A cascaded classification approach using transfer learning and feature engineering for improved breast cancer classification

Chokri Ferkous [a,1,*], Ouissal Fadel [b,2], Abderrahmane Kefali [b,3], Hayet-Farida Merouani [c,4]

[a] Laboratoire des Sciences et Technologies de l'Information et de la Communication LabSTIC, Université 8 Mai 1945 Guelma, Guelma, Algeria
[b] Department of Computer Science, Université 8 Mai 1945 Guelma, Guelma, Algeria
[c] Department of Computer Science, Université Badji Mokhtar, Annaba, Algeria
[1] ferkous.chokri@univ-guelma.dz; [2] fadelouissal@gmail.com; [3] kefali.abderrahmane@univ-guelma.dz; [4] hayet.merouani@univ-annaba.dz
* corresponding author

ARTICLE INFO

Article history
Received July 20, 2024
Revised October 20, 2025
Accepted November 10, 2025
Available online February 28, 2026

Keywords
Deep learning
Transfer learning
Cascaded classification
Mammography
Haralick features

ABSTRACT

The primary objective of this study is to design a cascaded classification framework that integrates deep-learning representations with handcrafted and clinical features to enhance the reliability and accuracy of breast cancer detection in mammographic screening. A multi-source mammography dataset comprising four databases was used to ensure diversity and reduce bias. The proposed system operates in two stages. In the first stage, transfer learning models (VGG16, ResNet50, and EfficientNet_B0) were evaluated using ROC-AUC, PR-AUC, calibration curves, and bootstrap confidence intervals. EfficientNet_B0, which achieved the best balance between discrimination and calibration, was selected as the feature extractor. In the second stage, the malignancy probability was combined with Haralick texture features, patient age, and breast density, and classified using SVM, Random Forest, MLP, Decision Tree, and Logistic Regression. Model robustness was verified through multi-run experiments (five random seeds) and subgroup analyses by age and density. Among the CNN models, EfficientNet_B0 yielded the best performance (accuracy = 0.9438, ROC-AUC = 0.944, PR-AUC = 0.960). In the second stage, although Random Forest achieved the highest accuracy (0.9556 ± 0.002), SVM obtained the highest mean ROC-AUC (0.980 ± 0.001) with stable accuracy (0.9539 ± 0.001) and the most significant p-values, indicating superior robustness and generalization. The proposed cascaded framework effectively combines deep, handcrafted, and clinical features to improve mammogram classification performance. The SVM-based model demonstrates strong calibration, stability, and subgroup consistency, highlighting its potential for deployment in computer-aided mammography screening systems that assist radiologists in early breast cancer detection.

© 2026 The Author(s).
This is an open access article under the CC–BY-SA license.

1. Introduction

Breast cancer, the most common cancer affecting women globally, poses a significant health threat. Timely detection and diagnosis of breast abnormalities are pivotal in enhancing patient outcomes and increasing survival rates. In recent years, significant advancements in medical imaging technologies, especially digital imaging and mammography, have revolutionised breast cancer screening and diagnosis. Digital mammography provides high-resolution imaging of breast tissue, enabling precise identification and analysis of abnormalities. These digital images can be processed and analysed using computer-aided

techniques, making them the gold standard for breast cancer screening. A computer-aided diagnostic (CAD) system has been developed to improve the accuracy and efficiency of breast abnormality identification, potentially leading to earlier and more precise diagnoses. These tools provide additional support for radiologists, offering a second pair of eyes to analyse mammograms and potentially enhancing diagnostic accuracy.

Researchers have increasingly explored machine learning (ML) solutions for breast cancer classification. These approaches show promise in supporting healthcare professionals by accurately classifying abnormalities in mammograms. Furthermore, the introduction of Deep Learning (DL) has advanced the discipline substantially. Convolutional Neural Networks CNNs autonomously learn to extract crucial information from medical images, enabling accurate classification of abnormalities. Their ability to capture intricate textures and deformations within images has significantly boosted the accuracy of lesion identification.

Research on mammography analysis was conducted by Chakravarthy and Rajaguru [1], who differentiated between abnormal and negative (normal) mammograms and further classified abnormal cases as benign or malignant. They used 570 digital mammograms from the CBIS-DDSM dataset, 322 from the MIAS database, and 179 from the INbreast dataset. Their approach, combining ResNet-18 features with the ICS-ELM (Improved Crow-Search Optimised Extreme Learning Machine) algorithm, outperformed existing methods, achieving accuracies of 97.19% on DDSM, 98.13% on MIAS, and 98.26% on INbreast.

Jaber et al. [2] introduced a CAD system that combines the AlexNet architecture with the grey-level co-occurrence matrix (GLCM) for extracting distinct texture features from breast tissue. To enhance precision, they employed an ensemble of MK-SVM. In their approach, they fused features extracted from mammography images, resulting in a feature vector comprising 4096 features from AlexNet and 20 textural features from the GLCM. The model achieved an accuracy of 96.26%, with a precision of 94%, a recall of 96%, and an F1-score of 95%. Li et al. [3] proposed an EfficientNet breast cancer classification model that incorporates the CBAM attention mechanism and a multiscale fusion method. The researchers found that combining EfficientNet_B5 with CBAM and multiscale fusion improves classification performance, offering valuable insights for future research in medical image processing and breast cancer classification. The results show that the accuracy of the EfficientNet model increased from 64.71% (baseline) to 70.58% (EfficientNet+CBAM) and 76.47% (EfficientNet+CBAM+ multiscale fusion). The accuracy achieved for step0 was 75.52%, step1 80.91%, step2 74.21% and step3 68.95%.

Jones et al. [4] aimed to assess whether combining handcrafted and automated features could enhance CAD performance. They gathered a dataset comprising 1535 (740 malignant and 795 benign). Two types of features were extracted from the ROI. The first set consisted of 40 radiomic features, encompassing statistical features describing pixel-intensity distributions and textural features illustrating the spatial arrangement of intensities. The second set included automated features generated from a VGG16 model via transfer learning. They employed two VGG16 models to extract automated features, one using pseudo-ROIs and the other using three stacked original ROIs without pre-processing. Using these features, five linear support vector machines (SVMs) were constructed. The fusion SVM, which integrated pseudo-ROIs, achieved the highest abnormalities classification performance with an accuracy of 70.4%.

Kumari and Jagadesh [5] addressed the challenge of feature extraction using an innovative approach, the Advanced Grey Level Cooccurrence Matrix (AGLCM), applied to pre-processed mammograms, and used machine learning algorithms for image classification. Their proposed process comprises four steps: mammograms acquisition, pre-processing (utilising Contrast Limited Advanced Histogram Equalisation), feature extraction (AGLCM, intensity, and shape-based features), and classification. In their experiments using the MIAS dataset (112 negative and 210 abnormal), they employed the eXtreme Gradient Boosting (XGBoost) classifier and compared it with other classifiers such as Random Forest (RF), K-Nearest Neighbour (KNN), Artificial Neural Networks (ANN), and Support Vector Machine (SVM). The combination of CLAHE, AGLCM, and XGBoost achieved a remarkable accuracy of 95.6%.

In this study, we present a novel CAD system that uses a hybrid approach to classify breast abnormalities as benign or malignant, combining machine learning and transfer learning. This system is innovative because of its well-thought-out design, which combines feature extraction and classification techniques. To appraise the system's robustness and performance, it is extensively tested on four datasets: RSNA, Mini-DDSM, INBreast, and CMMD. The paper follows the following organizational structure: Section 1 introduces the topic; Sections 2 and 3 detail the materials, architecture, experimental results, and discussion; and Section 4 concludes the work.

2. Method

The framework architecture proposed in this study is structured into five essential stages that collectively form a robust and systematic pipeline for mammogram image analysis. The process begins with **dataset collection**, where relevant mammogram images are gathered to ensure sufficient variability and representativeness. This is followed by mammogram preprocessing, which includes operations such as noise reduction, contrast enhancement, and normalization to improve image quality and prepare the data for reliable feature extraction. The first-stage classification then applies transfer learning using pretrained deep learning models to generate high-level representations and perform an initial categorisation of the images. Next, the model constructs the feature vector by aggregating both deep features and additional descriptors into a structured representation. The final classification stage utilizes Haralick texture features to refine decision-making, enabling a more discriminative and interpretable assessment of mammographic abnormalities. Detailed implementation of these stages is elaborated in the following subsections. For a comprehensive overview of the model's design and workflow, please refer to the graphical abstract shown in Fig. 1.

