



OPEN ACCESS

EDITED BY

Antonino Musolino,
University of Bologna, Italy

REVIEWED BY

Filippo Merloni,
Istituto Tumori della Romagna IRST IRCCS,
Italy
Olga Serra,
Scientific Institute of Romagna for the Study
and Treatment of Tumors (IRCCS), Italy

*CORRESPONDENCE

Mohan L. Jayatilake
✉ jayatilml@pdn.ac.lk

RECEIVED 06 March 2025

ACCEPTED 08 May 2025

PUBLISHED 03 June 2025

CITATION

Paripooranan N, Nirasha WB, Perera HRP,
Vijithananda SM, Hewavithana PB,
Sherminie LPG and Jayatilake ML (2025)
Machine learning-based classification
model to differentiate subtypes of
invasive breast cancer using MRI.
Front. Oncol. 15:1588787.
doi: 10.3389/fonc.2025.1588787

COPYRIGHT

© 2025 Paripooranan, Nirasha, Perera,
Vijithananda, Hewavithana, Sherminie and
Jayatilake. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Machine learning-based classification model to differentiate subtypes of invasive breast cancer using MRI

Nadesalingam Paripooranan¹,
Warnakulasuriya Buddhini Nirasha², H. R. P. Perera²,
Sahan M. Vijithananda¹, P. Badra Hewavithana¹,
Lahanda Purage Givanthika Sherminie²
and Mohan L. Jayatilake^{2*}

¹Department of Radiology, Faculty of Medicine, University of Peradeniya, Peradeniya, Sri Lanka,

²Department of Radiography/Radiotherapy, Faculty of Allied Health Sciences, University of Peradeniya, Peradeniya, Sri Lanka

Introduction: Breast cancer is considered one of the most lethal diseases among women worldwide. Invasive Ductal Carcinoma (IDC) and Invasive Lobular Carcinoma (ILC) are the two most prominent subtypes of breast cancer. They differ in epidemiology, molecular alterations, and clinicopathological features. Patient treatment and management also differ due to these variations.

Aim: The study aimed to develop a predictive model to differentiate IDC and ILC using machine learning techniques based on the morphological features of the contralateral breast. Methods- 143 magnetic resonance imaging (MRI) images were sourced from the “DUKE Breast-Cancer” collection on the Cancer Imaging Archive website. Regions of interest were drawn on each slice to compute the morphological features of the contralateral breast using the 3D Slicer application. Supervised learning methods were applied to the morphological features to build a predictive model incorporating a Random Forest Classifier to differentiate IDC and ILC. Hyperparameters were tuned to optimize the model.

Results: The model was able to differentiate IDC and ILC with an accuracy of 79% and an Area Under the Curve of 0.851 on the Receiver Operating Characteristic Curve. Among the morphological features, the total volume of the contralateral breast, surface area of the contralateral breast, breast density, and the ratio of the total volume of the contralateral breast to its surface area had higher F-scores, indicating that the dimensions of the contralateral breast could be an important factor in differentiating IDC and ILC.

Conclusion: This study successfully developed and optimized a predictive model based on breast morphological features to differentiate IDC and ILC using machine learning methods.

KEYWORDS

machine learning, breast MRI, invasive breast cancer, invasive ductal carcinoma, invasive lobular carcinoma

1 Introduction

Breast cancer is the most common malignancy among women worldwide and also the second most common cancer overall, following only lung cancer (1). According to the World Health Organization (WHO), approximately 2.3 million women were diagnosed with breast cancer in 2020, and it accounted for an estimated 685,000 deaths globally (2). Despite advancements in early detection, it is projected that in 20 years, more than 60% of new breast cancer cases and 70% of deaths will occur in low and middle income countries (3).

Breast cancer is a heterogeneous disease, primarily originating from the inner lining of the breast's fibroglandular tissues, which are responsible for milk production and supply. Heterogeneity can be reflected in the diversity of histological and molecular subtypes of breast cancer (4). Invasive breast cancer (IBC) is one of the major histological categories of breast cancer and can be further divided into various subtypes. IBC cells are proficient in invading through the fibroglandular tissue wall into the surrounding connective and fatty tissues. However, it should be noted that IBC can present without necessarily spreading beyond the breast and adjacent lymph nodes (5). The study of the different characteristics of IBC subtypes can lead to understanding which strategies will be more effective in developing therapies that could improve patient outcomes.

Invasive ductal carcinoma (IDC) and invasive lobular carcinoma (ILC) are the two most common types of IBC that differ in their epidemiology, molecular alterations, clinicopathological features, and natural history (6). ILC is the second most common type after IDC. IDC occurs in approximately three-fourths of patients, while ILC occurs in less than one-tenth of patients (7). ILC is differentiated from IDC with increased frequency of multifocality, bilaterality, multicentricity, and difficulty in determining its margins during clinical examinations (8, 9). Past research has shown that ILC patients were older (10). This could be due to the low proliferative rate or

difficulty detecting ILC (11). Tumor size was found to be larger in ILC (11, 12). The incidence of contralateral breast cancer was also higher in ILC (11). The metastatic pattern was also found to be different for ILC and IDC since ILC is less likely to affect the lungs, pleura, and CNS compared to IDC (11). Due to these differences, patient management also differs between IDC and ILC. For ILC patients, conservation therapy is not recommended unless it is early-stage due to the multicentric, multifocal characteristic. Radiation therapy controls the microscopic foci of multicentric disease, as mentioned in several studies (13). The neoadjuvant setting was used in several retrospective studies, which have shown that patients with IDC benefit much more from chemotherapy than those with ILC (14, 15).

Breasts are composed of glandular, fibrous, and adipose tissues. The glandular tissues form the lobules and ducts, which are the primary components of milk production. The lobules are present in the background of fibrous, and adipose tissues. Adipose tissue plays an integral role in the morphogenesis of the breast and participates in mammary epithelial cell differentiation (16). Breast volume and breast density, two features of this study, are related to the dimensions of these tissues. Breast density refers to the ratio of fibroglandular tissues to fatty tissues of a patient's breast and may vary between individuals and over the course of a patient's life (17). It is considered an independent risk factor for breast cancer (18). Breast volume has been suggested as a controversial risk factor for the last few decades. The study by Li et al. found that breast cancer might not be entirely related to breast volume (19). But, according to the study by Schutt et al., breast size was correlated to the occurrence of breast cancer (20).

