Artificial Intelligence Review (2025) 58:240
https://doi.org/10.1007/s10462-025-11232-8

Recent trends on mammogram breast density analysis using deep learning models: neoteric review

S. Jeba Prasanna Idas[1] · K. Hemalatha[2] · Jayakumar Naveenkumar[1] · T. Joshva Devadas[1]

Accepted: 10 April 2025 / Published online: 10 May 2025

Abstract
Breast cancer is a globally prevalent and potentially fatal illness affecting women. Timely identification of screening mammography may decrease the occurrence of incorrect positive results and enhance the rate of patient survival. Nevertheless, the density of breast tissue in mammograms can impact the precision and effectiveness of detecting breast cancer. This paper examines the existing body of research on the analysis of breast density in mammograms utilising advanced deep learning models, including convolutional neural networks (CNN), transfer learning (TL), and ensemble learning (EL). Additionally, it examines various datasets and evaluation measures employed in the investigations. The study demonstrates that deep learning models can attain exceptional accuracy in categorising breast density. However, they encounter obstacles such as limited data availability, intricate model structures, and difficulties in interpreting the results. The research asserts that categorising breast density is an essential undertaking in order to enhance the identification and survival rates of breast cancer. Further investigation is warranted to examine the most effective deep learning structures, data augmentation methods, and interpretable models for this undertaking.

Keywords Mammogram · Breast density classification · Convolutional neural network · Transfer learning · Ensemble learning

1 Introduction

Breast cancer exhibits a high prevalence within the population of India, and its incidence is similar in cities and villages. Studies in the year of 2020, suggested that around 2.3 million women would be diagnosed with breast cancer and 6,85,000 were succumbing to the disease (Tice et al. 2015). Due to the higher death rate, steps have been initiated to reduce the

✉ T. Joshva Devadas
joshvadevadas.t@vit.ac.in

[1] School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, India

[2] School of Electronics Engineering, Vellore Institute of Technology, Vellore, India

death rate by detecting breast cancer at an early stage. The early detection of breast cancer is typically achieved through the use of mammography, which is considered a widely accepted and reliable imaging technique in the medical field (Oza et al. 2021). Screening mammograms and diagnostic mammograms are the two various techniques involved in mammography (Mathur and Taurin 2022). Screening mammograms can be used when the women have no signs or symptoms related to breast cancer (Murtaza Mendes and Matela 2021). A diagnostic mammogram is used when a lump or any other symptoms are found in the breast.

Women can gain awareness and understanding of breast cancer and breast density through risk factors associated with the disease. Both modifiable and non-modifiable risk variables have been identified (Kaiser et al. 2019). During a Patient's lifetime, people are unable to alter risk variables such as family history, genetic alterations, mensuration, and menopause. Lifestyle choices of women may lead to cancer but that would change or lower the risk of the development of breast cancer when a healthy lifestyle is practiced. This can be achieved by breastfeeding children, maintaining a healthy weight, getting enough sleep, exercising frequently, staying against alcohol, and having a proper diet.

The application of CAD in the field of medical image processing and deep learning has significantly simplified the detection of breast cancer. In previous studies, breast density was estimated and classified with the use of feature extraction and a variety of segmentation techniques. However, segmenting and classifying breast density remains challenging due to the low quality of images and the intervention of radiologists. Deep Learning negates the necessity for human involvement (Braithwaite et al. 2018). Since feature extraction and selection are integrated within the network architecture itself. Moreover, the development of conventional neural networks has proven to be particularly effective in classification tasks. Over the past years, a large number of researches has focussed on breast cancer, especially on breast cancer detection and classification using mammogram images. However, there is a limited body of research that specifically examines the classification of breast density and its potential influence on the progression of breast cancer. The primary focus of the survey is to examine and explore the interrelation between mammographic density and the risk factors associated with breast cancer. Additionally, the present study aims to conduct a comprehensive review of the application of CNN, TL, and EL, in breast density classification methods (Lester et al. 2022).

1.1 Sources of breast cancer

Breast cancer is a multifaceted ailment that arises from a confluence of hereditary predisposition and environmental influences. The precise aetiology of breast cancer remains elusive, while it is hypothesised to stem from alterations in the DNA of breast tissue cells. In normal cells, DNA serves as a blueprint for cellular growth and apoptosis, whereas in cancer cells, it imparts altered instructions that promote rapid cell proliferation. This can result in the development of a neoplasm, which has the ability to infiltrate and eradicate normal bodily tissue (Kressin et al. 2022). Cancer cells can eventually detach and disseminate to different regions of the body, leading to the development of metastatic cancer. The genetic alterations associated with breast cancer predominantly occur in the milk ducts, which serve as conduits for transporting milk to the nipple, and in the milk glands, responsible for producing breast milk. Although infrequent, other breast cells can undergo malignant transformation and become cancerous.