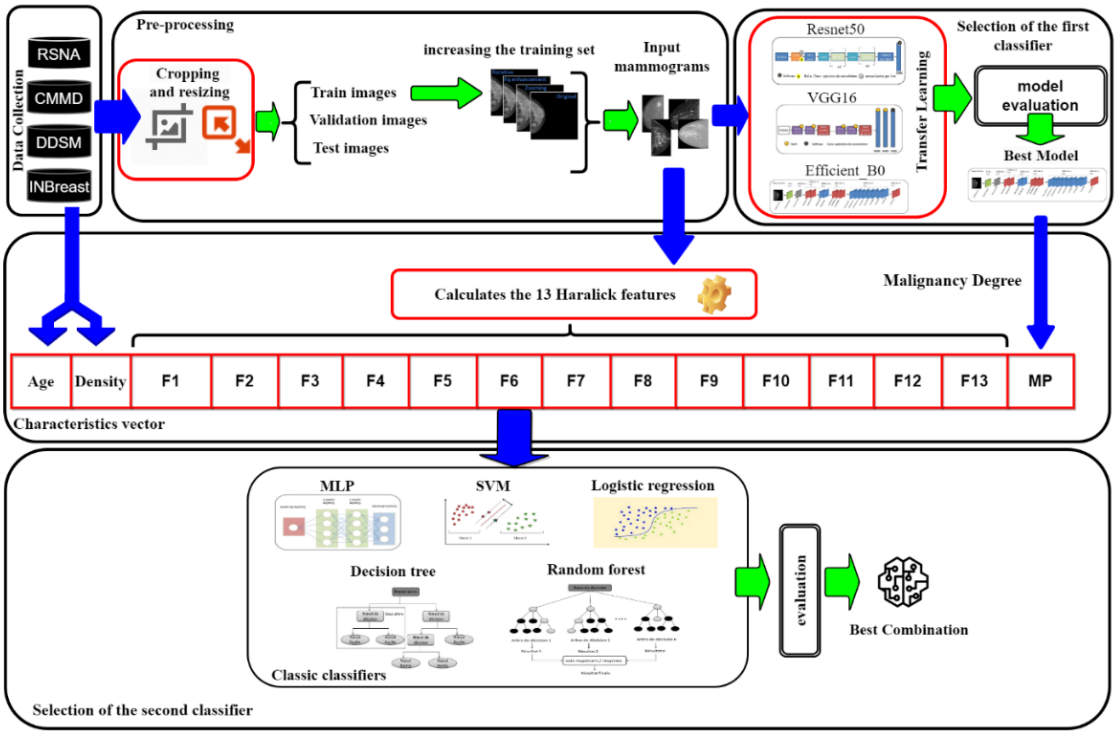

Fig. 1. Proposed CAD experimentation architecture

2.1. Dataset collection

The available datasets show a significant imbalance, with far fewer confirmed malignant abnormalities identified by biopsy than benign ones. To address this imbalance, we leveraged four datasets (RSNA, Mini-DDSM, INBREAST, and CMMD) to optimize the inclusion of malignant cases.

The Radiological Society of North America mammography dataset RSNA [6] includes radiographic images of women's breasts showing benign and malignant masses in both views. This collection contains mammograms from 11,913 patients. The database contains 53548 benign and 1158 malignant mammograms. CMMD or Chinese Mammography Database [7] contains mammograms from 1775 women presenting with benign or malignant breast abnormalities. Each abnormality is biopsy-confirmed and categorized as either benign or malignant. The dataset includes 4150 malignant and 1172 benign abnormalities. Mini-DDSM is a lighter version of the well-known DDSM database (Digital Database for Screening Mammography). However, the image files are in JPEG and PNG formats, generated with special software. This database includes 2716 malignant and 2684 benign mammograms [8]. The INBreast database [9] comprises 410 mammograms. These cases include 90 women with both breasts affected (each with four mammograms) and 25 mastectomy patients (each with two mammograms). The dataset incorporates various lesion types, including masses, calcifications, asymmetries, and distortions. Additionally, this database includes 343 malignant and 67 benign mammograms.

These datasets were gathered independently, in different countries and collection periods, and do not contain any overlapping or reposted cases. The principal dataset used in this study is RSNA, which initially had an unbalanced class distribution, with significantly more samples in the benign class than in the malignant class. To address this imbalance, additional cancer mammograms were collected from the DDSM, CMMD, and INBreast databases. This resulted in a final dataset that included mammograms from all four databases. There were specifically 8367 malignant and 8367 benign mammograms. The inclusion of images from various databases was intended to mitigate the over-learning issue observed when working with an unbalanced dataset.

2.2. Mammograms pre-processing

In the preprocessing step, we incorporate the following procedures:

2.2.1. The region of interest cropping

The ROI is an essential notion in medical image analysis. It refers to important or remarkable areas of an image. We perform mammogram cropping to eliminate unused portions of the mammograms, focusing on the ROI. This process is illustrated in Fig. 2, which shows the results of each step.

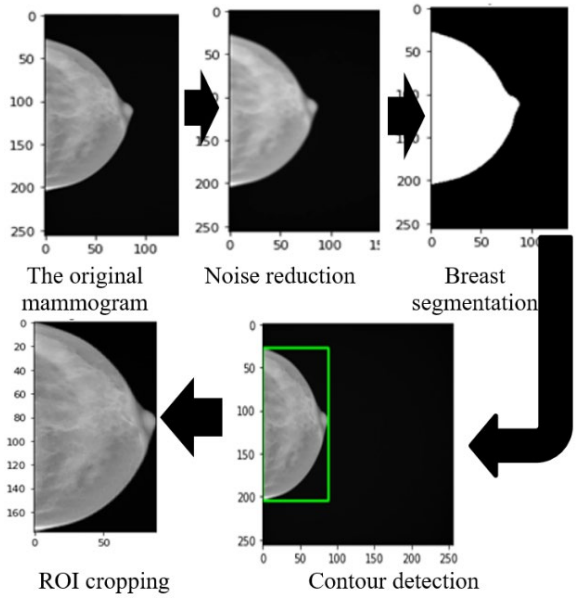

Fig. 2. Region-of-interest cropping process

The preprocessing stage of the proposed method involves a sequence of carefully designed steps to enhance image quality and isolate the relevant anatomical region. First, **noise reduction** is performed using a Gaussian filter, which smooths the mammogram, minimizing high-frequency noise while preserving important structural information. Next, **breast segmentation** is performed using Otsu's

thresholding to generate a binary mask that cleanly separates breast tissue from the background. Following segmentation, **contour detection** is applied to the binary mask to identify all possible contours, from which the contour with the largest surface area is selected as the true breast boundary. **ROI cropping** is performed by computing a bounding rectangle around the selected contour, then obtaining the coordinates of the upper-left corner and the region's width and height. These parameters define the precise area of interest (ROI) used for subsequent analysis.

2.2.2. ROI resizing

Images generally have different heights and widths after cropping. When images are supplied to a deep neural network for further analysis or classification, this variance in image size can prove problematic. It is advisable to resize images to a standard size to solve this problem. Resizing images to a uniform size ensures that the deep neural network can process them efficiently. This step is crucial when using deep learning models that require fixed input dimensions.

2.2.3. Dataset Split

To ensure the integrity of model evaluation, the dataset was split at the patient level. All images belonging to the same patient, including multiple mammographic views CC and MLO) and both laterality pairs (left and right breasts), were assigned to a single subset (training, validation, or testing). This procedure prevents any potential data leakage between subsets and ensures that the model's performance accurately reflects generalization to unseen patients. The preprocessed images are then divided into three subsets: training set (11712 mammograms), validation set (3348 mammograms), and test set (1674 mammograms).

