precise and efficient as possible.

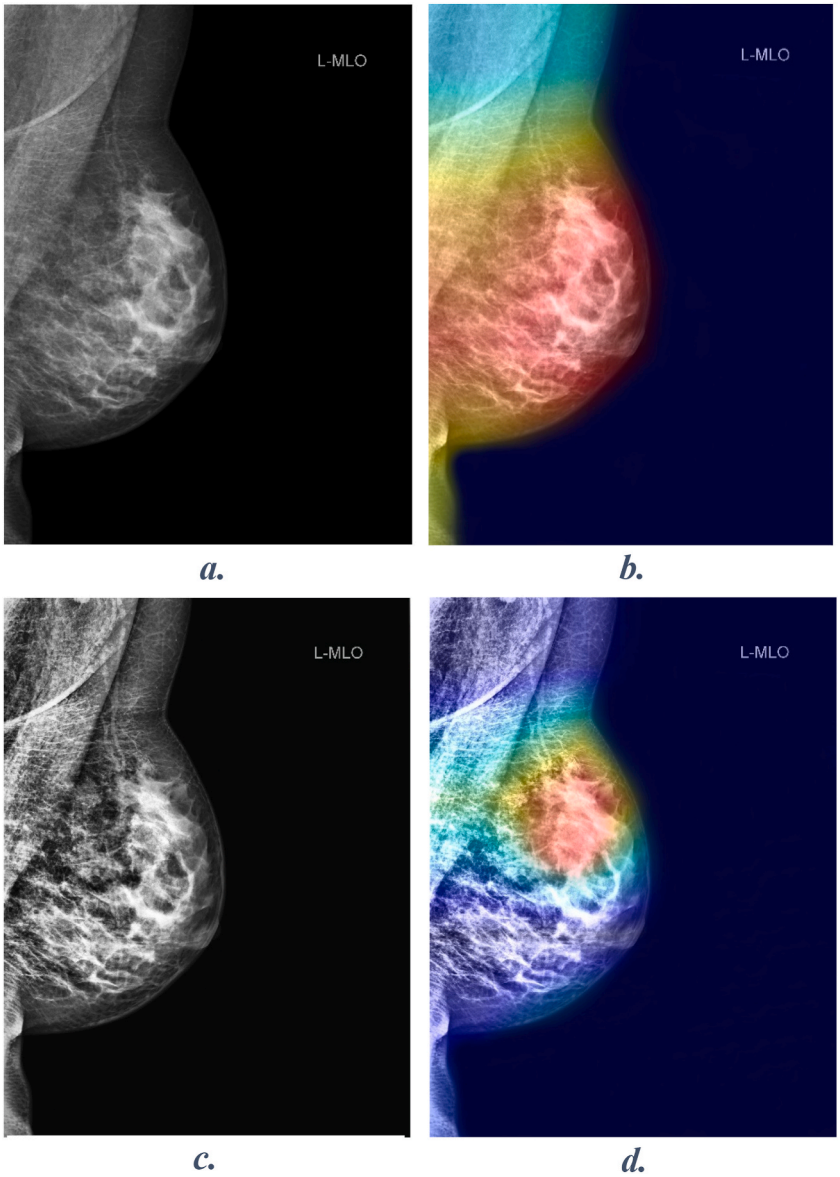


Fig. 6. Comparison of original and CLAHE-enhanced mammograms with their respective Grad-CAM heatmaps. (a) Original mammogram image; (b) Grad-CAM heatmap of the original image; (c) CLAHE-enhanced mammogram image; (d) Grad-CAM heatmap of the CLAHE-enhanced image.

6. Conclusion

This study presents a novel hybrid ensemble deep-learning model for breast cancer detection and classification through mammogram analysis. By integrating advanced image enhancement techniques with deep learning algorithms, the proposed model demonstrates significant improvements in diagnostic accuracy and efficiency compared to existing approaches. The implementation of a sequential image enhancement process, including morphological erosion, CLAHE, Gaussian Blur, and unsharp masking, substantially improved the quality and diagnostic clarity of mammogram images. This enhancement process, combined with the integration of the edRVFL algorithm and pre-trained CNN architectures (VGG16 and DenseNet121), yielded high accuracy in both benign and malignant lesion detection.