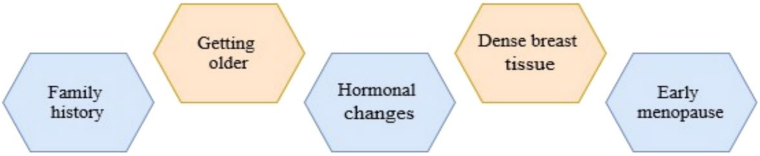


Fig. 1 Risk factors of breast cancer

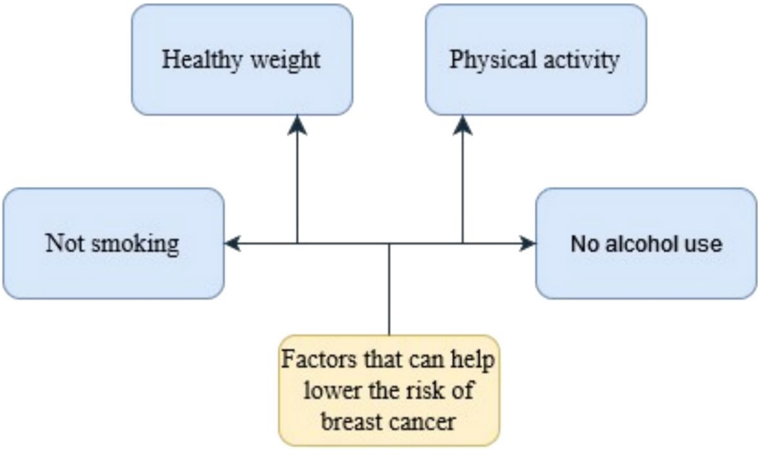


Fig. 2 Factors that can help lower the risk of breast cancer

Factors that increase the risk of breast cancer include having a family history of breast cancer, having a personal experience of breast cancer, and having a personal history of specific breast disorders. Females have a higher propensity for developing breast cancer compared to males (Miles et al. 2019). Additionally, the presence of dense breast tissue, characterised by a combination of fatty and dense tissue, can impede the detection of breast cancer during mammography. Consuming alcohol elevates the likelihood of developing breast cancer, while delaying the age at which one has their first child may also heighten the risk.

Genetic mutations that are passed down from parents, such as BRCA1 and BRCA2, can elevate the likelihood of developing breast cancer and other types of cancer. Combining oestrogen and progesterone in menopausal hormone therapy may potentially elevate the likelihood of developing breast cancer, obesity is associated with a greater chance of developing breast cancer, whereas exposure to radiation on the chest during childhood or early adulthood can also elevate the risk (Saffari et al. 2020).

To summarise, Fig. 1 depicts the risk of breast cancer can be influenced by factors such as familial history, personal history of breast cancer, and genetic predisposition. Engaging in a discussion with healthcare specialists is crucial to ascertain the optimal approach for managing and preventing this disease.

Figure 2 illustrates four lifestyle habits that have the potential to reduce the chance of developing breast cancer. These factors encompass the maintenance of a desirable body weight, participation in consistent physical exercise, cessation of smoking, and restriction of alcohol consumption. Women of all age groups should adhere to a well-balanced diet and engage in consistent physical exercise. Cessation of smoking is a pivotal determinant in mitigating the likelihood of developing breast cancer. Consuming a moderate amount of alcohol, which is defined as no more than one drink per day for women, is also advantageous. Additional factors, such as genetics and family history, can add to the level of risk. It is advisable to seek guidance from the doctor for tailored treatment programmes (Ulagamuthalvi et al. 2022).

The present paper is structured in the following manner. Section 2 provides a comprehensive review of several surveys that have been published in the field of deep models for mammogram breast density analysis. Section 3 shows how mammography breast density affects breast cancer detection and survival. Section 4 describes the architecture of deep models like CNN, pre-trained model and Ensemble model Sect. 5 illustrates the detailed review of published CNN-based solutions in the area of breast density classification. Section 6 examines the various available datasets, performance evaluation, and its usage level. Section 7 presents an overview of the research findings derived from the comprehensive survey. Lastly, Sect. 8 provides the concluding remarks and outlines potential avenues for further research.

1.2 Survey constraints

The scope and limitations of the data collection used in this survey are outlined here. In this study, we review and analyze literature proposed for breast cancer detection using deep learning models in the past two decades. There are several major digital repositories utilized by the survey to choose a wide range of deep learning models, including: Explore, Science Direct, CrossRef, MedPub, and Google Scholar. Also, the survey examines the literature that supports the development of interpretable models as well as methods for optimizing hyperparameters. Primarily the survey considers the studies that apply learning-based models using the dataset available to the public. Also, the focus of the study is chosen based on the mammogram breast density classification method along with the application of CNN, TL, EL, Contemporary models and Interpretable models. Additionally, segmentation, prediction, image retrieval, and preprocessing techniques are not included in this survey so more attention may be paid to the topic at hand.

2 Research motivation

Medical image classification has reached its peak with the advent of deep learning and it has become the focus of many recent researches. Breast cancer detection and classification methods have been the subject of extensive investigation during the past two decades. However, breast density has a strong association with breast cancer, and the performance of the classification model and the imaging modalities are affected due to the dense breast. Therefore, a deep learning model for breast density categorization will be the primary emphasis of this study. A single survey may not encompass all research publications in a particular domain (Daly et al. 2021). The primary aim of the survey is to lay a solid groundwork for future research in the selected topic and to investigate the existing body of knowledge in this area. This review also seeks to emphasize the strategies and techniques employed in implementation, as well as indicate gaps for prospective future research. Therefore, this evaluation presents in-depth responses to the following concerns.