2.2.4. Training subset augmentation

This process improves the robustness and generalization capability of the deep learning model. Two augmentation methods were used in particular: horizontal flip and random rotation. The data augmentation process is exclusively applied to the training subset, resulting in the creation of 23424 new mammograms. The objective of this process is to enhance the diversity and size of the training subset, thereby improving the generalization and robustness of machine learning models. Exposing our models to a broader range of abnormal variations during training enhances their adaptability and effectiveness in real-world scenarios.

2.3. First stage classification

Transfer learning has been widely utilized in the literature for breast cancer detection, demonstrating promising results. Among the well-known models used in CAD systems, notable mentions include:

VGG16 is a well-known convolutional neural network architecture known for its simplicity and effectiveness. It features a homogeneous topology, meaning all convolutional layers have the same filter size (either 2×2 or 3×3). This model has approximately 138 million parameters. VGG16 is the most frequently utilised architecture in the literature for CAD systems in breast cancer classification. Several studies have employed the VGG16 for breast cancer classification, including works [4], [10].

ResNet uses up to 152 layers, which can be very deep. It consists of 25 million parameters. ResNet supports hundreds of convolutional layers. ResNet's basic building blocks are the convolution and identity blocks. The convolution block uses the ReLU activation function and batch normalization. Among the studies that integrated the ResNet50 architecture for breast abnormalities classification are those conducted by [1], [11]. The EfficientNet-B0 model uses multiple convolution layers with a 3×3 receptive field and the reverse bottleneck moving convolution layer (MBconv) to capture features across layers. It accepts a 224×224×3 input image. EfficientNet-B0 has been employed in numerous studies on CAD systems for breast cancer [3], [12], [13].

The first stage of the proposed framework focused on evaluating and selecting the most suitable transfer learning model for mammogram classification, comparing EfficientNet_B0, ResNet, and VGG16. The analysis encompassed standard performance metrics, ROC-AUC and PR-AUC, calibration quality, and robustness across multiple training runs.

The three models were tested in this context. We used pre-trained versions of these models to leverage transfer learning. To retain the learned features of the three models, the pre-trained layers were frozen. A flattening layer, a dense layer with 256 neurons using the rectified linear unit activation function, a dropout layer to prevent overfitting using a rate of 0.5, and a final dense layer using a sigmoid activation function were added to our pre-trained models. We aimed to compare the performance of these three models when tested to determine which was most effective for our classification problem. The three models were compiled using the Adam optimizer with a learning rate of 0.001. Binary cross-entropy was utilized as the loss function.

Operating points were selected based on clinical screening priorities, favoring higher recall to minimize missed malignancies. The thresholds chosen for each model (EfficientNet_B0: 0.4, ResNet: 0.5, VGG16: 0.4) reflect an optimized trade-off between sensitivity and specificity suitable for pre-screening applications.

Among the three models, EfficientNet_B0 consistently achieved the best overall performance, with an accuracy of 0.9438, precision of 0.9895, recall of 0.8973, and F1-score of 0.9411. The confusion matrices of the three models are given in Fig. 3. These results suggest that EfficientNet_B0 effectively balances sensitivity and specificity, which is particularly important in mammography screening, where minimizing false negatives (missed malignancies) is critical.

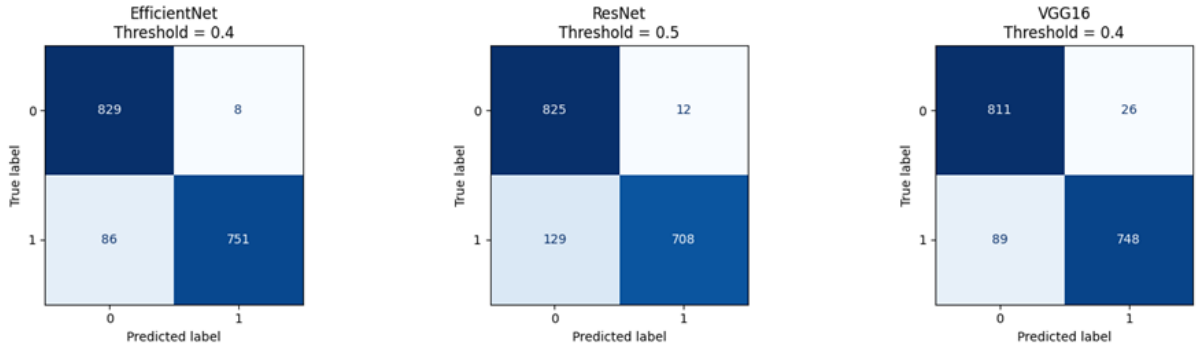

Fig. 3. The confusion matrices of EfficientNet_B0, VGG16 and ResNet

Table 1 summarises the tree models' performance, focusing on key metrics like accuracy, precision, recall, and F1-score.

Table 1. Transfer learning results

Measurements	VGG16	RESNET	EfficientNet-B0
Accuracy	93.13%	91.58%	94.38%
Recall	89.37%	84.59%	89.73%
Precision	96.64%	98.33%	98.95%
F1 score	92.86%	90.94%	94.11%

By comparison, VGG16 achieved competitive results (acc=0.9313, F1=0.9286) but demonstrated slightly lower recall (0.8937), implying a marginally reduced ability to detect malignant cases. ResNet, while still performant (acc=0.9158, F1=0.9094), lagged behind the other two architectures, suggesting that deeper residual connections might not have conferred a significant advantage given the dataset size and variability.

The ROC-AUC and PR-AUC metrics (Fig. 4 and Fig. 5) further demonstrate the superior performance of EfficientNet_B0 compared to the other evaluated models. In terms of ROC-AUC, EfficientNet_B0 achieves the highest score of 0.944, outperforming VGG16 (0.936) and ResNet (0.916), indicating its stronger ability to discriminate between malignant and benign cases across different classification thresholds. Similarly, the PR-AUC results reinforce this trend, with EfficientNet_B0 attaining a score of **0.960**, followed by VGG16 (**0.944**) and ResNet (**0.940**), showcasing its better

precision–recall balance, particularly under imbalanced data conditions typical of medical imaging. These metrics collectively validate EfficientNet_B0 as the most reliable model in this comparative analysis.

These results highlight EfficientNet_B0's stronger discriminative power and its ability to maintain high precision across a range of recall values, which is particularly valuable under class imbalance conditions where malignant samples are scarce. For the SVM classifier, the ROC-AUC and PR-AUC metrics were computed from probability estimates obtained via Platt scaling, which maps SVM decision scores to calibrated posterior probabilities via a sigmoid function. The calibration curves in Fig. 6 demonstrate how well each model's predicted probabilities align with the true positive case frequency. Ideally, a well-calibrated model produces predicted probabilities that reflect true event likelihoods.

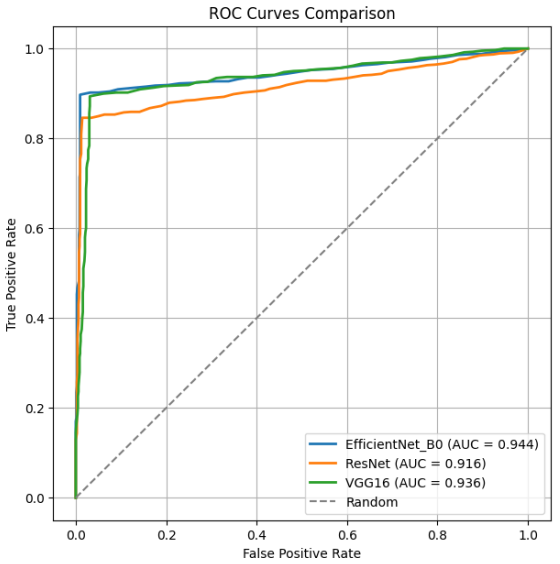

Fig. 4. The ROC curves of (EfficientNet_B0, VGG16 and ResNet).