The total breast volume is calculated using several methods such as the water displacement method, three-dimensional Ultrasound scan, mammography, Computed Tomography, and Magnetic Resonance Imaging. Among these, MRI is considered the modality with the highest accuracy (16). According to recent studies, breast MRI is considered a superior modality for measuring fibroglandular tissue volume and breast density because of its ability to segment the whole breast from the body and segment fibroglandular tissues within the breast, slice by slice. Breast MRI has become a revolutionary investigative tool in the diagnosis of breast cancer due to its better performance compared to mammography and US (21). Breast MRI adds anatomical characterization, information on tumor vascularity and perfusion

Abbreviations: IBC, Invasive breast carcinoma; IDC, Invasive ductal carcinoma; ILC, Invasive lobular carcinoma; CLB, Contralateral breast; US, Ultra sound; MRI, Magnetic resonance imaging; CT, Computed tomography; DICOM, Digital Imaging and Communications in Medicine; TCIA- The Cancer Imaging Archive; ROI, Region of interest; ROC, Receiver operating characteristic AUC, Area under the curve RF, Random Forest.

dynamics, and excellent lesion detection. Even though breast MRI provides sensitivity above 88.1%, its specificity fluctuates between 83% and 98.4%, critically dependent on observer experience (22). The sensitivity of breast MRI is usually higher than 95%, making it vital for detection before symptoms emerge, especially in cases where other imaging modalities are not precise enough. Similarly, it has higher sensitivity in detecting breast cancers in denser breasts (21).

The main challenges in breast cancer screening and imaging diagnosis are the complexity and wide range of image features, varying quality of images, and inconsistent interpretations by different radiologists and medical institutions. Artificial Intelligence could be incorporated to increase the accuracy and precision of these interpretations. Furthermore, personalized predictions and recommendations derived from individual patient data through machine learning (ML) can optimize treatment strategies, enhancing overall patient care. By handling complex, high-dimensional datasets, ML algorithms can extract relevant features from MRI images, facilitating the creation of robust classification models. Several ML algorithms are popular in this field. The Random Forest Classifier (22–24), Support Vector Machine (25), and k-nearest neighbor (26) are some of the algorithms used in medical image processing and building predictive models. The Random Forest Classifier is a collection of decision trees where each tree is trained individually with sub datasets randomly, and the outputs of all the individual trees are combined to make a prediction (27). In the training phase, a large number of decision trees are created initially, which is called bootstrap aggregating. Each tree is trained on a random subset of the data and grows until it reaches its maximum depth. In the testing phase, each case is passed through all the trees in the forest and a class prediction is given for each tree. The final output is determined by the majority vote. Jacques Wainer conducted a comprehensive study on 14 classification algorithms to determine which algorithm performs best under specific conditions. Random Forest along with, Gradient Boosting Machines, and Radial Basis Function Support Vector Machines were the top performing algorithms. Random Forest was found to be more robust and easier to use. It was also found to perform better with high-dimensional datasets (28).

In recent years, many studies have been conducted using ML in the field of breast cancer. There were a few studies that focused on IDC and ILC differentiation using ML methods. In a study conducted in 2022 by Gunawan et al., a Support Vector Machine was used on 156 IDC and 8 ILC mammographic data to differentiate IDC and ILC (29). They selected 9 mammographic physical parameters of the image: entropy, contrast, MomentAng, MomentDiff, mean, deviation, EntropyHDiff, MomentHDiff, and MeanHdiff. Their model scored an accuracy of 76.56%. Maiti et al. did a study in 2024 in which he explored the texture features extracted from DCE-MRI images to differentiate IDC from ILC (30). These texture features included shape feature, gray level dependence matrix (GLDM), gray level co-occurrence matrix (GLCM), First order, gray level run length matrix (GLRLM), gray level size zone matrix (GLSZM), and neighboring gray tone difference matrix (NGTDM). The study was done on a dataset

that included 58 patients (30 with IDC and 28 with ILC). They acquired 0.998, 97.21%, and 96.2% for Area under Curve(AUC), sensitivity, and specificity respectively. Faraz et al., extracted a total of 342 radiomic features from MRI images(DCE-MRI, subtraction and T2 weighted images) and selected 32 relevant features to differentiate IDC from ILC (31). They achieved an AUC of 0.73 when using SVM to classify a dataset of MR images from 323 patients.

Previous studies primarily focused on demographic data, laboratory-based data, and radiomic features of the tumor, with limited research on the morphological features of the breast (29–31). These studies were focused on the radiomic features of the tumor and mammographic physical parameters of the image. This has been the norm for the last few years. Prior to this, there were no studies in this field using only morphological features of the contralateral breast. Dimensions of the ipsilateral breasts, like total volume and fibroglandular volume, would vary because of tumors. Using threshold function on breast MRI images to delineate fibroglandular tissues would be influenced by the presence of tumors. Hence, contralateral breasts were chosen for this study. Bilateral breasts are considered physiologically almost identical, although slight asymmetries may be observable in most women. These asymmetries can manifest in size, shape, or volume and are typically within a normal range. Coltman et al. stated in their study that there was no significant difference in the mean breast volume between the women's left and right breasts (32).

The aim of this study was to leverage ML techniques to develop a predictive model capable of accurately distinguishing between IDC and ILC, based on morphological features of the contralateral breast.

2 Methods

This retrospective, quantitative study included 143 female patients selected from a database of 922. The patients who were 40 years or older, had undergone a breast MRI examination, and were diagnosed with IBC. The cohort consisted of 117 patients with IDC and 26 with ILC, classified based on histopathological reports available publicly in the “DUKE Breast-Cancer-MRI” database from TCIA website. The patients who had undergone surgical intervention in the contralateral breast (CLB), those with bilateral breast cancer, MRI image artefacts, significant breast density asymmetry, known benign mass lesions in the CLB, or those lacking axial T1-weighted pre-contrast MRI images were excluded. The flowchart of the patient selection protocol is illustrated as [Supplementary Figure S1](#) in supplementary materials.

The Ethics Review Committee of the Faculty of Allied Health Sciences, University of Peradeniya, Sri Lanka, approved the study protocol. All procedures were performed per relevant guidelines and regulations.

Breast MRI images were sourced from the openly accessible “Duke Breast-Cancer-MRI” collection on the Cancer Imaging Archive website (33). This study adhered to The Cancer Imaging Archive's data usage policy and restrictions. The DUKE database was collected at Duke Hospital, USA, between 2000 and 2014.

Images were acquired using 1.5T or 3T scanner from General Electric and Siemens. Digital Imaging and Communications in Medicine (DICOM) images of the axial T1-weighted pre-contrast sequence were included in the study. The workflow of the supervised learning method used in the development of the IDC and ILC classification model is illustrated in Figure 1.

2.1 Image processing and region of interest selection

3D Slicer software (version 5.2.2) was used as the image processing tool. The investigator manually selected the ROI within the total breast area of the CLB as illustrated in Figure 2. The anterior border was defined by the skin margin, and the posterior border was delineated at the level of the sternum. Pectoralis muscles, identified by their grey color in MRI images, were excluded from the ROI selection process, with validation provided by a board-certified radiologist. The ROI for the fibroglandular area of the CLB was marked in each slice using the

thresholding function of the 3D Slicer application and by manually delineating the fibroglandular area. A comprehensive 3D image of the CLB was generated by combining all the slices of the MRI image.