1. Why mammogram breast density is a prominent factor in breast cancer detection?
2. Whether the association of mammogram density and survival rate is correlated?
3. Identify various methodologies used in breast density classification.
4. Address the significance of Deep Learning models in determining breast density.
5. Expose the many measures the academics have used to evaluate the efficacy of the Deep Learning categorization algorithm.
6. To what extent the model can be trained on the datasets available in the current deep learning architecture?

2.1 Flow of the survey

Figure 3 depicts the review process flow and design model. To begin with, this system model contains mammographic breast density to analyse the correlation between breast density and breast cancer. The subsequent step is to use the open-source Dataset for mammographic imaging to process the input images. The survey also includes a deep learning-based system for correctly categorizing breast density which includes CNN, TL, and EL. Finally, the study digs into the evaluation measures and provides a performance analysis of the numerous pre-trained architectures, including AlexNet, ResNet, MobileNet, and so on.

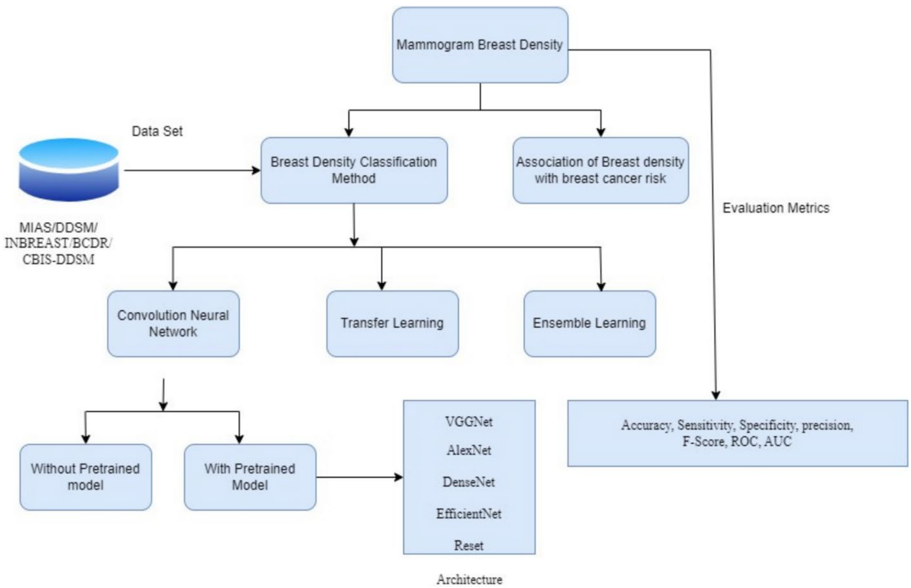


Fig. 3 Flow of survey

3 Related works

Few analyses have been conducted during the past two decades in the field of image processing with an emphasis on the deep model architecture and its performance. The diagnosis and staging of breast cancer were the primary topics of this survey. This review focuses on recent studies that have linked mammographic breast density to increased cancer detection and better patient outcomes. The works of literature published in the field of mammography breast density classification methods and various deep learning models are included in the survey to gain additional insights. The interrelation of breast density and breast cancer has been the subject of a literature review (Bond-Smith and Stone 2018) the review mainly focuses on methodological challenges that arose when using mammogram breast density. Moreover, the survey considers 165 samples from the published articles and is also explained using the Bayesian approach. The advantages of using continuous data for different patient populations are also described.

In the past two decades, Li has surveyed the articles and proposed a computer-aided diagnosis for breast density measurement method on mammograms. Several commonly used methods with difficulties, challenges, limitations, metrics, and disadvantages were also addressed (Xue et al. 2020). Abdelrahman has reviewed the computer-assisted detection using a convolution neural network for identifying breast cancer. Classifying breast density, detecting breast asymmetry, observing calcifications, and identifying and categorizing masses were the four main focuses of the study. The survey also provided a road map for providing a CNN-based solution to enhance mammographic diagnosis of breast cancer and discussed a CNN-CAD algorithm based on a Food and Drug Administration (FDA) approved model (Abdelrahman et al. 2021).

Rehman reviewed some machine-learning techniques for breast mammogram grading. The survey considers 110 papers for analysis to find out the techniques that are suitable for breast density detection and classification. In addition, the survey draws attention to a range of imaging modalities and factors that may prove helpful in determining mammographic grading (ur Rehman et al. 2022). Breast density on mammograms has been connected to a high rise in breast cancer risk, as described in a recent study by Allison. The survey aimed to highlight the interconnection of breast density on mammograms and the danger of developing breast cancer. The survey also contrasts the benefits and drawbacks of various mammogram density detection methods used in several studies (Allison et al. 2022).

Puliti evaluated the risk of breast cancer and volumetric mammogram density. Breast cancer incidence was determined by analyzing data from 16,752 women who had their first screening mammogram between the ages of 49 and 54. Breast density was found to be a significant risk factor for developing breast cancer. The study finds that breast density is strongly associated with breast cancer risk (Puliti et al. 2018). Kehm studied a cohort of mothers and daughters in Santigo, Chile. With a sample size of 42 mother-daughter couples. Clinical dual-energy X-ray absorptiometry (DXA) was used to calculate the percentage of fibro glandular volume, while optical spectroscopy (OS) was employed to measure collagen, water, and lipid concentration. The findings demonstrated that information on breast density and breast tissue is related yet separate from the obtained OS, DXA, and mammogram (Kehm et al. 2022).