The hybrid model achieved impressive accuracy rates of 99.7 % for benign cases and 99.8 % for malignant cases when tested on enhanced images, with processing times as low as 0.75 s. These results highlight the model's potential for real-time clinical applications. Furthermore, clinical validation against histopathological findings demonstrated the model's robustness, with accuracy rates of 97.6 % for benign cases and 98.6 % for malignant cases. The incorporation of YOLOv5 and MedSAM segmentation models further enhanced the model's ability to extract, classify, and visualize breast lesions accurately.

These results indicate that the proposed hybrid model has the potential to significantly improve breast cancer detection and

classification in clinical settings. By providing rapid, accurate analysis of mammograms, this approach could support earlier diagnosis and treatment of breast cancer, potentially improving patient outcomes. However, limitations such as the need for diverse training datasets and potential variability in imaging equipment must be addressed in future research.

Ethical Statement

This study was reviewed and approved by the Al-Quds University Institutional Review Board (IRB) committee, with the approval number: 305/REC/2023, dated March 2023.

Data availability statement

The study utilized the publicly available VinDr-Mammo dataset available at www.vindr.ai/datasets/mammo.
The authors do not have permission to share the clinical data.

Funding

This research did not receive any specific funding.

Declaration of generative AI and AI-assisted technologies

During the preparation of this work, the authors used GPT-3.5 to improve readability and language. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

CRediT authorship contribution statement

Radwan Qasrawi: Writing – review & editing, Writing – original draft, Validation, Software, Project administration, Methodology, Investigation, Formal analysis, Conceptualization. **Omar Daraghmeh:** Visualization, Validation, Formal analysis, Data curation. **Ibrahem Qdaih:** Validation, Data curation. **Suliman Thwib:** Visualization, Validation, Software, Methodology, Formal analysis. **Stephanny Vicuna Polo:** Writing – review & editing, Writing – original draft, Resources. **Haneen Owienah:** Validation, Data curation. **Diala Abu Al-Halawa:** Validation, Methodology, Data curation. **Siham Atari:** Writing – review & editing, Visualization, Validation, Resources.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to acknowledge the Dunya Center for Cancer in Palestine for their vast support in the implementation of this research, as well as the doctors, radiologists, and field workers who have supported the successful completion of the study.

References

[1] H. Zerouaoui, A. Idri, Reviewing machine learning and image processing based decision-making systems for breast cancer imaging, J. Med. Syst. 45 (2021), https://doi.org/10.1007/s10916-020-01689-1.
[2] C.D. Lehman, S. Mercaldo, L.R. Lamb, T.A. King, L.W. Ellisen, M. Specht, R.M. Tamimi, Deep learning vs traditional breast cancer risk models to support risk-based mammography screening, J Natl Cancer Inst 114 (2022) 1355–1363, https://doi.org/10.1093/jnci/djac142.
[3] A. Yala, P.G. Mikhael, F. Strand, G. Lin, K. Smith, Y.L. Wan, L. Lamb, K. Hughes, C. Lehman, R. Barzilay, Toward robust mammography-based models for breast cancer risk, Sci. Transl. Med. 13 (2021) 1–12, https://doi.org/10.1126/scitranslmed.aba4373.
[4] F. Janan, M. Brady, RICE: a method for quantitative mammographic image enhancement, Med. Image Anal. 71 (2021), https://doi.org/10.1016/j.media.2021.102043.
[5] J.H. Yoon, E.K. Kim, Deep learning-based artificial intelligence for mammography, Korean J. Radiol. 22 (2021) 1225–1239, https://doi.org/10.3348/KJR.2020.1210.
[6] P.E. Jebarani, N. Umadevi, H. Dang, M. Pomplun, A novel hybrid K-means and GMM machine learning model for breast cancer detection, IEEE Access 9 (2021) 146153–146162, https://doi.org/10.1109/ACCESS.2021.3123425.
[7] S. Prakash, M.V. Kumar, R.S. Ram, M. Zivkovic, N. Bacanin, M. Antonijevic, Hybrid GLFIL enhancement and encoder animal migration classification for breast cancer detection, Comput. Syst. Sci. Eng. 41 (2022) 735–749, https://doi.org/10.32604/csse.2022.020533.
[8] N. Dhungel, G. Carneiro, A.P. Bradley, Deep structured learning for mass segmentation from mammograms, in: 2015 IEEE International Conference on Image Processing (ICIP), IEEE, 2015, pp. 2950–2954.
[9] D. Arefan, A.A. Mohamed, W.A. Berg, M.L. Zuley, J.H. Sumkin, S. Wu, Deep learning modeling using normal mammograms for predicting breast cancer risk, Med. Phys. 47 (2020) 110–118, https://doi.org/10.1002/mp.13886.
[10] J. Dabass, S. Arora, R. Vig, M. Hanmandlu, Mammogram image enhancement using entropy and CLAHE based intuitionistic fuzzy method, in: 2019 6th International Conference on Signal Processing and Integrated Networks, SPIN 2019, 2019, pp. 24–29, https://doi.org/10.1109/SPIN.2019.8711696.