2.2 Quantification of morphological features

The 3D Slicer application's quantification feature was used to determine parameters for each patient's CLB. Specifically, the total volume of the CLB, fibroglandular tissue volume of the CLB, surface area of the CLB, and surface area of the fibroglandular tissue of the CLB were computed. MRI breast density, a crucial metric in breast health assessment, was calculated as the ratio of fibroglandular tissue volume to the overall breast volume. In addition, the ratio of fibroglandular tissue volume of the CLB, to the surface area of the fibroglandular tissue of the CLB, and the ratio of total volume of the CLB, to the surface area of the CLB were calculated.

2.3 Machine learning model development

After feature computation, supervised learning methods were applied to build a model for predicting types of IBC (IDC or ILC). The model was developed using Python 3.12 software. Given the imbalance between IDC and ILC cases, the Synthetic Minority Oversampling Technique (SMOTE) was applied to balance the dataset. The resampled dataset was stratified and split into training and testing subsets. Different proportions of training and testing data set combined with different feature selections were used to find the optimal splitting ratio (as shown in Table 1). A training and testing data set with the proportion of 80:20 was chosen, with all 7 features selected to build the model.

2.4 Feature selection and model training

The ANOVA F-test was used to select a subset of normalized features most pertinent to distinguishing between IDC and ILC. All 7 features were selected based on a tenfold cross-validation method which was employed to evaluate various algorithms, including Support Vector Classifier (SVC), Random Forest (RF), Gaussian Naive Bayes (GaussianNB), K-Nearest Neighbors (KNN), Decision Tree Classifier (DT), Logistic Regression (LR), and Linear Discriminant Analysis (LDA). The RF classifier, which provided the best performance, was used to develop the final model. The selected RF classifier is then validated on the testing set.

2.5 Model optimization and evaluation

GridSearchCV is a sklearn technique which uses grid search algorithm to find the optimal hyperparameters for the model (34). These hyperparameters include min_samples_split, n_estimators, max_depth, max_features, bootstrap, and min_samples_leaf. Each

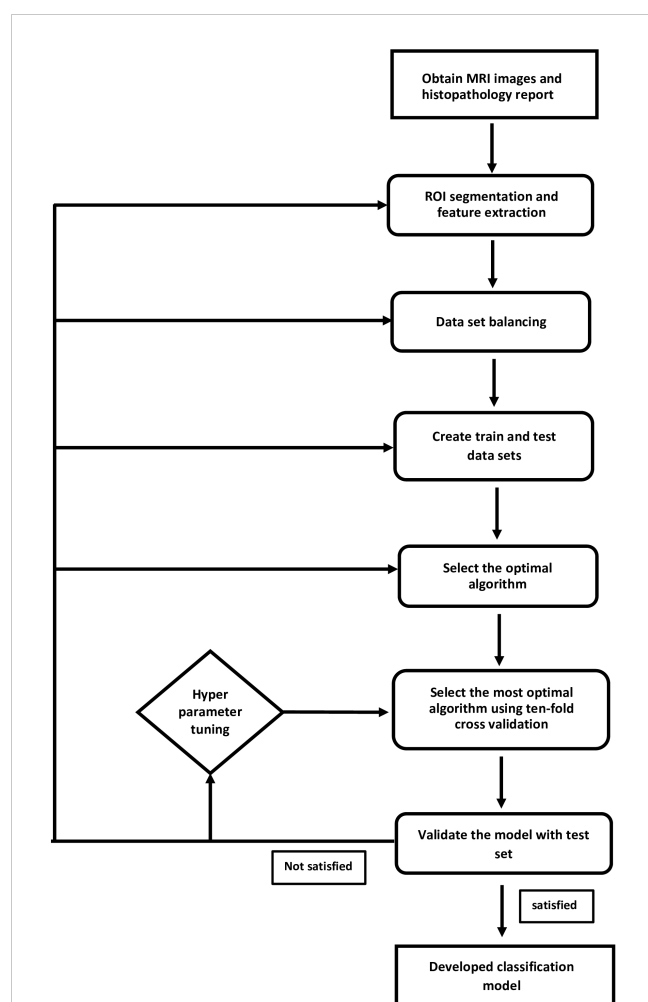


FIGURE 1
Supervised learning method to classify IDC and ILC was applied. This flowchart illustrates the steps followed to establish a prediction model.