The breast density analysis using AI-based CAD is compared and contrasted in a recent study. The author classifies 488 mammograms taken in a single institution on Asian women, which are categorized into BIRADS (Breast Imaging Reporting and Data System) density categories. Like the agreement seen between the Volpara tool and radiologist, the results demonstrate that AI-CAD density assessment exhibits fair agreement with those of radiologists (Lee et al. 2022). Rampun offers a complete analysis of the findings as well as a thorough evaluation of the relevant literature and methodology. It also describes breast density categorization by incorporating the Local Septenary Pattern (LSP). Also, it analyzes different methods of encoding local patterns in mammograms to categorize breast density. LSP beats the other methods on the MIAS and in the breast dataset with the highest accuracy values of 83.3% and 80.5% respectively (Rampun et al. 2020). Table 1 depicts the association of MBDWBC, MBDCM, CNN, TL and EL.

4 State of mammogram breast density and breast cancer

This section outlines the mammogram density followed by awareness and significance of mammograms with respect to breast cancer. Subsequently, this section goes on to detail the interconnection of mammogram breast density and the likelihood of rising breast cancer. Despite the fact, that a great number of studies concentrate on determining the location of breast cancer, segmentation, and classing, but interconnection of mammogram density and breast cancer as well as the categorization of breast density are the primary focus of this survey.

4.1 Breast density

Breast tissue is composed of Adipose, fibro glandular, and connective tissue (Cohn and Terry 2019). Women's breast density is not always proportional to breast size, perhaps breast density cannot be felt or touched. Moreover, dense breasts are common among women and it is not abnormal but it lowers the sensitivity of mammograms (Vargas-Hakim et al. 2021). Because of this effect experienced during screening mammography, finding unusual tissue in heavy breast mammograms is a difficult process. The masking effect occurs when mammogram tissue covers up malignant cells in the breast (Posso et al. 2019). Due to this masking effect abnormalities may likely blend with normal breast tissues. To measure the density radiologist assigns a certain level of breast density by computing the ratio of fibro glandular tissue (dense tissue) and connective tissue (non-dense or fatty tissue) (Wengert et al. 2018).

Researchers are naturally curious about whether or not there is a correlation between breast density and breast cancer. The Breast Imaging Reporting and Data System (BIRADS), created by the American College of Radiology is widely used in the scientific community for clinical classification of mammographic density (Kyanko et al. 2020). Dense breast is classed as extremely dense (B), scattered (A), heterogenous (c), or extremely dense (D) based on the BI-RADS scale. Researchers typically utilize BI-RADS fifth edition to categorize breast density, make predictions about breast density, and examine the correlation between mammogram density and breast cancer risk (Lin et al. 2023; Verma et al. 2021). Table 2 depicts the BIRADS scale of breast density.

Table 1 Existing review articles in the Association of Mammogram Breast Density with Breast Cancer (MBDWBC), Mammogram Breast Density Classification Method (MBDCM), CNN, TL, EL

Review article	Description	MBDCM	MBDWBC	CNN	TL	EL
Yu and Ye (2022)	Review on epidemiological factors associated with breast density		✓			
Bond-Smith and Stone (2018)	Discussed methodological challenges associated with breast density		✓			
Xue et al. (2020)	Analysis of breast density, focussing mainly on both qualitative and quantitative measurement approaches		✓			
Abdelrahman et al. (2021)	The Survey was conducted to assess the existing knowledge base on CNN in mammography	✓		✓		
ur Rehman et al. (2022)	Survey on various machine learning techniques employed in the assessment of image grading	✓	✓			
Puliti et al. (2018)	This research investigates the interconnection of volumetric breast density and the likelihood of breast cancer		✓			
Kehm et al. (2022)	Comparative investigation of breast density methods	✓	✓			
Lee et al. (2022)	A comparative study of AI-based breast density methods	✓	✓			
Rampun et al. (2020)	A survey on investigation of channel encoding techniques in breast density classification method	✓				
Shamshiri et al. (2023)	A Critical analysis of the biological factors implicated in breast density		✓			
Alison et al. (2022)	Investigate the correlation between breast cancer associated with tumour macrophage		✓			
Bodewes et al. (2022)	Study examines the interconnection of mammographic breast density and the risk of developing cancer		✓			
Mendes and Matela (2021)	A Review of mammogram-based breast cancer risk assessment		✓			
Nazari and Mukherjee (2018)	The administration of a survey investigates the relationship between breast cancer and breast density	✓				
This survey	This study underscores the relationship between mammogram density and breast cancer as well as the various classification methods on deep learning models	✓	✓	✓	✓	✓

Table 2 BI-RADS scale of breast density

BI-RADS class	Density range	Breast density class
A	00–25	Entirely fatty
B	26–50	Fibro glandular tissue
C	51–75	Heterogeneously dense
D	76–100	Extremely dense

4.2 Mammogram breast density

The widely used technique for identifying malignancy in women is mammography. Screening mammograms cannot detect all types of cancer in women because their sensitivity is affected by the dense cells present in the breast tissue. Radiologists use diagnostic mammography to determine breast density (MBD) by comparing the amount of fatty tissue (radiolucent) to the amount of epithelial and stromal components (radio-opaque in a breast). Epithelial and stromal elements which filter X-rays efficiently and absorb more energy, appear white or radio-opaque. By contrast, white fatty tissue appears black on the radiograph. Very little fatty tissue presents in dense breasts. Compared to less dense breasts with more fatty tissue, they have a higher risk of developing cancer. It also implies that both false positive and false negative occurrences in mammography interpretations are higher in dense breast tissue cases (Kumar et al. 2019; Yamada et al. 2022; Kim et al. 2022; Dayaratna and Jackson 2022; Moini et al. 2022; Pizzato et al. 2022).