[11] N. Kharel, A. Alsadoon, P.W.C. Prasad, A. Elchouemi, Early diagnosis of breast cancer using contrast limited adaptive histogram equalization (CLAHE) and Morphology methods, in: 2017 8th International Conference on Information and Communication Systems, ICICS 2017, 2017, pp. 120–124, https://doi.org/10.1109/IACS.2017.7921957.
[12] V.D.P. Jasti, A.S. Zamani, K. Arumugam, M. Naved, H. Pallathadka, F. Sammy, A. Raghuvanshi, K. Kaliyaperumal, Computational technique based on machine learning and image processing for medical image analysis of breast cancer diagnosis, Secur. Commun. Network. 2022 (2022), https://doi.org/10.1155/2022/1918379.
[13] X. Wang, I. Ahmad, D. Javeed, S.A. Zaidi, F.M. Alotaibi, M.E. Ghoneim, Y.I. Daradkeh, J. Asghar, E.T. Eldin, Intelligent hybrid deep learning model for breast cancer detection, Electronics (Switzerland) 11 (2022), https://doi.org/10.3390/electronics11172767.
[14] A. Akselrod-Ballin, M. Chorev, Y. Shoshan, A. Spiro, A. Hazan, R. Melamed, E. Barkan, E. Herzel, S. Naor, E. Karavani, G. Koren, Y. Goldschmidt, V. Shalev, M. Rosen-Zvi, M. Guindy, Predicting breast cancer by applying deep learning to linked health records and mammograms, Radiology 292 (2019) 331–342, https://doi.org/10.1148/radiol.2019182622.
[15] T. Kyono, F.J. Gilbert, M. van der Schaar, Improving workflow efficiency for mammography using machine learning, J. Am. Coll. Radiol. 17 (2020) 56–63, https://doi.org/10.1016/j.jacr.2019.05.012.
[16] D. Oyewola, D. Hakimi, K. Adeboye, M.D. Shehu, Using five machine learning for breast cancer biopsy predictions based on mammographic diagnosis, International Journal of Engineering Technologies IJET 2 (2017) 142–145, https://doi.org/10.19072/ijet.280563.
[17] M.L. Giger, N. Karssemeijer, J.A. Schnabel, Breast image analysis for risk assessment, detection, diagnosis, and treatment of cancer, Annu. Rev. Biomed. Eng. 15 (2013) 327–357, https://doi.org/10.1146/annurev-bioeng-071812-152416.
[18] M.A. Elshafey, T.E. Ghoniemy, A hybrid ensemble deep learning approach for reliable breast cancer detection, International Journal of Advances in Intelligent Informatics 7 (2021) 112.
[19] S. Zhang, L. Yao, A. Sun, Y. Tay, Deep learning based recommender system: a survey and new perspectives, ACM Comput. Surv. 52 (2019), https://doi.org/10.1145/3285029.
[20] T. Liu, J. Huang, T. Liao, R. Pu, S. Liu, Y. Peng, A hybrid deep learning model for predicting molecular subtypes of human breast cancer using multimodal data, Irbm 43 (2022) 62–74, https://doi.org/10.1016/j.irbm.2020.12.002.
[21] G. Jayandhi, J.S. Leena Jasmine, R. Seetharaman, S.M. Joans, R. Priscilla Joy, Efficient breast cancer prediction using hybrid deep learning in mammographic images, in: Proceedings of the International Conference on Electronics and Renewable Systems, ICEARS 2022, 2022, pp. 1366–1371, https://doi.org/10.1109/ICEARS53579.2022.9752046.