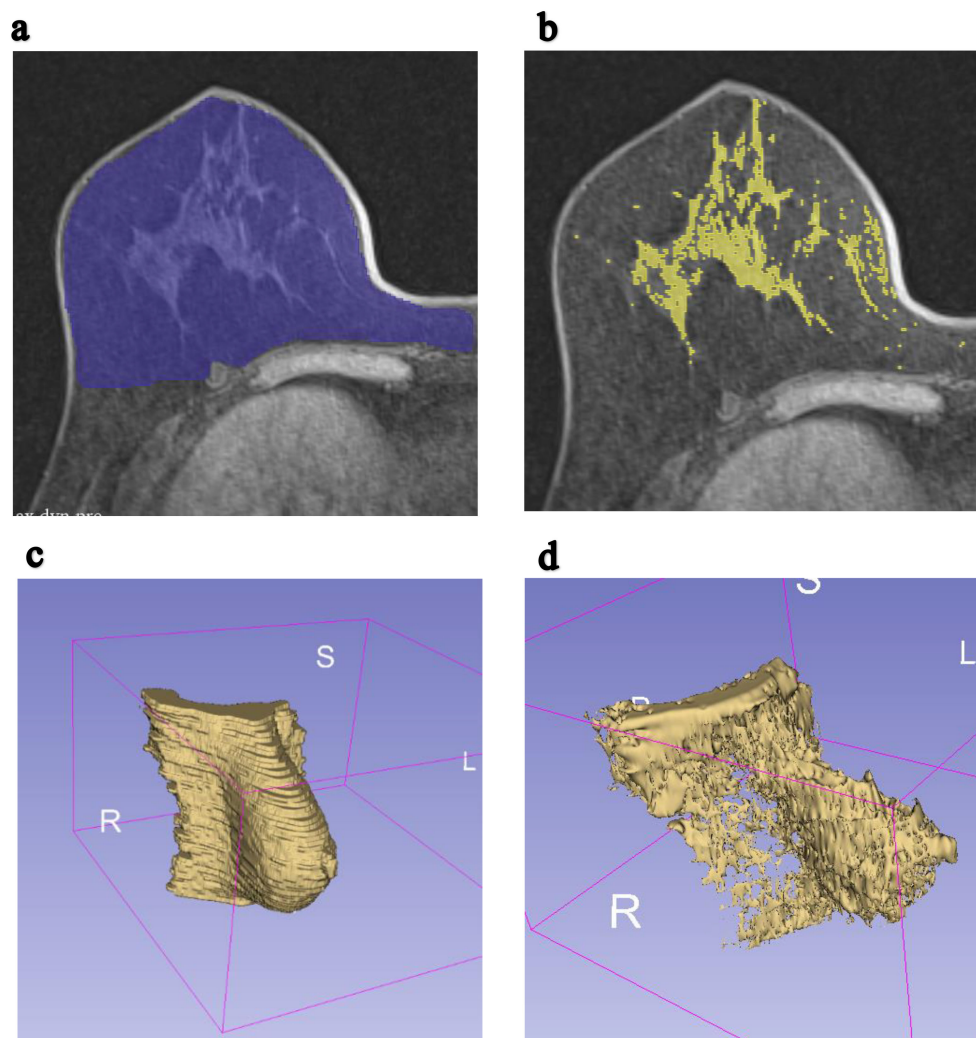


FIGURE 2

Example of MRI T1-weighted breast axial images were segmented to compute total breast volume and fibroglandular volume. **(a)** ROI for a single slice of the right breast in the axial plane. **(b)** ROI for a single slice of the fibroglandular tissue of the right breast in the axial plane. **(c)** The constructed 3D image of total volume of the contralateral breast. **(d)** The constructed 3D image of fibroglandular volume of the contralateral breast.

hyperparameter was tested within predefined ranges; min_samples_split: from 2 to 5, n_estimators: from 100 to 200 (with the steps of 10), max_depth: 10 to 100 (with the steps of 10), max_features: 3, 5, 10, boot_strap: True, False, min_samples_leaf: 1, 2, 4. Precision and recall scores of the model were acquired.

The model's performance was evaluated using several metrics, including accuracy, precision, recall, and F1 score as stated here by the Equations 1–4.

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (1)$$

Where TP, TN, FP and FN specify true positive, true negative, false positive and false negative, respectively. Here, Accuracy represents the proportion of correct predictions out of the total number of predictions made by the ML model.

$$Precision = \frac{TP}{(TP + FP)} \quad (2)$$

Where TP and FP represent true positive and false positive. Precision represents the proportion of the correctly made p positive predictions made by the ML model.

$$Recall = \frac{TP}{(TP + FN)} \quad (3)$$

Where TP and FN indicate true positive and false negative. Recall indicate the correctly predicted positive outcomes out of all the positive individuals.

$$F1 = 2 \cdot \left[\frac{(Precision \cdot Recall)}{(Precision + Recall)} \right] \quad (4)$$

F1 score indicates the balance of precision and recall.

TABLE 1 Results from tenfold cross validation and test set validation based on various combination of train and test data splits and feature selection.

Testing set (%)	Features	RFC ten-fold cross validation score	Accuracy
15	4	77.13	72.22
	5	78.26	58.33
	6	81.78	66.66
	7	81.76	69.44
20	4	74.325	68.08
	5	75.46	70.21
	6	78.07	74.46
	7	79.12	70.21
25	4	72.38	69.49
	5	74.24	69.49
	6	75.94	72.81
	7	75.42	76.27
30	4	71.98	69.01
	5	71.36	74.64
	6	73.78	69.01
	7	70.73	69.01
35	4	69.16	63.41
	5	71.11	75.6
	6	71.83	69.51
	7	69.16	75.6
40	4	69.28	74.46
	5	71.42	76.59
	6	69.28	79.78
	7	70.71	79.78

The precision-recall curve was used to illustrate the trade-off between sensitivity and specificity. The final decision threshold was set at 0.5 to maximize the model’s performance. The area under the ROC curve (AUC-ROC) was also evaluated to further assess the model’s efficacy.

3 Results

3.1 Feature selection and ANOVA F-Test Scores

The ANOVA F-test was used to identify the most significant features for distinguishing between IDC and ILC. As shown in [Table 2](#), the feature “Total volume of the contralateral breast” reported the highest ANOVA F-test score (mean value = 8.343721), indicating a strong correlation with the IDC and ILC

TABLE 2 ANOVA F-test scores of 7 morphological features used in the study for feature selection.

Feature	ANOVA F-test score
Total volume of the contralateral breast	8.343721
Fibroglandular tissue volume of the contralateral breast	0.001393
Breast density	6.382038
Surface area of the contralateral breast	6.921108
Surface area of fibroglandular tissue of the contralateral breast	0.817569
Fibroglandular volume of the contralateral breast/Surface area of fibroglandular tissue	1.735226
Total volume of the contralateral breast/Surface area of the contralateral breast	5.226712

differentiation. In contrast, “Fibroglandular volume of the contralateral breast” (mean value = 0.001393) reported the lowest score, suggesting minimal correlation with the cancer types. Another notable feature is “Surface Area of the contralateral breast” with a score of 6.921108.

3.2 Algorithm performance

Several ML algorithms were evaluated to determine the most effective model for predicting IDC and ILC. [Table 3](#) summarizes the accuracy and standard deviation of each algorithm based on ten-fold cross-validation.

The Random Forest Classifier demonstrated the highest mean accuracy of 79.13%, outperforming other algorithms such as Decision Tree Classifier(71.19%), Support Vector Classifier (71.05%), and Gaussian Bayes (63.65%).This high performance, combined with a relatively low standard deviation, indicated the Random Forest Classifier’s robustness and reliability in distinguishing between IDC and ILC.

TABLE 3 Performance summary of different algorithms under cross validation (mean accuracy and standard deviation).

Algorithm	Mean Accuracy (%)	Standard Deviation
Logistic Regression	56.6	0.128070
Linear Discriminant Analysis	57.16	0.121735
k-Nearest Neighbors Classifier	65.17	0.112232
Decision Tree Classifier	71.19	0.116184
Gaussian Bayes	63.65	0.131818
Support Vector Classifier	71.05	0.143569
Random Forest Classifier	79.13	0.096389

3.3 Model evaluation and performance metrics

The Random Forest Classifier, identified as the best-performing model, was further evaluated using the test set. The evaluation metrics included precision, recall, and F1-score for both IDC and ILC. Table 4 presents the detailed performance metrics.

The Random Forest model achieved an overall accuracy of 70% with IDC and ILC classifications. ILC exhibited a higher precision score (75%) while IDC exhibited a higher Recall score (78%). The F1-score, which balances both precision and recall, is higher for IDC (72%) compared to ILC (68%).

3.4 Hyperparameter tuning and decision threshold adjustment

To further enhance the model's performance, hyperparameter tuning and decision threshold adjustments were conducted. Precision and Recall scores were maximized by running GridSearchCV on the model to find the optimal hyperparameters within a pre-defined range. Table 5 describes the details of optimized hyperparameters for maximizing precision and recall, respectively.

After tuning, the Random Forest model with maximum precision correctly predicted 19 cases of IDC and 17 cases of ILC, with 4 false negatives and 7 false positives. The performance metrics for this optimized model are shown in Table 6.