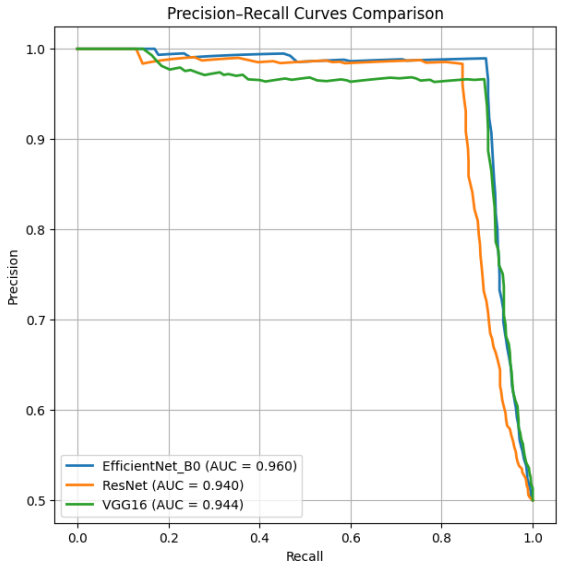

Fig. 5. The PR curves of (EfficientNet_B0, VGG16 and ResNet).

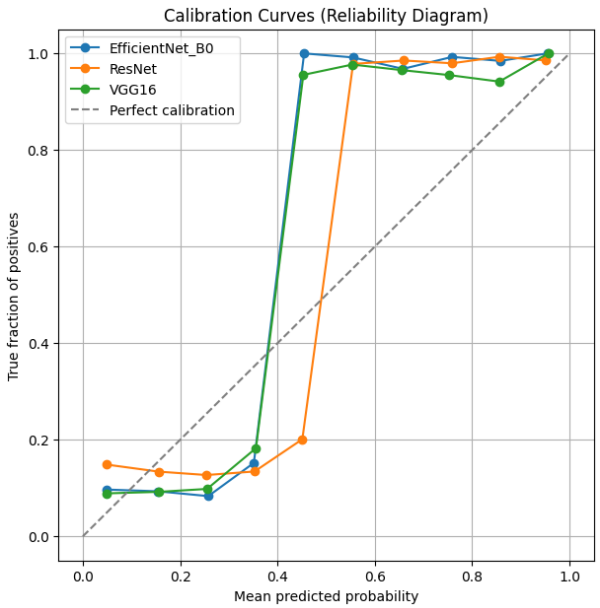

Fig. 6. The calibration curves of (EfficientNet_B0, VGG16, and ResNet).

EfficientNet_B0 exhibits near-perfect calibration in higher probability bins (≥0.45), with true positive fractions approaching 1.0. This indicates reliable probability estimates for high-confidence predictions.

VGG16 shows a similar trend, but with a slightly less smooth transition in mid-probability regions (0.25–0.35). ResNet tends to underpredict malignancy in the lower bins (true fraction lower than mean predicted probability), implying less reliable probability scaling at the low end. Overall, EfficientNet_B0 demonstrates the best probability calibration, an important characteristic for risk-based screening systems where threshold tuning is critical. The bootstrap confidence intervals (95% CIs) for ROC-AUC provide insight into the statistical robustness of each model (Table 2).

Table 2. The bootstrap confidence intervals (95% CIs) for ROC-AUC.

Model	Mean_AUC	95%_CI_Lower	95%_CI_Upper
EfficientNet_B0	0.943754	0.931707	0.95691
ResNet	0.915202	0.896683	0.928241
VGG16	0.93593	0.922476	0.950662

EfficientNet_B0 not only achieved the highest mean ROC-AUC but also displayed the narrowest CI, indicating both superior performance and greater stability. The results across five random seeds reveal limited performance variability, confirming model reproducibility (Table 3).

Table 3. The results across five random seeds.

Model	ROC-AUC (Mean ± Std)	Accuracy (Mean ± Std)	p-value (vs. best model)
EfficientNet_B0	0.943±0.001	0.9422±0.002	-
RESNET	0.915±0.001	0.9140±0.001	6.43E-06
VGG16	0.935±0.001	0.9300±0.001	1.86E-07

The minor standard deviations confirm consistent results across runs, while the statistically significant p-values indicate that the differences between EfficientNet_B0 and the other models are not due to random chance. The first-stage results demonstrate that EfficientNet_B0 outperforms both VGG16 and ResNet across nearly all evaluation dimensions, including discriminative ability, calibration, stability, and statistical significance. Consequently, it was selected as the feature-extracting backbone for the subsequent cascaded classification stage. Its balanced precision-recall characteristics and reliable probability estimates make it the most suitable foundation for downstream integration with clinical and texture-based features (Haralick, age, and density).

2.4. Features vector

Before entering the final classification stage, which relies on the classical machine learning approaches, the feature vector is constructed. It incorporates:

2.4.1. Malignancy probability

In cascaded classification, the output of the initial classification stage serves as input to the subsequent stage. In this work, instead of using the predicted class, we utilize the probability of malignancy (MP) as the input for the second classifier. In this study, the malignancy probability (MP) generated by the EfficientNet_B0 model is interpreted as a continuous score ranging from 0 to 1, representing the model's confidence in the malignancy of a given abnormality. An MP value of 1 indicates that the model predicts an extremely high likelihood of malignancy, corresponding to a 100% malignant assessment. Conversely, an MP value of 0 indicates a very low likelihood of malignancy, indicating the abnormality is 100% benign. For intermediate values, such as MP = 0.6, the model estimates a 60% probability of malignancy, suggesting moderate suspicion that may warrant further clinical evaluation.

2.4.2. Age and the mammographic density

The patient's age and mammographic density are also considered, as they can influence the BI-RADS (Breast Imaging Reporting and Data System) classification [14]. Patient age and mammographic density, both clinical features, are derived from medical records or imaging data.

2.5. Haralick features

The feature vector also comprises 13 Haralick features [15]. Those features are derived from the normalized Grey-Level Co-occurrence Matrix (GLCM). This statistical method analyses the spatial distribution of grey levels within a specific area by constructing co-occurrence matrices. These matrices g(i,j,d, θ) capture the frequency of encountering pairs of pixels with specific grey levels (i and j) at a particular distance (d) and orientation (θ) from each other. When evaluating texture direction, the angle can generally only have the values 0°, 45°, 90° and 135°. Consequently, separate SGLDM matrices are produced for different θ and d values. This example (Fig. 7) demonstrates calculating the SGLDM for a specific case ($\theta = 0$ and $d = 1$).

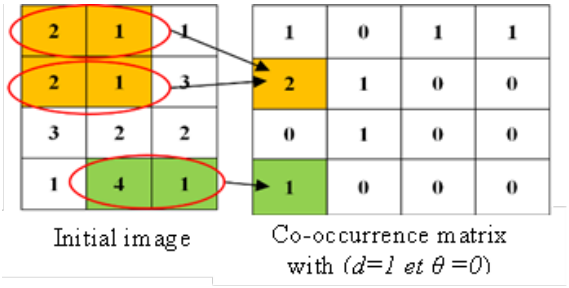

Fig. 7. The principle of the co-occurrence matrix.

The SGLDM matrix is of size G x G, where G represents the highest grey level. Within this matrix, the value at (2, 1) is 2. This indicates that there are two occurrences in the image where two horizontally adjacent pixels have grey levels of 2 and 1, respectively. Similarly, the value at (4, 1) is 1, indicating that only one instance exists in the image where two neighboring pixels have grey levels of 4 and 1, horizontally adjacent. Haralick features are calculated from co-occurrence matrices ($M^d_{i,j}$). In this work, we employed a set of 13 texture features (Table 4).

2.6. The Final Classification Stage

In the final classification stage of the cascaded model, the objective is to leverage the input vector to select the most appropriate classification method from some approaches. SVMs are a widely used machine learning method for both classification and regression tasks. SVMs classify data by finding a decision boundary in a high-dimensional space that maximizes the separation between different classes. This boundary is defined by strategically placing support vectors, which are data points closest to the boundary from each class. SVMs are extensively employed in breast abnormality classification, as demonstrated by the works of [16], [17], [18].

A decision tree visually depicts choices and their outcomes, resembling an actual tree structure. The branches represent decision rules based on specific criteria, while the nodes represent events or options. Each node group features belonging to a certain classification, with branches leading to different possibilities based on the node's value. [19]. Decision trees have emerged as a valuable tool for breast cancer classification due to their interpretability and ability to handle diverse data types. Studies like [20] achieved promising accuracy rates using decision trees, while [21] further improved performance by incorporating texture features. Random forests combine predictions from multiple decision trees, built in parallel, for improved accuracy. This reduces overlearning and increases prediction accuracy. Random forests work by building multiple diverse decision trees. Several studies have effectively leveraged random forests for breast cancer classification, as demonstrated by the work of [22], [23], [24]. Logistic regression is a widely used probabilistic statistical approach to classification problems in machine learning. The logistic function, commonly known as the sigmoid function, estimates the probability of an entry falling into a specific category. One of the advantages of logistic regression is its simplicity and efficiency, making it a popular option for many applications. Logistic regression is a powerful tool for classifying data into discrete categories [25]. it has been employed in numerous studies for CAD systems in breast cancer classification, such as [26], [27],[25].