4.3 Mammogram breast density correlation with breast cancer

Mammography has an overall sensitivity of 70 to 80% for a woman at reduced risk for breast cancer (Wanders et al. 2017). According to BIRADS, Women who fall under category A will have 80–90% of sensitivity (Kocer 2021). Whereas having the high density (30–48% were categorized into category D. Women who fall under category D may likely to develop breast cancer 4–6 times more than the one who falls under category A. So, women having higher breast density who fall under the category C and D should be aware of the risk of breast cancer (Schifferdecker et al. 2019).

The association between mammographic breast density and breast cancer was investigated in a cohort study of the Saudi population. The study by Aloufi incorporates both automatic and visual assessment of breast density and it uses approximately 1140 mammogram data. Finally, the study result shows that the mammogram breast density is highly associated with breast cancer. According to Kolb screening mammograms detected breast cancer in 11,130 women who showed no signs of disease. Women with extremely dense breasts have dropped to 48% of mammogram sensitivity. However, when compared to the normal breast the obtained mammogram sensitivity is 78%. The author concludes that women with exceptionally dense breasts are at a higher risk of developing breast cancer due to their decreased sensitivity. It has been suggested by Saftlas that the proportion of mammographic densities visible in the breast is not more reliable than a qualitative analysis of mammographic patterns for determining the risk of developing breast cancer (Aloufi et al. 2022; Zhang et al. 2022; Duffy et al. 2018).

Ali evaluated the distribution of breast density among women in Sulaimaniyah, Iraq (Ali et al. 2022). 750 women who underwent mammogram routine at the Sulaimaniyah Cancer Institute. The findings indicate that 54% of breast cancer cases of BIRADS Classes C and D. Moreover, Age, BMI, and Family were all correlated with breast density. I observed a strong association between mammogram density and sensitivity of mammography (Li et al. 2022). Hence it is observed that mammogram breast density was the major risk factor while detecting breast cancer. Puliti made a cohort study with 16,752 women under the age of 49–54 with two rounds of screening programs. The incidence of breast cancer formed in the dense breast is found in the second round of screening. Moreover, the author calculated breast density by using automated tools and found a strong correlation between increasing density and an increased risk of breast cancer.

Another Study was conducted with 329 breast patients over a period of 8 years. At different times it is observed that mammogram density for all the breast cancer patients falls under categories D to D. Only 19% of the tumors were correctly diagnosed and 81% of breast cancer was missed due to the overlap of dense tissue. In addition, it should be noted that the risk of breast cancer can also be influenced by variations in breast density. Advani has addressed some factors that influence mammogram breast density namely age, Body Mass Index (BMI). Hormonal radiation Therapy (HRT), family history, and menopause (Advani et al. 2021). The association of BMI is evaluated with the density of the mammogram. Further, the risk factor of breast cancer may also vary based on the association and the influence of breast density. The correlation between the mammography sensitivity and the breast cancer risk is presented clearly in Table 3.

5 Deep Learning models in Breast Density Classification

This section provides an overview and general architecture of convolution neural networks, Transfer Learning, and Ensemble Learning along with its pros and cons. These models significantly help to overcome the pitfalls in cancer detection at the early stage by learning from the existing information.

Table 3 Relationship of breast density with breast cancer

Review articles	Population used	Observation
Aloufi et al. (2022)	1140 screening mammograms collected from Saudi females	There was a significant correlation observed between elevated mammographic density and the likelihood of developing breast cancer
Ali et al. (2022)	750 women screening mammograms from Sulaimaniyah Breast Cancer, Iraq	The breast density profile of Sulaimaniyah, Iraq revealed that increased risk of breast cancer
Kim et al. (2022)	3.9 million Korean women's mammograms	Women who possess dense breast tissue are at an increased susceptibility to developing breast cancer
Mai Tran et al. (2022)	A total of 48.35507 women's mammograms have been collected. Among those 79,153 reports originated from a history of breast cancer	Women Possess a familial background characterized by a prevalence of breast cancer cases
Puliti et al. (2018)	16,752 women under the age of 49–54	The density of breast tissue significantly influences the occurrence and progression of breast cancer
Hanis et al. (2022)	Mammogram reports collected from Hospital University Sains, Malaysia	The density of mammograms has been identified as a substantial indicator of the likelihood of developing breast cancer
Choi et al. (2021)	290,448 women's report were taken from the Korean National Cancer Screening Program (KNCSP)	Korean-specific natural history parameters of breast cancer with higher dense breast
Barnard et al. (2022)	160,804 women with mammogram images were estimated based on racial	Association between Body Mass Index (BMI) and Breast density
Zhang et al. (2022)	11,130 women's mammograms have been analyzed	A significant correlation has been observed, indicating that individuals with a fully dense category have an increased probability of developing breast cancer

5.1 CNN

The convolutional neural network (CNN) is a popular Deep Learning architecture for image classification and recognition. Convolutional layers, pooling layers, and fully connected layers are just a few types of layers that make up a CNN architecture. The convolution layer integrates two functions, f and g combines two functions f and g to generate an output. The convolution layer applies a filter to the image to extract relevant features. As there are several sequences of convolution layers, the progress from one input layer to one output layer is performed accurately and precisely to extract features and this process is repeated to all the layers present in the CNN. When CNN applies filters to the input image, the output would be a complex feature map (Gargouri et al. 2022; Greenspan et al. 2016; Khan et al. 2020). The feature map is obtained by computing the following expression.