[22] L. Narayanan, S. Krishnan, H. Robinson, A hybrid deep learning based assist system for detection and classification of breast cancer from mammogram images, Int. Arab J. Inf. Technol. 19 (2022).
[23] F. Yan, H. Huang, W. Pedrycz, K. Hirota, A disease diagnosis system for smart healthcare based on fuzzy clustering and battle royale optimization, Appl. Soft Comput. 151 (2024) 111123, https://doi.org/10.1016/J.ASOC.2023.111123.
[24] J. de Nazaré Silva, A.O. de Carvalho Filho, A. Corrêa Silva, A. Cardoso de Paiva, M. Gattass, Automatic detection of masses in mammograms using quality threshold clustering, correlogram function, and SVM, J Digit Imaging 28 (2015) 323, https://doi.org/10.1007/S10278-014-9739-3.
[25] W. Borges De Sampaio, A. Corrêa Silva, A. Cardoso De Paiva, M. Gattass, Detection of masses in mammograms with adaption to breast density using genetic algorithm, phylogenetic trees, LBP and SVM, Expert Syst. Appl. 42 (2015) 8911–8928, https://doi.org/10.1016/j.eswa.2015.07.046.
[26] N. Dhungel, G. Carneiro, A.P. Bradley, Deep structured learning for mass segmentation from mammograms, in: 2015 IEEE International Conference on Image Processing (ICIP), IEEE, 2015, pp. 2950–2954.
[27] M.A. Al-antari, M.A. Al-masni, M.T. Choi, S.M. Han, T.S. Kim, A fully integrated computer-aided diagnosis system for digital X-ray mammograms via deep learning detection, segmentation, and classification, Int J Med Inform 117 (2018) 44–54, https://doi.org/10.1016/J.IJMEDINF.2018.06.003.
[28] H. Nasir Khan, A. Raza Shahid, B. Raza, A.H. Dar, H. Alquhayz, Multi-View Feature Fusion Based Four Views Model for Mammogram Classification Using Convolutional Neural Network, (n.d.). https://doi.org/10.1109/ACCESS.2019.2953318.
[29] X. Yu, N. Zeng, S. Liu, Y.-D. Zhang, Utilization of DenseNet201 for Diagnosis of Breast Abnormality, vol. 30, 2019, pp. 1135–1144, https://doi.org/10.1007/s00138-019-01042-8.
[30] R. Song, T. Li, Y. Wang, Mammographic classification based on XGBoost and DCNN with multi features, IEEE Access 8 (2020) 75011–75021, https://doi.org/10.1109/ACCESS.2020.2986546.
[31] G. Altan, International Journal of INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING Deep Learning-based Mammogram Classification for Breast Cancer, Original Research Paper International Journal of Intelligent Systems and Applications in Engineering IJISAE 2020 (n.d.) 171–176. https://doi.org/10.1039/b000000x.
[32] G. Altan, Breast cancer diagnosis using deep belief networks on ROI images, Pamukkale University Journal of Engineering Sciences 28 (2022) 286–291, https://doi.org/10.5505/pajes.2021.38668.
[33] D.A. Ragab, O. Attallah, M. Sharkas, J. Ren, S. Marshall, A framework for breast cancer classification using Multi-DCNNs, Comput. Biol. Med. 131 (2021), https://doi.org/10.1016/J.COMPBIOMED.2021.104245.
[34] L. Narayanan, S. Krishnan, H. Robinson, A hybrid deep learning based assist system for detection and classification of breast cancer from mammogram images, Int. Arab J. Inf. Technol. 19 (2022).