Table 4. Haralick's textural features were used in this work.

N°	Feature	Function	
F1	Angular Second Moment	$ASM^d = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1}(M_{i,j}^d)^2$	(1)
F2	Entropy	$Entropie^d = -\sum_{i=0}^{G-1}\sum_{j=0}^{G-1}M_{i,j}^d \log_{M_{i,j}^d}$	(2)
F3	Contrast	$Cont^d = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1}((i-j)^2 \times M_{i,j}^d)$	(3)
F4	Correlation	$Cor^d = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1}M_{i,j}^d\left[\frac{(i-\mu_i^d)(j-\mu_j^d)}{\sigma_i^d\sigma_j^d}\right]$	(4)
F5	Inverse Difference Moment	$IDM^d = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1}\left[\frac{M_{i,j}^d}{1+(i-j)^2}\right]$	(5)
F6	Variance	$Var^d = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1}M_{i,j}^d(i-\mu^d)^2$	(6)
F7	Sum average	$SA^d = \frac{1}{2}\sum_{i=0}^{G-1}\sum_{j=0}^{G-1}M_{i,j}^d(i+j)$	(7)
F8	Sum entropy	$SE^d = -\sum_{i=0}^{G-1}\sum_{j=0}^{G-1}(i+j)M_{i,j}^d \log_{(i+j)M_{i,j}^d}$	(8)
F9	Sum variance	$SV^d = -\sum_{i=0}^{G-1}\sum_{j=0}^{G-1}\left[(i+j)-SE^d\right]^2 M_{i,j}^d$	(9)
F10	Difference entropy	$DE^d = -\sum_{i=0}^{G-1}\sum_{j=0}^{G-1}(i-j)M_{i,j}^d \log_{(i-j)M_{i,j}^d}$	(10)
F11	Difference variance	$DV^d = -\sum_{i=0}^{G-1}\sum_{j=0}^{G-1}\left[(i-j)-DE^d\right]^2 M_{i,j}^d$	(11)
F12	Information Measure of correlation 1	$IMC1 = \frac{(Entropie^d - HXY1)}{Max(HX,HY)}$ Where HX and HY are respectively the entropies of $M^d{}_i$ and $M^d{}_j$ And $HXY1 = -\sum_{i=0}^{G-1}\sum_{j=0}^{G-1}M_{i,j}^d \log_{((i\times j)M_{i,j}^d)}$	(12) (13)
F13	Information Measure of correlation 2	$IMC2 = (1-\exp^{-2(HXY2-Entropie^d)})^{1/2}$ Where $HXY2 = -\sum_{i=0}^{G-1}\sum_{j=0}^{G-1}(i\times j)M_{i,j}^d \log_{((i\times j)M_{i,j}^d)}$	(14) (15)

The most generalized neural network is called a multilayer perceptron (MLP). The fundamental aim of the multilayer perceptron is to develop a model that accurately maps input data to output data using prior data; consequently, the model can be used to generate output data even when the required output is unknown. MLPs have established themselves as a popular tool for breast abnormality classification, as evidenced by the research of [28], [29].

3. Results and Discussion

A CSV file was then generated to consolidate all extracted information, serving as a structured dataset for the final classification stage. This file includes 13 Haralick features, which describe the spatial distribution and textural patterns of grey levels within the mammogram, offering valuable insight into underlying tissue characteristics. Additionally, the patient's age is incorporated as an important clinical factor known to influence breast cancer risk. The dataset also records mammographic density, a key breast tissue attribute that can affect the visibility and interpretation of abnormalities on imaging. Finally, the file includes the malignancy probability (MP) predicted by the EfficientNet_B0 model, providing a

deep-learning–derived estimate of cancer likelihood. Together, these features form a comprehensive and multimodal input set for the final classification process.

The second stage of the proposed framework aimed to leverage both malignancy probability from EfficientNet_B0 and complementary clinical and texture-based descriptors (Haralick features, age, and mammographic density) to enhance classification performance. The outputs of the best-performing model from the first stage (EfficientNet_B0) were integrated within a cascaded classification architecture, and multiple classical supervised learning techniques were compared to determine the optimal final classifier. The same training set of the first stage is used to train five different machine learning approaches: MLP, SVM, Logistic Regression, Decision Tree, and Random Forest. Table 5 summarizes the results.

Table 5. Results of classical classifiers in cascade with EfficientNet_B0.

Models	Accuracy	Recall	Precision	F1 score
MLP	95.52%	91.40%	99.61%	95.33%
SVM	95.58%	91.40%	**99.74%**	95.39%
Random Forest	**95.70%**	91.76%	99.61%	**95.52%**
Logistic regression	95.46%	91.64%	99.22%	95.28%
Decision Trees	95.34%	**92.35%**	98.22%	95.20%

All classifiers demonstrated excellent performance, indicating that the hybrid feature representation substantially improved discrimination between benign and malignant mammograms. The confusion matrices of the three models are given in Fig.8.

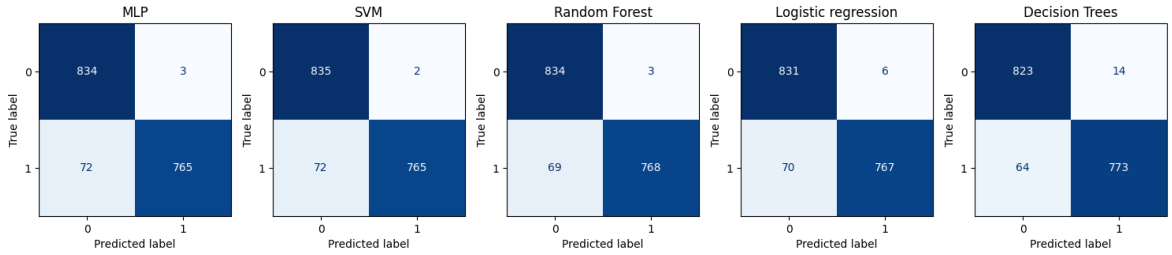

Fig. 8. The confusion matrices of the 5 classifiers.

Random Forest (RF) achieved the highest overall performance with an accuracy of 0.9570, precision of 0.9961, recall of 0.9176, and F1-score of 0.9552, followed closely by SVM and MLP. SVM reached comparable results (acc=0.9558, F1=0.9539), while MLP, Logistic Regression (LR), and Decision Trees (DT) also maintained strong accuracy (≥0.953).

This marginal performance gap reflects the robustness of the constructed feature space, which effectively captures both morphological and contextual aspects of breast tissue. Importantly, the high recall values across all models confirm the framework's suitability for screening scenarios, where sensitivity to malignant cases is a critical requirement. The discriminative ability of all models was evaluated using ROC-AUC and PR-AUC metrics. SVM achieved the best results in both measures (ROC-AUC = 0.981, PR-AUC = 0.985), closely followed by Random Forest (ROC-AUC = 0.978, PR-AUC = 0.982). The remaining models (MLP, Logistic Regression, and Decision Trees) obtained nearly identical performance, with ROC-AUC values around 0.976 and PR-AUC between 0.979 and 0.981.

These high AUC scores indicate that all models can effectively distinguish between malignant and benign cases across varying decision thresholds, reinforcing the effectiveness of the cascaded framework in enhancing separability. The calibration curves and their numerical equivalents reveal how well predicted probabilities align with actual malignancy frequencies. Across models, calibration quality improved significantly compared with the first stage.