5.1.1 Feature map = Input image x feature detector

When the input image passes through a convolutional layer then the image would transform into a feature map or activation map and send the same to the pooling layer for reducing dimension. The activation function determines which bits of information progress to the next neuron, as identical to the neuron model of the human brain. Ech neuron in a neural network takes the value produced by the neurons in the layer below it as input and passes on the result of its processing to the layer above it (Joshva Devadas and Arumugam 2010).

$$(f \times g)\,(x, y) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} (i, j)\, f(x - i, x - j)$$

It is formally expressed as a discrete bi-dimensional convolution operation between two functions, f and g. Where i and j are the row and column indexes of the pixel, and x and y are the two variables of f and g respectively. The primary purpose of the pooling layer is to reduce the dimension of an image (Lawrence and Zhang 2019). Utilizing a pooling layer accelerates computation, conserves memory, and guards against overfitting. Pooling layers come in two common varieties: Max pooling and Average pooling. The Max pooling layer reduces background noise by returning the highest value from the portion of the image that the kernel has covered. It eliminates the noisy activations, dimensionality reduction, and denoising. The average pooling layer, on the other hand, displays the average of all the values from the region of the image that the kernel has covered (Shrestha and Mahmood 2019).

To obtain more low-level features, the number of convolution layers and pooling layers may be expanded. However, this will require additional computational power depending on the complexity of the image. To classify the images, the output is later flattened and fed into a standard neural network. For learning non-linear combinations of high-level features, fully connected layers are incorporated thus reducing the cost. Due to this image is flattened into a column vector and the flattened output is fed forward as input to the neural network. The training process is carried out with the back propagation method and the same is applied to all the iterations. Now, the model can categorize images using the SoftMax Classification method across several epochs by identifying dominant and specific low-level features (Simonyan and Zisserman 2015; Thomaz et al. 2017; Li et al. 2021a).

It's common knowledge that deep learning models need a lot of training data to function properly. Since deep learning is effective when there are several training possibilities applied to the model (Vargas-Hakim et al. 2021). The CNN can be trained from scratch if large enough training samples are given to them. The most challenging task involved selecting hyperparameters, such as the number of layers, dropout rates for each layer, filter sizes, acquisition of knowledge regarding to choose the regularization parameters and kind of suitable activation function. As a result, the entire training process usually takes a long time and requires a strong GPU. The CNN can be trained from scratch if large enough training samples are given to them. The most challenging task involved selecting hyperparameters, such as the number of layers, dropout rates for each layer, filter sizes, acquisition of knowledge regarding to choose the regularization parameters, and the kind of suitable activation function. As a result, the entire training process usually takes a long time and requires a strong GPU (Turay and Vladimirova 2022). Table 4 depicts the comparison of various approaches.

5.2 Transfer learning

In 1976, Stevo Bozinovski and Ante Fulgosi introduced a mathematical and geometrical model of transfer learning. The use of transfer learning in the context of training a neural network using a dataset representing letters of computer terminals was first published in another paper in 1981. Experimental research confirms the existence of both positive and negative types of transfer learning, both of which made use of datasets including images of letters A-Z. In 1993, Thrun developed the discriminability-based transfer (DBT) algorithm to emphasize the significance of Transfer Learning. In 2016, Andrew Ng stated that transfer Learning would be the next factor in determining the commercial success of Machine Learning. Recently, Transfer learning has been eminent in the field of deep Learning (Bozinovski 2020; Prasad et al. 2021; Li et al. 2019).

One common method for training a CNN model is Transfer Learning. The main role of Transfer Learning is to adapt the data from a network that has already been trained to carry out a similar function but different task (Rafiq and Albert 2022). Transfer Learning is more efficient and simpler to implement because it does not require a large labeled data set for training. The three most common conditions can be used to facilitate transfer learning (Shah 2020). They are shallow tuning, fine-tuning and deep learning. Shallow tuning just modifies the classification layer to make it suitable for the new task, while leaving the weights of the other layer to remain same. Fine-tuning is a method of gradually training the

Table 4 Comparison of various approaches

Models	Advantages	Disadvantages
CNN without a pre-pre-trained model	Fully automated with reduced manpower Simple and easy to understand and implement	Substantial dataset is required to train the model Computation time is more
Transfer learning	Model efficiency is high Only less amount of data is required Less training time	Pre-trained model may not fit some specific case Training time is more to train the pre-trained model A substantial quantity of data is required to train a model from its original state
Ensemble learning	Higher accuracy when combining various model	Very difficult to interpret Computation time is increased Complexity increases due to stacking

subsequent layers by adjusting the learning parameters until a major performance gain is achieved. Lastly, deep learning tries to relearn all the weights of the deployed pre-trained network from end to end (Kandel and Castelli 2020). Figure 4 depicts the architecture of transfer Learning.