[35] F. Yan, H. Huang, W. Pedrycz, K. Hirota, Automated breast cancer detection in mammography using ensemble classifier and feature weighting algorithms, Expert Syst. Appl. 227 (2023) 120282, https://doi.org/10.1016/J.ESWA.2023.120282.
[36] L.-K.-N. Nguyen, C. Kumar, B. Jiang, N. Zimmermann, L.-K.-N. Nguyen, C. Kumar, B. Jiang, N. Zimmermann, Implementation of systems thinking in public policy: a systematic review, Systems 11 (2023) 64, https://doi.org/10.3390/SYSTEMS11020064, 11 (2023) 64.
[37] J. Heinig, R. Witteler, R. Schmitz, L. Kiesel, J. Steinhard, Accuracy of classification of breast ultrasound findings based on criteria used for BI-RADS, Ultrasound Obstet. Gynecol. 32 (2008) 573–578, https://doi.org/10.1002/UOG.5191.
[38] M. Mustra, M. Grgic, R.M. Rangayyan, Review of recent advances in segmentation of the breast boundary and the pectoral muscle in mammograms, Med. Biol. Eng. Comput. 54 (2016) 1003–1024, https://doi.org/10.1007/s11517-015-1411-7.
[39] I.K. Maitra, S. Nag, S.K. Bandyopadhyay, R. Cross, A novel edge detection algorithm for digital mammogram, International Journal of Information and Communication Technology Research 2 (2012) 207–215.
[40] V. Bhateja, M. Misra, S. Urooj, Human visual system based unsharp masking for enhancement of mammographic images, J. Comput. Sci. 21 (2017) 387–393, https://doi.org/10.1016/j.jocs.2016.07.015.
[41] M. Hu, J. Herng Chion, P.N. Suganthan, R.K. Katuwal, Ensemble deep random vector functional link neural network for regression, IEEE Trans Syst Man Cybern Syst 53 (2023) 2604–2615, https://doi.org/10.1109/TSMC.2022.3213628.
[42] L. Zhang, P.N. Suganthan, A comprehensive evaluation of random vector functional link networks, Inf. Sci. 367–368 (2016) 1094–1105, https://doi.org/10.1016/j.ins.2015.09.025.
[43] A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A.C. Berg, W.-Y. Lo, Segment Anything, 2023. ArXiv Preprint ArXiv:2304.02643.
[44] J. Ma, B. Wang, Segment Anything in Medical Images, 2023. ArXiv Preprint ArXiv:2304.12306.
[45] H. Huang, Z. You, H. Cai, J. Xu, D. Lin, Fast detection method for prostate cancer cells based on an integrated ResNet50 and YoloV5 framework, Comput Methods Programs Biomed 226 (2022) 107184.
[46] C.E. Cardenas, R.E. McCarroll, L.E. Court, B.A. Elgohari, H. Elhalawani, C.D. Fuller, M.J. Kamal, M.A.M. Meheissen, A.S.R. Mohamed, A. Rao, Deep learning algorithm for auto-delineation of high-risk oropharyngeal clinical target volumes with built-in dice similarity coefficient parameter optimization function, Int. J. Radiat. Oncol. Biol. Phys. 101 (2018) 468–478.
[47] T.H.H. Aldhyani, R. Nair, E. Alzain, H. Alkahtani, D. Koundal, Deep learning model for the detection of real time breast cancer images using improved dilation-based method, Diagnostics 12 (2022), https://doi.org/10.3390/diagnostics12102505.