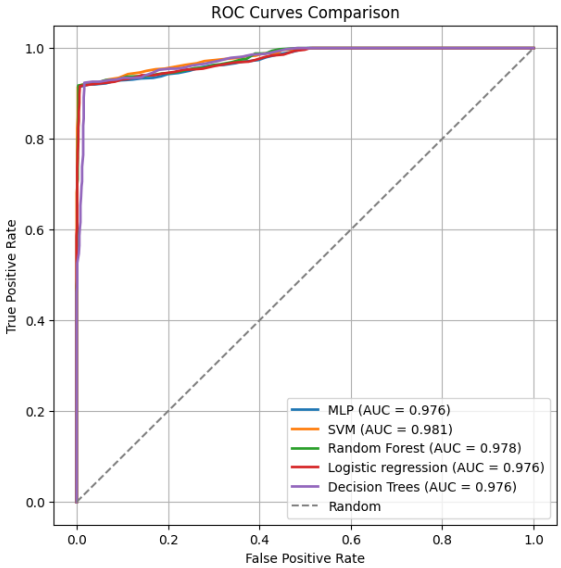

Fig. 9. The ROC curves of the 5 classifiers.

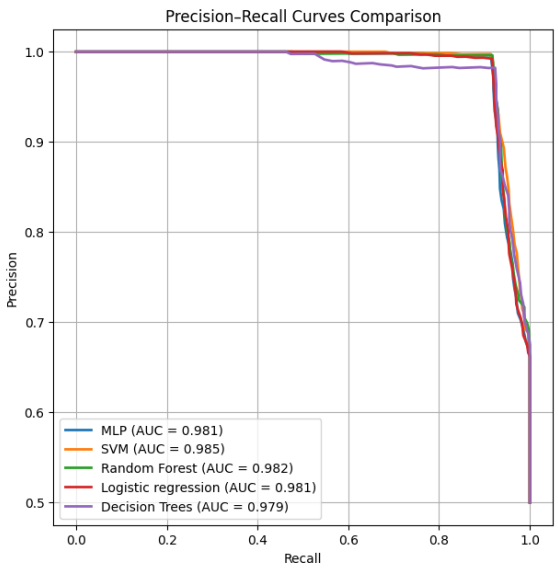

Fig. 10. The PR curves of the 5 classifiers.

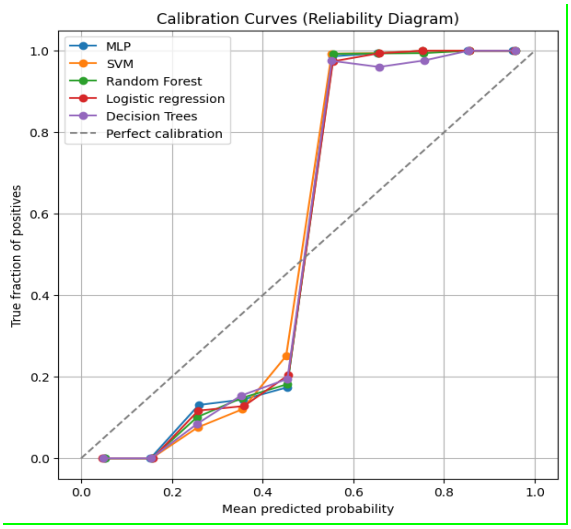

Fig. 11. The calibration curves of the 5 classifiers.

SVM and Random Forest show well-aligned probability estimates for higher probability bins (≥0.5), with true positive fractions close to 1.0, suggesting reliable confidence levels in positive predictions. MLP and Logistic Regression also exhibit good calibration behavior, though with a slight underestimation in mid-range bins (0.25–0.45). Decision Trees remain slightly less stable in the mid-probability range but converge well at higher confidence thresholds.

Overall, SVM demonstrated the most consistent calibration, an essential property for screening applications in which decision thresholds may need to be tuned to risk tolerance. The 95% bootstrap confidence intervals further confirm the statistical reliability of the performance metrics:

Table 6. The bootstrap confidence intervals (95% CIs) for ROC-AUC.

Model	Mean ROC-AUC	95% CI Lower	95% CI Upper
SVM	**0.9803**	0.9749	0.9849
Random Forest	0.9777	0.9703	0.9838
MLP	0.9756	0.9696	0.9804
Logistic Regression	0.9757	0.9693	0.9812
Decision Trees	0.9755	0.9692	0.9809

The narrow confidence intervals across all classifiers indicate low variance and strong statistical confidence in the reported results. The SVM model's upper bound (0.9849) further emphasizes its superior discriminative stability. To evaluate robustness, each model was trained using five random seeds. Results confirm minimal variability across runs.

Table 7. The bootstrap confidence intervals (95% CIs) for ROC-AUC.

Model	ROC-AUC (Mean ± Std)	Accuracy (Mean ± Std)	p-value (vs. best model)
MLP	0.974±0.001	0.9533±0.001	2.51E-05
SVM	**0.980±0.001**	0.9539±0.001	-
Random Forest	0.977±0.001	**0.9556±0.002**	2.89E-03
Logistic regression	0.974±0.001	0.9524±0.001	4.50E-04
Decision Trees	0.974±0.001	0.9511±0.001	2.17E-05

Overall, all classifiers demonstrated strong performance, with accuracy values ranging from 95.1% to 95.7% and ROC-AUC values of 0.97–0.98, indicating that the fusion of the malignancy probability from the first-stage classifier and handcrafted features significantly enhanced discriminative capability compared to using transfer learning alone. Among them, the Random Forest achieved the highest mean accuracy (0.9556 ± 0.002). At the same time, the SVM achieved the best overall ROC-AUC (0.980 ± 0.001) and exhibited the most stable performance across runs (the lowest standard deviation and the most significant p-values among the models).

Calibration analysis confirmed that SVM and Random Forest provided the most reliable probability estimates, with predicted probabilities closely matching the true positive fraction. Moreover, multi-run variability analysis (using five random seeds) highlighted that SVM consistently maintained superior robustness and statistical significance (p-values < 0.01 versus other models), which is crucial for clinical reliability. Accuracy performance across mammographic density categories (A–D) revealed notable consistency (Table 8).

Table 8. Classification accuracy by breast density level.

Model	A	B	C	D
MLP	96.97%	94.19%	**97.16%**	90.91%
SVM	**98.79%**	94.45%	96.86%	87.88%
Random Forest	**98.79%**	**95.23%**	96.26%	87.88%
Logistic regression	**98.79%**	94.58%	96.41%	87.88%
Decision Trees	**98.79%**	94.32%	96.11%	**90.91%**

Performance remains high across all densities, with slightly lower accuracy in the high-density category D, consistent with the known diagnostic difficulty in dense breasts due to masking effects. The SVM and Random Forest models maintained the best overall performance and stability across categories. Subgroup evaluation by age (Table.9) demonstrated high and stable accuracy across most ranges, with a modest decrease in older age groups (>75 years):

Table 9. Classification accuracy according to age range.

Models	≥45 years	45–60 (years)	60–75 (years)	>75 years
MLP	96.32%	96.38%	96.58%	84.25%
SVM	96.32%	96.38%	96.58%	85.04%
Random Forest	97.99%	95.48%	96.41%	88.19%
Logistic regression	97.32%	94.87%	96.58%	88.98%
Decision Trees	95.99%	95.93%	95.90%	88.19%

These results suggest that model performance remains robust across diverse age ranges, though reduced accuracy in older patients likely reflects smaller sample sizes (127 mammograms) and greater tissue compositional variability. SVM achieved the best trade-off between discriminative power,

calibration reliability, and statistical robustness. Random Forest provided marginally higher accuracy but with slightly higher variability. All models exhibited strong calibration and generalization across density and age subgroups. The integration of EfficientNet_B0 with handcrafted and demographic attributes significantly improved performance compared to the first stage, confirming the strength of the cascaded architecture for clinical screening applications.

In conclusion, the second-stage experiments confirm that combining multiple source features enhances the reliability of breast cancer classification. The SVM classifier, with its margin-based decision boundaries and excellent calibration, emerges as the optimal final model.

This model not only maximizes diagnostic accuracy but also maintains consistency across clinically relevant subgroups, making it well-suited for deployment in automated mammographic screening pipelines. The results unequivocally demonstrate the superior performance of the cascade classification model over individual classifiers, particularly with the EfficientNet_B0 and Random Forest classifiers in cascade. Our approach achieves a significant accuracy improvement, reaching 95.70%, a +1.31% gain over efficientNet_B0 (Table 10).