The field of medicine has recently given a significant amount of attention to the concept of employing transfer Learning rather than training a full CNN it starts with random initialization (Abdelhafiz et al. 2019). Abdelhafiz stated that due to the high cost and small size of the available datasets, medical image analysis is increasingly turning towards transfer learning to improve accuracy and efficiency. In addition, the time it takes radiologists to collect data and categorize it can be substantial. Training a deep CNN also requires a large amount of memory and CPU time.

5.3 Pretrained model

In computer vision, transfer Learning is often implemented through the use of pre-trained models. Pretrained models, as defined by Abd-Elsalam are those that have already been trained to solve a certain problem using a large benchmark dataset. Table 5 shows the summary of the Pretrained model (Abd-Elsalam et al. 2020).

5.3.1 AlexNet

AlexNet was the first pre-trained CNN to achieve performance levels superior to the current gold standard approaches for classification and object detection tasks. Alex Krizhevsky and his colleagues proposed this model in 2013 hence the name is AlexNet. It is comprised of eight layers, each with its own set of learnable parameters. The model has five distinct layers. Including a max pooling level, and a fully connected layer. Each of these levels, excluding the output level, makes use of the rectified Linear Unit (Relu) activation function. The Relu's activation function was also found to dramatically accelerate the training process by

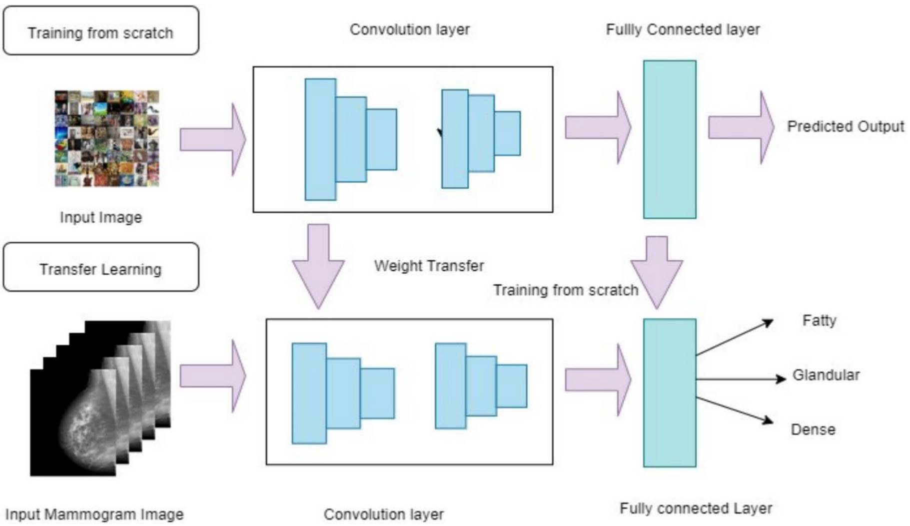


Fig. 4 Transfer learning

Table 5 Pretrained model summaries

Model	Model variant	Parameters	Size (MB)	Depth	Top 5 accuracy	Top 1 accuracy	Input size	Activation on function	Dropout rate
LeNet (1998)	LeNet-5	60,000		5	73	66	32,321	Tan h	–
AlexNet (2012)	AlexNet	62.3 M	227	8	84.6	63.3	227,227	Relu	0.5
VGGNet (2014)	VGG 16	138	528	23	90.1	71.13	224,224	Relu	–
	VGG 19	143.6	549	26	90	71.13			
ResNet (2015)	ResNet-50	25.6	548	168	92.1	74.9	224,224	Relu	–
	ResNet-152s	60.4	232	152	94.29	78.57			
GoogleNet (2015)	Inception-V3	23.8	92	159	93.7	77.9	229,229	Relu	0.7
Xception (2016)	Xception	23.3	88	126	94.5	79	229,229	–	–
DenseNet (2017)	DenseNet-121	8.2	33	121	93.34	76.39	224,224	Relu	–
	DenseNet-169	14.3	57	169	93.2	76.2			
	DenseNet-201	20.2	80	201	93.6	77.3	224,224	Relu6	–
MobileNet (2017)	MobileNet-V1	4.2	16	88	90.1	71.3			
	MobileNet-V2	3.5	13	88	92.1	74.9			
ShuffleNet (2017)	ShuffleNet						224,224	Relu	–
EfficientNet (2019)	EfficientNet-b0	5.3	88	–	93.5	76.3	224,224	Relu	–
	EfficientNet-b3	12	29	–	95.6	81.7			
	EfficientNet-b7	66	–	–	97.1	84.4			

a factor of about six. Additionally, dropout layer is used to prevent overfitting of the model. The image net dataset is used to train the model. Almost 14 million images spread across and 1000 varieties make up the ImageNet collection (Krizhevsky et al. 2012).