[48] Y.J. Suh, J. Jung, B.J. Cho, Y.J. Automated breast cancer Suh, J. Jung, B.J. Cho, Automated breast cancer detection in digital mammograms of various densities via deep learning, J. Personalized Med. 10 (4) (2020) 1–11, https://doi.org/10.3390/jpm10040211detection.
[49] H.T. Nguyen, H.Q. Nguyen, H.H. Pham, K. Lam, L.T. Le, M. Dao, V. Vu, VinDr-Mammo: a large-scale benchmark dataset for computer-aided diagnosis in full-field digital mammography, Sci. Data 10 (2023) 1–10, https://doi.org/10.1038/s41597-023-02100-7.
[50] S.R. Kebede, F.G. Waldamichael, T.G. Debelee, M. Aleme, W. Bedane, B. Mezgebu, Z.C. Merga, Dual view deep learning for enhanced breast cancer screening using mammography, Sci. Rep. 14 (1) (2024) 1–15, https://doi.org/10.1038/s41598-023-50797-8, 14 (2024).
[51] B. Pardamean, T.W. Cenggoro, R. Rahutomo, A. Budiarto, E.K. Karuppiah, Transfer learning from chest X-ray pre-trained convolutional neural network for learning mammogram data, Procedia Comput. Sci. 135 (2018) 400–407, https://doi.org/10.1016/J.PROCS.2018.08.190.
[52] A.A. Hekal, A. Elnakib, H.E.D. Moustafa, Automated early breast cancer detection and classification system, Signal Image Video Process 15 (2021) 1497–1505, https://doi.org/10.1007/S11760-021-01882-W.
[53] F. Janan, M. Brady, RICE: a method for quantitative mammographic image enhancement, Med. Image Anal. 71 (2021), https://doi.org/10.1016/j.media.2021.102043.
[54] N. Dhungel, G. Carneiro, A.P. Bradley, A deep learning approach for the analysis of masses in mammograms with minimal user intervention, Med. Image Anal. 37 (2017) 114–128, https://doi.org/10.1016/J.MEDIA.2017.01.009.
[55] F. Gao, T. Wu, J. Li, B. Zheng, L. Ruan, D. Shang, B. Patel, SD-CNN: a shallow-deep CNN for improved breast cancer diagnosis, Comput Med Imaging Graph 70 (2018) 53–62, https://doi.org/10.1016/J.COMPMEDIMAG.2018.09.004.
[56] T. Liu, J. Huang, T. Liao, R. Pu, S. Liu, Y. Peng, A hybrid deep learning model for predicting molecular subtypes of human breast cancer using multimodal data, Irbm 43 (2022) 62–74, https://doi.org/10.1016/j.irbm.2020.12.002.
[57] J. Chu, H. Min, L. Liu, W. Lu, A novel computer aided breast mass detection scheme based on morphological enhancement and SLIC superpixel segmentation, Med. Phys. 42 (2015) 3859–3869, https://doi.org/10.1118/1.4921612.
[58] H. Liu, Y. Chen, Y. Zhang, L. Wang, R. Luo, H. Wu, C. Wu, H. Zhang, W. Tan, H. Yin, D. Wang, A Deep Learning Model Integrating Mammography and Clinical Factors Facilitates the Malignancy Prediction of BI-RADS 4 Microcalcifications in Breast Cancer Screening Artificial Intelligence AUC Area under the Receiver Operating Characteristic Curve BI-RADS, 2021, pp. 5902–5912, https://doi.org/10.1007/s00330-020-07659-y.
[59] Y. Qiu, Y. Wang, S. Yan, M. Tan, S. Cheng, H. Liu, B. Zheng, An initial investigation on developing a new method to predict short-term breast cancer risk based on deep learning technology, Medical Imaging 2016, Computer-Aided Diagnosis 9785 (2016) 978521, https://doi.org/10.1117/12.2216275.
[60] J.H. Yoon, E.K. Kim, Deep learning-based artificial intelligence for mammography, Korean J. Radiol. 22 (2021) 1225–1239, https://doi.org/10.3348/KJR.2020.1210.
[61] L. Shen, L.R. Margolies, J.H. Rothstein, E. Fluder, R. McBride, W. Sieh, Deep learning to improve breast cancer detection on screening mammography, Sci. Rep. 9 (1) (2019) 1–12, https://doi.org/10.1038/s41598-019-48995-4, 9 (2019).