Table 10. Results comparison with state-of-the-art methods.

Reference	Used dataset	Dataset size	Accuracy
Chakravarthy & Rajaguru, 2022 [1]	CBIS_DDSM+ MIAS + INbreast	1071	97.19% for DDSM , 98.13% for MIAS, 98.26% for INbreast
Alyami et al., 2022 [2]	MIAS	322	96.26%
Li et al., 2021 [3]	Private dataset	300	76.47%
Jones et al., 2022 [4]	Private dataset	1535	70.4%,
Kumari and Jagadesh, 2022 [5]	MIAS	322	95.6%,
Our work	**Mini-DDSM INbreast CMMD RSNA**	**16734**	**95.70%**

Several previous works, including those by Li et al. [3], Alyami et al.[2], and Kumari & Jagadesh [5] have utilised limited datasets. This limitation can lead to overfitting. While fine-tuning pre-trained models on smaller datasets is a powerful transfer learning technique, finding the optimal balance during this process can be difficult, especially when data is limited. This can lead to model instability and low accuracy, as seen in Jones et al. [4], where accuracy didn't exceed 70.4%.

Transfer learning, particularly with deep neural networks, typically requires large datasets to effectively capture and transfer meaningful features. The scarcity of such datasets in breast cancer classification can hinder the model's ability to generalize well. While Chakravarthy & Rajaguru [1] achieved 98.26% accuracy using selected mammograms from three datasets (Table 11), this approach may not be scalable or generalizable due to the data selection strategy.

As summarized in Table 11, prior studies [1]–[5] have explored various combinations of handcrafted, deep, or hybrid models for mammogram classification. Most methods rely on either handcrafted texture features fused with shallow classifiers [2], [5], or deep features optimized or fused with metaheuristics or radiomics descriptors [1], [4]. Li et al. [3] integrated attention and multi-scale fusion but did not explore decision-level cascading. In contrast, our work introduces a two-stage cascade in which the EfficientNet-B0 output probability is fed into SVM or decision tree models. This architecture leverages the representational power of CNNs while enabling interpretable decision-level reasoning, thereby strengthening originality compared with prior hybrid or single-stage designs. Our proposed methodology addresses these challenges by leveraging:

- This work presents a computer-assisted diagnosis system inspired by radiological examinations, leveraging the BI-RADS lexicon for feature selection.
- Our approach leverages well-established classification methods and well-documented feature extraction techniques that align with the BI-RADS lexicon and have a demonstrably positive

impact on classification accuracy, as evidenced by prior research. Notably, integrating age and density features has been shown to significantly improve results, an aspect often overlooked in existing computer-aided diagnosis (CAD) systems in this field.

- In this study, cascaded classification has proven effective in the context of breast cancer classification. This approach combines the strengths of multiple classifiers, enabling them to learn from each other's errors and potentially achieve higher accuracy than any single classifier alone.
- All malignant mammograms from the four datasets are comprehensively included, without exception. To achieve a balanced final dataset, benign mammograms are randomly selected from the same datasets.

Table 11. Comparative summary of architectural differences between our proposed model and related works.

Ref	Model / Architecture	Feature Extraction Strategy	Key Architectural
Chakravarthy & Rajaguru, 2022 [1]	ResNet-18 + Improved Crow-Search Optimized ELM (ICS-ELM)	Deep features from ResNet-18	Hybrid deep + metaheuristic optimisation.
Alyami et al., 2022 [2]	AlexNet + GLCM texture fusion + MK-SVM	CNN deep features + handcrafted texture features (GLCM)	Cloud framework focus; handcrafted + CNN fusion.
Li et al., 2021 [3]	EfficientNet + CBAM + multi-scale fusion	Deep features with attention and multi-scale fusion	Uses attention mechanism and scale fusion.
Jones et al., 2022 [4]	VGG16 + SVM	Deep (VGG16) + handcrafted radiomic features	Feature-level fusion of deep + radiomics.
Kumari and Jagadesh, 2022 [5]	CLAHE + AGLCM + XGBoost	Texture-based handcrafted features	Purely handcrafted pipeline.
Our work	**Cascaded EfficientNet-B0 + Decision Tree/SVM**	**Deep CNN features or malignancy probability + Haralik features**	**End-to-end cascade integrating CNN probability output with Haralick and clinical features.**

Our strategy achieved 95.70% accuracy in breast cancer classification. This result suggests its potential effectiveness in addressing the challenges of limited data.

4. Conclusion

Beyond performance metrics, patient safety remains a central concern in automated breast cancer screening. False negatives may delay essential treatment, while false positives can expose patients to emotional distress and unnecessary medical interventions. The strong calibration and sensitivity of our proposed framework help mitigate these risks, providing more trustworthy malignancy predictions to assist radiologists in clinical decision-making. This study presented a two-stage cascaded classification framework designed to enhance the accuracy and reliability of breast cancer screening from mammographic images. The first stage employed transfer learning models (EfficientNet_B0, ResNet, and VGG16) to extract the malignancy probabilities from mammograms. Among these, EfficientNet_B0 demonstrated the best overall performance (ROC-AUC = 0.944 ± 0.001, Accuracy = 0.942 ± 0.020) and was selected for subsequent processing. In the second stage, the output of the selected deep model was combined with Haralick texture features, mammographic density, and patient age to form a comprehensive feature set. Various classical machine learning algorithms were then evaluated, including MLP, SVM, Random Forest, Logistic Regression, and Decision Tree. Results revealed that the SVM achieved the best overall balance between discriminative power and stability (ROC-AUC = 0.980 ± 0.001), with excellent calibration and robustness across multiple runs and subgroups. Although Random Forest achieved a marginally higher accuracy, the SVM was selected as the optimal classifier due to its

superior generalisation and consistent performance across different breast density categories and age groups. The proposed cascaded framework effectively integrates deep and handcrafted features with clinical parameters, leading to a marked improvement in classification performance. This approach addresses key challenges in breast cancer detection, particularly the scarcity of malignant samples and the variability in tissue density, offering a promising, interpretable solution for computer-aided screening systems. Future work will focus on expanding the dataset diversity and evaluating the system's clinical applicability in real-world screening workflows.

Declarations

Author contribution. All authors contributed equally to the main contributor to this paper. All authors read and approved the final paper.
Funding statement. Not Applicable.
Conflict of interest. The authors declare that they have no conflict of interest.
Additional information. No additional information is available for this paper.
Conflict of interest. The authors declare that they have no conflict of interest.
Additional information. No additional information is available for this paper.

Data and Software Availability Statements

This work leverages four publicly available datasets. Details on the datasets are provided in the following:

1. **IN-BREAST** : Data openly available in a public repository https://www.kaggle.com/datasets/quachnam/inbreast-roi-mammography/versions/1 , reference number [9].
2. **CMMD** : Data openly available in a public repository https://www.kaggle.com/datasets/tommyngx/cmmd2022, reference number [7].
3. **Mini-DDSM** : Data openly available in a public repository https://www.kaggle.com/datasets/cheddad/miniddsm2 , reference number [8].
4. **RSNA Screening Mammography Breast Cancer Detection:** Data openly available in a public repository https://www.kaggle.com/competitions/rsna-breast-cancer-detection/data , reference number [6].

References

[1] S. R. Sannasi Chakravarthy and H. Rajaguru, "Automatic Detection and Classification of Mammograms Using Improved Extreme Learning Machine with Deep Learning," *IRBM*, vol. 43, no. 1, pp. 49–61, Feb. 2022, doi: 10.1016/j.irbm.2020.12.004.

[2] J. Alyami *et al.*, "Cloud Computing-Based Framework for Breast Tumor Image Classification Using Fusion of AlexNet and GLCM Texture Features with Ensemble Multi-Kernel Support Vector Machine (MK-SVM)," *Comput. Intell. Neurosci.*, vol. 2022, no. Aug, pp. 1–9, Aug. 2022, doi: 10.1155/2022/7403302.

[3] R. Li, S. Wang, Z. Wang, and L. Zhang, "Breast cancer X-ray image staging: based on efficient net with multi-scale fusion and cbam attention," *J. Phys. Conf. Ser.*, vol. 2082, no. 1, p. 012006, Nov. 2021, doi: 10.1088/1742-6596/2082/1/012006.