For maximum recall, the model accurately predicted 19 cases of IDC and 18 cases of ILC, with 4 false negatives and 6 false positives. The performance metrics are detailed in Table 7.

Further adjustments were made to the decision threshold based on the precision-recall curve (illustrated in Figure 3), which optimized the balance between precision and recall. The decision threshold was adjusted to 0.5 to achieve optimal values. The final model achieved an accuracy of 79%.

The ROC curve, shown in Figure 4, plots the true positive rate against the false positive rate across different decision thresholds, providing insight into the trade-offs between precision and recall. An AUC value of 0.851 was obtained which indicates the model's ability to distinguish positive and negative outcomes.

The confusion matrix in Figure 5 visually represents the performance of the optimized classification model. The matrix

TABLE 5 Optimized hyperparameters for maximum precision and maximum recall.

Hyperparameter	Maximized precision	Maximized recall
Max Depth	20	30
Max Features	3	7
Min samples leaf	1	1
Min Samples Split	2	2
Number of Estimators	133	177

highlights the 19 true negatives, 18 true positives, 6 false positives, and 4 false negatives. In this aspect, true negatives and true positives are identified as correctly predicted IDCs and correctly predicted ILCs, respectively. False positives are identified as actual ILCs, falsely predicted to be IDCs, and false negatives are identified as actual IDCs, falsely predicted to be ILCs. The confusion matrix corroborates the model's high accuracy and balanced performance in distinguishing between IDC and ILC.

4 Discussion

Past studies incorporating the RF Classifier in building ML models for breast cancer prediction reported various accuracies and sensitivities (32, 35). However, they focused entirely on the affected breast, while the focus was on the contralateral breast in this study. Among the morphological features extracted from the contralateral breast, total volume of the breast, surface area of the breast, and breast density had higher ANOVA F-test scores, indicating higher predictive power when differentiating IDC and ILC. A study by Egan et al. observed a significant positive association between breast size and breast cancer (36). Wanders et al., in their study, found a correlation between breast density, and both interval and screening breast cancers (37). A study by Nara et al. found a significant correlation between fibroglandular volume and breast cancer (38). Past studies related to classifying IDC and ILC primarily focused on the tumor's radiomic features, and they did not consider the morphological features of the breast. Gunawan et al. acquired an accuracy of 76% based on mammographic physical parameters (29). A study by Faraz et al. focused on radiomic features to achieve an AUC of 0.73 (31). We have selected morphological features of the contralateral breast in contrast to their studies and achieved a

TABLE 4 Performance of the random forest classifier on the test set.

Cancer Type	Precision (%)	Recall (%)	F1-Score (%)	Support
Ductal	67	78	72	23
Lobular	75	62	68	24
Accuracy			70	47
Macro Average	71	70	70	47
Weighted Average	71	70	70	47

TABLE 6 A summary of performance metrics for the model optimized for precision.

Cancer Type	Precision (%)	Recall (%)	F1-Score (%)	Support
Ductal	73	83	78	23
Lobular	81	71	76	24
Accuracy			77	47
Macro Average	77	77	77	47
Weighted Average	77	77	77	47

Precision, Recall, F1-score and support score were obtained.

modest accuracy of 79% and an AUC of 0.851. Compared to their models, our machine learning model performed slightly better in differentiating IDC and ILC. A study by Maiti et al. achieved an AUC of 0.998 for building a model based on the radiomic features of the tumor (30). Compared to other studies, they achieved an impressive AUC but didn't use machine learning or deep learning methods because of the smaller sample size. It can be reasoned that using ML model would not be ideal for smaller datasets like our study. In 2023, Bouasria et al. explored the optimal sample size needed for a RF model (39). They mentioned that a 300–500 sample size was recommended for optimal performance, but the accuracy difference above 50 sample size was shown to be minimal, not accounting for deeper analysis. Hence, the ML model could be advantageous since it could handle high dimensional datasets and precision-recall tradeoffs. Contralateral breast cancer refers to development of cancer in the opposite breast after a patient has already been diagnosed with cancer in one breast. Giannakeas et al., in their study observed that 3.2% of patients were diagnosed with the CLB cancer, and the 25-year actuarial risk of contralateral invasive breast cancer was 9.9% (40). There are several risk factors for the CBC cancer- age, genetic mutations, tumor type and hormone receptors. Several studies have discussed correlation between histological type of primary breast cancer(IDC and ILC), and the development of contralateral breast cancer (41, 42). Langland et al. observed that there were no risks in developing contralateral breast cancer based on primary cancer histology(IDC and ILC) (42). However, a study by Park et al., based on a retrospective analysis of 1071 patients found that the risk of developing contralateral breast cancer was higher in patients with ILC(7.1%) compared to those with IDC(1.5%). This finding aligns with an earlier study by Arpino et al. in 2004 which observed that

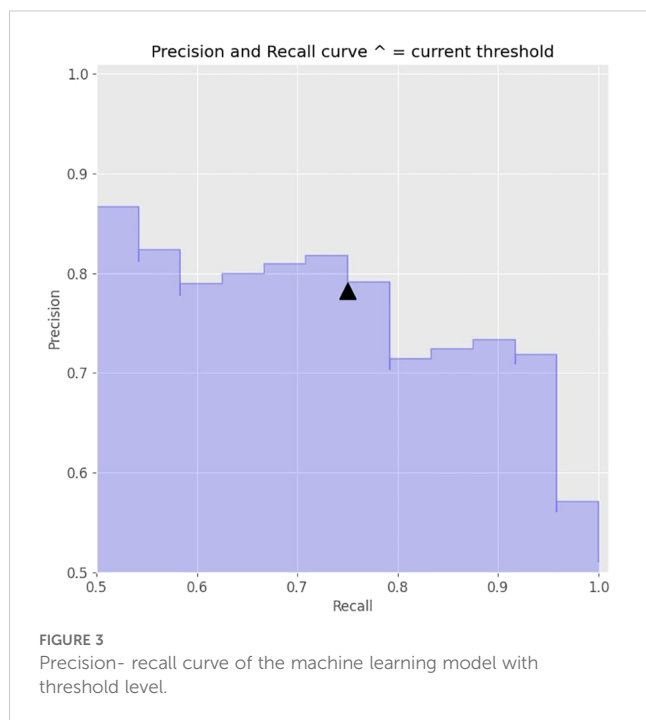
the incidence of CLB cancer was higher in ILC patients than in IDC patients (20.9% vs 11.9%) (11). Therefore, our study could be extended in the future to explore relationship between the morphological features of the CLB, histological types of primary cancer(IDC and ILC), and the risk of CLB cancer.