[4] M. A. Jones, R. Faiz, Y. Qiu, and B. Zheng, "Improving mammography lesion classification by optimal fusion of handcrafted and deep transfer learning features," *Phys. Med. Biol.*, vol. 67, no. 5, p. 054001, Feb. 2022, doi: 10.1088/1361-6560/AC5297.

[5] L. K. Kumari and B. N. Jagadesh, "A Robust Feature Extraction Technique for Breast Cancer Detection using Digital Mammograms based on Advanced GLCM Approach," *EAI Endorsed Trans. Pervasive Heal. Technol.*, vol. 8, no. 30, pp. 1–10, Jan. 2022, doi: 10.4108/EAI.11-1-2022.172813.

[6] "RSNA Screening Mammography Breast Cancer Detection AI Challenge (2023) | RSNA," Radiological Society of North America. Accessed: Nov. 30, 2025. [Online]. Available at: https://www.rsna.org/artificial-intelligence/ai-image-challenge/screening-mammography-breast-cancer-detection-ai-challenge

[7] "CMMD | The Chinese Mammography Database," NATIONAL CANCER INSTITUTE CIP Cancer Imaging Program. [Online]. Available at: https://www.cancerimagingarchive.net/collection/cmmd/#citations

[8] C. D. Lekamlage, F. Afzal, E. Westerberg, and A. Chaddad, "Mini-DDSM: Mammography-based Automatic Age Estimation," in *ACM International Conference Proceeding Series*, Association for Computing Machinery, Nov. 2020, pp. 1–6. doi: 10.1145/3441369.3441370;PAGEGROUP:STRING:PUBLICATION.

[9] I. C. Moreira, I. Amaral, I. Domingues, A. Cardoso, M. J. Cardoso, and J. S. Cardoso, "INbreast: Toward a Full-field Digital Mammographic Database.," *Acad. Radiol.*, vol. 19, no. 2, pp. 236–248, Feb. 2012, doi: 10.1016/j.acra.2011.09.014.

[10] A. B. M. A. Hossain, J. K. Nisha, and F. Johora, "Breast Cancer Classification from Ultrasound Images using VGG16 Model based Transfer Learning," *Int. J. Image, Graph. Signal Process.*, vol. 15, no. 1, pp. 12–22, Feb. 2023, doi: 10.5815/IJIGSP.2023.01.02.

[11] S. Civilibal, K. K. Cevik, and A. Bozkurt, "A deep learning approach for automatic detection, segmentation and classification of breast lesions from thermal images," *Expert Syst. Appl.*, vol. 212, no. February, p. 118774, Feb. 2023, doi: 10.1016/J.ESWA.2022.118774.

[12] S. J. Frank, "A deep learning architecture with an object-detection algorithm and a convolutional neural network for breast mass detection and visualization," *Healthc. Anal.*, vol. 3, no. November, p. 100186, Nov. 2023, doi: 10.1016/J.HEALTH.2023.100186.

[13] J. Peta and S. Koppu, "Explainable Soft Attentive EfficientNet for breast cancer classification in histopathological images," *Biomed. Signal Process. Control*, vol. 90, no. April, p. 105828, Apr. 2024, doi: 10.1016/J.BSPC.2023.105828.

[14] D. A. Spak, J. S. Plaxco, L. Santiago, M. J. Dryden, and B. E. Dogan, "BI-RADS® fifth edition: A summary of changes," *Diagn. Interv. Imaging*, vol. 98, no. 3, pp. 179–190, Mar. 2017, doi: 10.1016/J.DIII.2017.01.001.

[15] R. M. Haralick, I. Dinstein, and K. Shanmugam, "Textural Features for Image Classification," *IEEE Trans. Syst. Man Cybern.*, vol. SMC-3, no. 6, pp. 610–621, 1973, doi: 10.1109/TSMC.1973.4309314.

[16] M. A. Aswathy and M. Jagannath, "An SVM approach towards breast cancer classification from H&E-stained histopathology images based on integrated features," *Med. Biol. Eng. Comput. 2021 599*, vol. 59, no. 9, pp. 1773–1783, Jul. 2021, doi: 10.1007/S11517-021-02403-0.

[17] H. S. Laxmisagar and M. C. Hanumantharaju, "FPGA implementation of breast cancer detection using SVM linear classifier," *Multimed. Tools Appl. 2023 8226*, vol. 82, no. 26, pp. 41105–41128, Mar. 2023, doi: 10.1007/S11042-023-15121-6.

[18] A. Kumari, M. Akhtar, M. Tanveer, and M. Arshad, "Diagnosis of breast cancer using flexible pinball loss support vector machine," *Appl. Soft Comput.*, vol. 157, no. May, p. 111454, May 2024, doi: 10.1016/J.ASOC.2024.111454.

[19] N. K. Al-Salihy and T. Ibrikci, "Classifying breast cancer by using decision tree algorithms," *ACM Int. Conf. Proceeding Ser.*, pp. 144–148, Feb. 2017, doi: 10.1145/3056662.3056716.

[20] M. M. Ghiasi and S. Zendehboudi, "Application of decision tree-based ensemble learning in the classification of breast cancer," *Comput. Biol. Med.*, vol. 128, no. January, p. 104089, Jan. 2021, doi: 10.1016/J.COMPBIOMED.2020.104089.

[21] K. Juneja and C. Rana, "An improved weighted decision tree approach for breast cancer prediction," *Int. J. Inf. Technol. 2018 123*, vol. 12, no. 3, pp. 797–804, Apr. 2018, doi: 10.1007/S41870-018-0184-2.

[22] T. Ananth Kumar, G. Rajakumar, and T. S. Arun Samuel, "Analysis of breast cancer using grey level co-occurrence matrix and random forest classifier," *Int. J. Biomed. Eng. Technol.*, vol. 37, no. 2, pp. 176–184, 2021, doi: 10.1504/IJBET.2021.119503.

[23] J. Quist, L. Taylor, J. Staaf, and A. Grigoriadis, "Random Forest Modelling of High-Dimensional Mixed-Type Data for Breast Cancer Classification," *Cancers 2021, Vol. 13, Page 991*, vol. 13, no. 5, p. 991, Feb. 2021, doi: 10.3390/CANCERS13050991.

[24] S. N. S., "Prediction of Breast Cancer Through Random Forest," *Curr. Med. Imaging*, vol. 19, no. 10, pp. 1144–1155, Sep. 2022, doi: 10.2174/1573405618666220930150625.

[25] D. Houfani *et al.*, "An Improved Model for Breast Cancer Diagnosis by Combining PCA and Logistic Regression Techniques," *Int. J. Comput. Digit. Syst.*, vol. 20, no. March, pp. 2210–142, 2021, [Online]. Available at: https://rodin.uca.es/bitstream/handle/10498/29738/SC_2023_0475.pdf?sequence=1&isAllowed=y.

[26] L. Liu, "Research on logistic regression algorithm of breast cancer diagnose data by machine learning," in *Proceedings - 2018 International Conference on Robots and Intelligent System, ICRIS 2018*, Institute of Electrical and Electronics Engineers Inc., Jul. 2018, pp. 157–160. doi: 10.1109/ICRIS.2018.00049.

[27] F. Morais-Rodrigues *et al.*, "Analysis of the microarray gene expression for breast cancer progression after the application modified logistic regression," *Gene*, vol. 726, no. February, p. 144168, Feb. 2020, doi: 10.1016/J.GENE.2019.144168.

[28] S. Boumaraf, X. Liu, C. Ferkous, and X. Ma, "A New Computer-Aided Diagnosis System with Modified Genetic Feature Selection for BI-RADS Classification of Breast Masses in Mammograms," *Biomed Res. Int.*, vol. 2020, no. 1, p. 7695207, Jan. 2020, doi: 10.1155/2020/7695207.

[29] Z. Guo, L. Xu, and N. Ali Asgharzadeholiaee, "A Homogeneous Ensemble Classifier for Breast Cancer Detection Using Parameters Tuning of MLP Neural Network," *Appl. Artif. Intell.*, vol. 36, no. 1, p. 21, Dec. 2022, doi: 10.1080/08839514.2022.2031820.