We had an unequal number of IDC and ILC patients, with 117 and 26, respectively. The SMOTE oversampling method was used to balance the data set. The sample size of ILC was equalized to that of IDC, since IDC had the largest sample size in the population. Without balancing the dataset, ANOVA feature selection scores of 0.440811, 0.000123, 1.012839, 0.280833, 0.009705 and 0.198752 were obtained for total volume of the contralateral breast, fibroglandular volume of the contralateral breast, breast density, surface area of the contralateral breast, surface area of fibroglandular tissue, fibroglandular volume of the contralateral breast/surface area of fibroglandular tissue and total volume of the contralateral breast/surface area of the contralateral breast, respectively. The scores were very low across all the features and showed no distinct variations compared to the ANOVA F-test scores we obtained (see Table 2) after using the SMOTE oversampling method. Using SMOTE positively increased the performance of the ML model. When splitting train and test data, we had to go through all the combinations, as pointed out in Table 1, to find the ideal combination to build a robust model. Even though split ratios of 65:35 and 60:40 gave higher accuracies, the accuracies decreased significantly after hyperparameter tuning. Therefore, an 80:20 ratio was selected to reduce the variance and overfitting of the data. Overfitting is when the model performs exceptionally well on the training data set and poorly on the testing data set. Overfitting happens when the model tries to memorize the data instead of generalizing it in the training phase. Optimal

TABLE 7 A summary of performance metrics for the model optimized for recall.

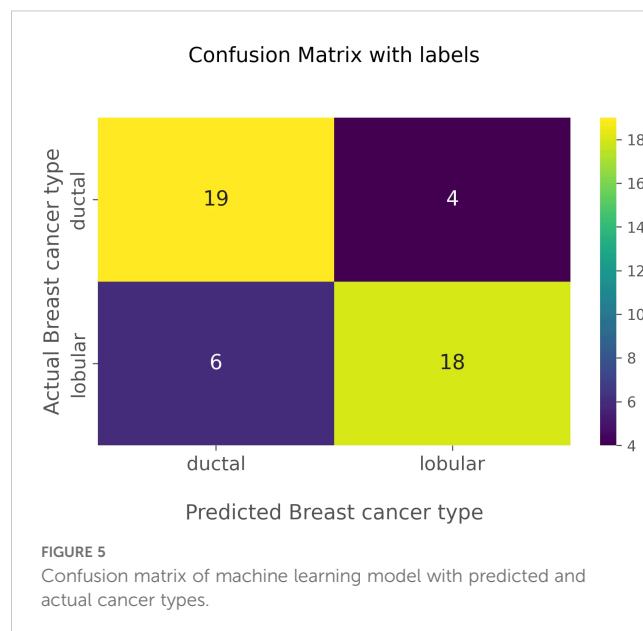
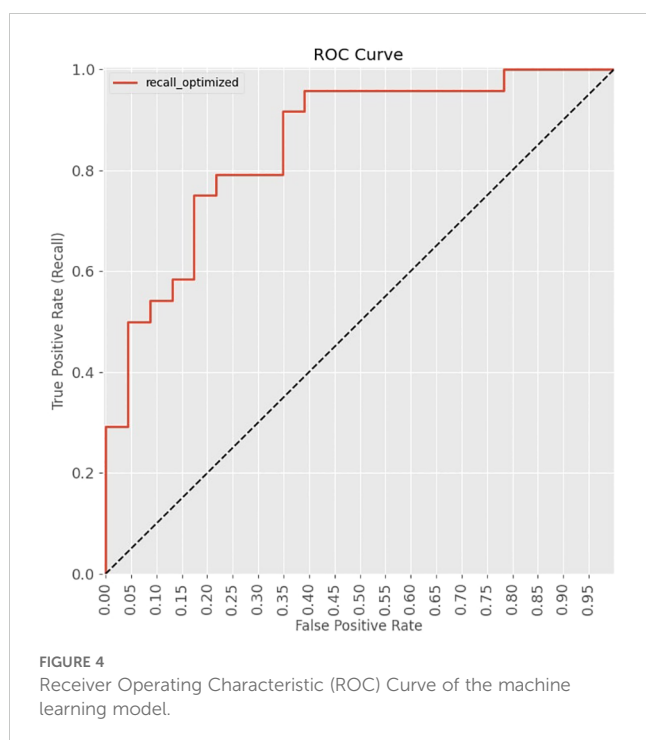
Cancer Type	Precision (%)	Recall (%)	F1-Score (%)	Support
Ductal	76	83	79	23
Lobular	82	75	78	24
Accuracy			79	47
Macro Average	79	79	79	47
Weighted Average	79	79	79	47

Precision, Recall, F1-score and support score were obtained.



hyperparameters were selected for the maximized precision and recall scores. Given the need to address IDC and ILC, it is essential to balance false positives and false negatives to optimize the performance of the machine learning model. Hence, precision and recall were maximized, and the threshold was set to 0.5 to acquire balanced precision and recall.

Based on this study, IDC and ILC can be differentiated without invasive procedures, which would immensely benefit medicine. This



could complement biopsy procedures moving forward. In situations where performing biopsies are considered challenging, this model could be of service in such circumstances. It should be evaluated on a large data set in the future to be considered reliable. Data acquisition can also be accomplished quickly since we only focus on the fat-suppressed T1 weighted imaging. This can be achieved within minutes, and patient comfort is not compromised. The acquisition takes approximately 1 minute for standard resolution and 6 minutes for higher resolution at 3T (43).

However, there are several limitations in this study. The difference in the dataset sizes of IDC and ILC could have affected the model's performance. Gunawan et al. also faced the same struggle with an unbalanced dataset (156 IDC and 8 ILC), which resulted in almost identical accuracy to our study (29). Decreasing the number of IDC cases to balance the dataset would also affect the model, as the ability to learn complex patterns would be compromised. As stated by Bouasria et al., decreasing the overall sample size to around 50 would marginally decrease the model's accuracy (39). Hence, we proceeded with 117 IDCs and 26 ILCs as our dataset. Even though data acquisition is quicker with fat-suppressed T1 weighted imaging, the time needed to extract the morphological features of the breast would still be longer when using 3D Slicer. Multiple recent studies have focused on building automatic segmentation models based on deep learning methods like deep convolutional neural networks (44–46). These automatic segmentation methods could be incorporated into our study to speed up the segmentation process further. Further advancements should be made to expedite the acquisition of morphological features of the breast. Another limitation of this machine learning model is the lack of an external data set comprising IDC and ILC patients for validation. External and multi-center validation is needed to gauge the accuracy and reliability of the model across various clinical settings. It would also prevent the overfitting of the model. Therefore, future studies should focus on incorporating external validation into this model.

5 Conclusion

This study developed and optimized a predictive model using morphological features of the contralateral breast and ML methods to differentiate IDC and ILC. The RF Classifier was identified as the most effective algorithm, achieving a high level of accuracy through extensive feature selection and hyperparameter tuning.

The significant predictive power of features such as the total volume of the CLB, surface area of the CLB and breast density underscores the importance of comprehensive feature selection in developing robust ML models. Classification of IDC and ILC through non-invasive method like this could be further developed in the future to complement biopsies in challenging situations.

Future directions should include expanding the dataset size and incorporating additional features to enhance model accuracy. By improving diagnostic accuracy, the developed model represents a valuable advancement in breast cancer management, ultimately contributing to better patient outcomes and more efficient healthcare delivery.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Ethics Review Committee, Faculty of Allied Health Sciences, University of Peradeniya, Peradeniya, Sri Lanka. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and institutional requirements.

Author contributions

NP: Conceptualization, Investigation, Methodology, Validation, Writing – original draft, Writing – review & editing. WN: Conceptualization, Investigation, Methodology, Writing – review & editing. HP: Conceptualization, Investigation, Methodology, Writing – review & editing. SV: Formal analysis, Methodology, Software, Writing – review & editing. PH: Supervision, Writing – review & editing. LS: Supervision, Writing – review & editing. MJ: Conceptualization, Project administration, Software, Supervision, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Acknowledgments

I would like to acknowledge the support provided by University of Peradeniya and express my gratitude to all the staff members of Department of Radiography/Radiotherapy for their guidance, resources and knowledge which were needed for the successful completion of this study.

Conflict of interest

The authors declare that the research was conducted without any commercial or financial relationships that could be interpreted as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2025.1588787/full#supplementary-material>

SUPPLEMENTARY FIGURE 1

This flow chart illustrates the patient selection protocol from "Duke Breast-Images-MRI" data base in TCIA website.

References

- Petracci E, Decarli A, Schairer C, Pfeiffer RM, Pee D, Masala G, et al. Risk factor modification and projections of absolute breast cancer risk. *J Natl Cancer Inst.* (2011) 103:1037–48. doi: 10.1093/jnci/djr172
- Lei S, Zheng R, Zhang S, Wang S, Chen R, Sun K, et al. Global patterns of breast cancer incidence and mortality: A population-based cancer registry data analysis from 2000 to 2020. *Cancer Commun.* (2021) 41:1183. doi: 10.1002/cac2.12207
- Harford JB. Breast-cancer early detection in low-income and middle-income countries: do what you can versus one size fits all. *Lancet Oncol.* (2011) 12:306–12. doi: 10.1016/S1470-2045(10)70273-4
- Weigelt B, Geyer FC, Reis-Filho JS. Histological types of breast cancer: How special are they? *Mol Oncol.* (2010) 4:192–208. doi: 10.1016/j.molonc.2010.04.004
- Malhotra GK, Zhao X, Band H, Band V. Histological, molecular and functional subtypes of breast cancers. *Cancer Biol Ther.* (2010) 10:955. doi: 10.4161/cbt.10.10.13879
- Barroso-Sousa R, Metzger-Filho O. Differences between invasive lobular and invasive ductal carcinoma of the breast: results and therapeutic implications. *Ther Adv Med Oncol.* (2016) 8:261–6. doi: 10.1177/1758834016644156
- Li CI, Anderson BO, Daling JR, Moe RE. Trends in incidence rates of invasive lobular and ductal breast carcinoma. *JAMA.* (2003) 289:1421–4. doi: 10.1001/jama.289.11.1421
- Manning P, Fazeli S, Lim V, Ladd WA, Eghtedari M, Chong A, et al. Invasive lobular carcinoma: A multimodality imaging primer. *Radiographics.* (2022) 42:E115–6. doi: 10.1148/rg.210058
- Cornford EJ, Wilson ARM, Athanassiou E, Galea M, Ellis IO, Elston CW, et al. Mammographic features of invasive lobular and invasive ductal carcinoma of the breast: a comparative analysis. *Br J Radiology.* (1995) 68:450–3. doi: 10.1259/0007-1285-68-809-450
- Li CI, Uribe DJ, Daling JR. Clinical characteristics of different histologic types of breast cancer. *Br J Cancer.* (2005) 93:1046–52. doi: 10.1038/sj.bjc.6602787
- Arpino G, Bardou VJ, Clark GM, Elledge RM. Infiltrating lobular carcinoma of the breast: Tumor characteristics and clinical outcome. *Breast Cancer Res.* (2004) 6:1–8. doi: 10.1186/bcr767
- Iorfida M, Maiorano E, Orvieto E, Maisonneuve P, Bottiglieri L, Rotmensz N, et al. Invasive lobular breast cancer: Subtypes and outcome. *Breast Cancer Res Treat.* (2012) 133:713–23. doi: 10.1007/s10549-012-2002-z
- Kurtz JM, Jacquemier J, Thorhorst J, Spitalier J -M, Amalric R, Hünig R, et al. Conservation therapy for breast cancers other than infiltrating ductal carcinoma. *Cancer.* (1989) 63:1630–5. doi: 10.1002/1097-0142(19890415)63:8<1630::AID-CNCR2820630833>3.0.CO;2-U
- Mathieu MC, Rouzier R, Llombart-Cussac A, Sideris L, Koscielny S, Travagli JP, et al. The poor responsiveness of infiltrating lobular breast carcinomas to neoadjuvant chemotherapy can be explained by their biological profile. *Eur J Cancer.* (2004) 40:342–51. doi: 10.1016/j.ejca.2003.08.015
- Cristofanilli M, Gonzalez-Angulo A, Sneige N, Kau SW, Broglio K, Theriault RL, et al. Invasive lobular carcinoma classic type: response to primary chemotherapy and survival outcomes. *J Clin Oncology.* (2005) 23:41–8. doi: 10.1200/JCO.2005.03.11
- Dorrius MD, Jansen-Van Der Weide MC, Van Ooijen PMA, Pijnappel RM, Oudkerk M. Computer-aided detection in breast MRI: a systematic review and meta-analysis. *Eur Radiol.* (2011) 21:1600. doi: 10.1007/s00330-011-2091-9
- Gilbert FJ, Hickman SE, Baxter GC, Allajbeu I, James J, Caraco C, et al. Opportunities in cancer imaging: risk-adapted breast imaging in screening. *Clin Radiol.* (2021) 76:763–73. doi: 10.1016/j.crad.2021.02.013
- McCormack VA, Dos Santos Silva I. Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. *Cancer Epidemiol Biomarkers Prev.* (2006) 15:1159–69. doi: 10.1158/1055-9965.EPI-06-0034
- Li X, Zhou C, Wu Y, Chen X. Relationship between formulaic breast volume and risk of breast cancer based on linear measurements. *BMC Cancer.* (2020) 20:1–8. doi: 10.1186/s12885-020-07499-5
- Scutt D, Manning JT, Whitehouse GH, Leinster SJ, Massey CP. The relationship between breast asymmetry, breast size and the occurrence of breast cancer. *Br J Radiology.* (1997) 70:1017–21. doi: 10.1259/bjr.70.838.9404205
- Chen HL, Zhou JQ, Chen Q, Deng YC. Comparison of the sensitivity of mammography, ultrasound, magnetic resonance imaging and combinations of these imaging modalities for the detection of small (≤ 2 cm) breast cancer. *Med (United States).* (2021) 100:E26531. doi: 10.1097/MD.00000000000026531
- Bluemke DA, Gatsonis CA, Chen MH, DeAngelis GA, DeBruhl N, Harms S, et al. Magnetic resonance imaging of the breast prior to biopsy. *JAMA.* (2004) 292:2735–42. doi: 10.1001/jama.292.22.2735
- Nguyen C, Wang Y, Nguyen HN, Nguyen C, Wang Y, Nguyen HN. Random forest classifier combined with feature selection for breast cancer diagnosis and prognostic. *J BioMed Sci Eng.* (2013) 6:551–60. doi: 10.4236/jbise.2013.65070
- Anisha PR, Reddy CKK, Apoorna K, Mangipudi CM. Early diagnosis of breast cancer prediction using random forest classifier. *IOP Conf Ser Mater Sci Eng.* (2021) 1116:12187. doi: 10.1088/1757-899X/1116/1/012187
- Mu T, Nandi AK. Breast cancer detection from FNA using SVM with different parameter tuning systems and SOM-RBF classifier. *J Franklin Inst.* (2007) 344:285–311. doi: 10.1016/j.jfranklin.2006.09.005
- Ghazavi SN, Liao TW. Medical data mining by fuzzy modeling with selected features. *Artif Intell Med.* (2008) 43:195–206. doi: 10.1016/j.artmed.2008.04.004
- Breiman L. Random forests. *Mach Learn.* (2001) 45:5–32. doi: 10.1023/A:1010933404324
- Wainer J. Comparison of 14 different families of classification algorithms on 115 binary datasets. (2016). doi: 10.48550/arXiv.1606.00930
- Gunawan AAN, Suardana P, Sulaiman A, Negara AANFK, Mahendra AANS, Negara AANFC. Classification of invasive lobular carcinoma (ILC) and invasive ductal carcinoma (IDC) using the support vector machine (SVM) method. *Appl Math Sci.* (2022) 16:261–71. doi: 10.12988/ams.2022.916783
- Maiti S, Nayak S, Hebbar KD, Pendem S. Differentiation of invasive ductal and lobular carcinoma of the breast using MRI radiomic features: a pilot study. *F1000Research.* (2024) 13:91. doi: 10.12688/f1000research.146052.2
- Faraz K, Dauce G, Bouhamama A, Leporq B, Sasaki H, Bito Y, et al. Characterization of breast tumors from MR images using radiomics and machine learning approaches. *J Personalized Med.* (2023) 13:1062. doi: 10.3390/jpm13071062
- Rabiei R, Ayyoubzadeh SM, Sohrabei S, Esmaili M, Atashi A. Prediction of breast cancer using machine learning approaches. *J BioMed Phys Eng.* (2022) 12:297. doi: 10.31661/jbpe.v0i0.2109-1403
- Saha A, Harowicz MR, Grimm LJ, Kim CE, Ghate SV, Walsh R, et al. A machine learning approach to radiogenomics of breast cancer: a study of 922 subjects and 529 DCE-MRI features. *Br J Cancer.* (2018) 119:508. doi: 10.1038/s41416-018-0185-8
- Zöller MA, Huber MF. Benchmark and survey of automated machine learning frameworks. *J Artif Intell Res.* (2021) 70:409–72. doi: 10.48550/arXiv.1904.12054
- Aroef C, Rivan Y, Rustam Z. Comparing random forest and support vector machines for breast cancer classification. *TELKOMNIKA (Telecommunication Computing Electron Control).* (2020) 18:815–21. doi: 10.12928/telkomnika.v18i2.14785
- Egan KM, Newcomb PA, Titus-Ernstoff L, Trentham-Dietz A, Baron JA, Willett WC, et al. The relation of breast size to breast cancer risk in postmenopausal women (United States). *Cancer Causes Control.* (1999) 10:115–8. doi: 10.1023/A:1008801131831
- Wanders JOP, Holland K, Karssemeijer N, Peeters PHM, Veldhuis WB, Mann RM, et al. The effect of volumetric breast density on the risk of screen-detected and interval breast cancers: A cohort study. *Breast Cancer Res.* (2017) 19:1–13. doi: 10.1186/s13058-017-0859-9
- Nara M, Fujioka T, Mori M, Aruga T, Tateishi U. Prediction of breast cancer risk by automated volumetric breast density measurement. *Jpn J Radiol.* (2023) 41:54–62. doi: 10.1007/s11604-022-01320-y
- Bouasria A, Bouslihim Y, Gupta S, Taghizadeh-Mehrjardi R, Hengl T. Predictive performance of machine learning model with varying sampling designs, sample sizes, and spatial extents. *Ecol Inform.* (2023) 78:102294. doi: 10.1016/j.ecoinf.2023.102294
- Giannakeas V, Lim DW, Narod SA. The risk of contralateral breast cancer: a SEER-based analysis. *Br J Cancer.* (2021) 125:601–10. doi: 10.1038/s41416-021-01417-7
- Dawson LA, Chow E, Goss PE. Evolving perspectives in contralateral breast cancer. *Eur J Cancer.* (1998) 34:2000–9. doi: 10.1016/S0959-8049(98)00208-1
- Langlands F, White J, Kearins O, Cheung S, Burns R, Horgan K, et al. Contralateral breast cancer: Incidence according to ductal or lobular phenotype of the primary. *Clin Radiol.* (2016) 71:159–63. doi: 10.1016/j.crad.2015.10.030
- Brown R, Storey P, Geppert C, McGorty K, Leite APK, Babb J, et al. Breast MRI at 7 Tesla with a bilateral coil and T1-weighted acquisition with robust fat suppression: Image evaluation and comparison with 3 Tesla. *Eur Radiol.* (2013) 23:2969–78. doi: 10.1007/s00330-013-2972-1
- Moeskops P, Wolterink JM, van der Velden BHM, Gilhuijs KGA, Leiner T, Viergever MA, et al. (2017). Deep learning for multi-task medical image segmentation in multiple modalities, in: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Germany: Springer Verlag. pp. 478–86, 9901 LNCS. doi: 10.48550/arXiv.1704.03379
- Piantadosi G, Sansone M, Fusco R, Sansone C. Multi-planar 3D breast segmentation in MRI via deep convolutional neural networks. *Artif Intell Med.* (2020) 103:101781. doi: 10.1016/j.artmed.2019.101781
- Guo YY, Huang YH, Wang Y, Huang J, Lai QQ, Li YZ. Breast MRI tumor automatic segmentation and triple-negative breast cancer discrimination algorithm based on deep learning. *Comput Math Methods Med.* (2022) 2022:2541358. doi: 10.1155/2022/2541358