5.3.2 VGGNet

Simonyan and Zisserman of Oxford University proposed the visual Geometry Group (VGG) network in 2014. The necessity to speed up the training process and decrease the number of parameters used in the subsequent layers led to the development of VGGNet. The impact of network depth was examined by academics from the University of Oxford under the assumption that conventional filters are relatively small. Moreover, they demonstrated that increasing the depth to 16–19 layers was associated with a considerable improvement in the outcome. The architecture input is a fixed-size input with the value 224×224. Moreover, the VGG model was able to increase the network's effectiveness and broaden its receptive field while simultaneously reducing the number of parameters. This was accomplished by stacking many layers of convolution with a very small kernel size. The authors examined the number of combinations with different depths of 9,11,16 and 19 layers (Simonyan and Zisserman 2015).

5.3.3 ResNet

To solve the vanishing gradient problem, the concept of residual blocks in design was developed. The skip connections strategy is used in the residual network. Some intermediate layers between the activation layer and the next layer are skipped over by the skip connection called residual block (Veit et al. 2016). ResNet is a stack of these surplus blocks. Among the many advantages of including this type of skip connection. One of the notable advantages of regularization is its ability to circumvent the influence of any layer on the performance of the design, provided that layer exists. Consequently, this property enables the training of deep neural networks without encountering the challenges associated with vanishing or exploding gradients. According to the authors, adopting this kind of network makes optimization easier and allows for a significant depth improvement. This network employs a 34-layer simple network architecture that was influenced by VGG-19 and the shortcut connection was added towards the end of the process. The presence of shortcut connections in the design leads to the subsequent transformation of the network into a residual network (Lizzi et al. 2019).

5.3.4 GoogleNet

One of the key developments in the field of CNN was the inception network. Inception network currently has three version, known as Inception Version 1, 2 and 3. The initial version of GoogleNet launched in 2014. The aim of this study is to identify the most effective local structure and subsequently develop it in sequential stages, leading to the development of a multi layered network. The following the publication of inception-v1 in 2014, the authors proceeded their model, focusing on enhancing its performance through improvements in accuracy and reductions in time complexity. In particular, Inception V3 was proposed in 2016 by Szegedy et al. (2017).

5.3.5 XceptionNet

Chollet who initially presented the Xception architecture. Xception is an acronym that stands for extreme inception. The distinguishing characteristic of this methodology lies in its utilisation of depth-wise separable convolution to substitute the inception modules within the inception network, afterward being followed by a further pointwise separable convolution. It is entirely, the network is made up of 71 layers, and it has 22.9 million different parameters. According to the findings of an experiment that was carried out by Kandel and Castelli in the year 2020. A notable benefit of this network is its ability to achieve depth while utilizing a limited number of parameters (Chollet 2017).

5.3.6 DenseNet

In 2017, Haung introduced DenseNet, which stands for densely connected convolutional network. The purpose of this study was to achieve the establishment of a high level of connectivity in the channel-wise concatenation. In this particular design, the input for each subsequent layer is the preceding feature map, thereby addressing the problem of vanishing gradients. The author stated that the inclusion of dense connections in the model also contributes to a reduction in the overall number of parameters utilized. This phenomenon occurs due to the network's usage of feature map data from the preceding stage and at each subsequent layer, instead of creating new parameters. The densely connected architecture of this network has achieved a reduction in parameter count by a factor of five while preserving the number of layers in comparison to the ResNet architecture (Huang and Lin 2021).

5.3.7 MobileNet

Howard was the one who initially proposed the idea of portable CNN architecture. The author suggested depth-wise separable convolution in place of the conventional convolutions used in the earliest models in order to create a lighter weight model. The point wise convolution and the depth wise convolution are two separate processes that make up the depth wise separable convolution. The resolution multiplier and the width multiplier are both examples of global hyper parameters that have been incorporated into this design as a means of controlling the input image's channel depth and resolution, respectively. While the model was being developed to meet the requirements of the user, these hyper parameters assisted in providing a trade-off between miniaturization of the model creator (Howard et al. 2017).

5.3.8 ShuffleNet

The ShuffleNet architecture was offered by the Megvii group in 2019, and two additional operational elements were also presented at that time. These features consisted of pointwise group convolution and channel shuffle, both of which were designed to reduce the amount of computation cost and maintain high-level accuracy. In recent years, CNN architectural designs have increasingly incorporated billions of floating-point calculations every second. This is done in an effort to improve accuracy. Because it can perform approximately 10–50 mega floating-point operations per second. ShuffleNet is ideally suited for use in mobile applications, where the available processing power is more constrained. As a result of this, the aforementioned architecture exhibits superiority over MobileNet in terms of reduced top-1 error rates during the process of ImageNet classification. Additionally, it exhibits a computational speed that is 13 times faster than AlexNet in practical applications, while maintaining an equivalent level of accuracy and generating results with comparable precision (Bobo et al. 2004).

5.3.9 EfficientNet

Tan and LE introduced scaling and in the same year, EfficientNet distinguishes from other network architectures. The authors demonstrated that it is possible to successfully scale up ResNet and MobileNet designs by using compound coefficients, which they did so by presenting this architecture. In order to provide further clarity, the designers have put forth a proposed methodology that ensures the proportional adjustment of all dimensions. While simultaneously preserving the healthy relationship between dimension and the network. The dimensions encompassed in this context consist of picture resolution. Which pertains to the size of the image, depth, which refers to the number of layers, and width, which denotes the number of channels (Tan and Le 2019).

5.4 Ensemble learning

Figure 5 illustrates, Deep Ensemble modelling combines the predictions of multiple neural network models to minimise generalisation error. It is used to boost the quality of the final model's performance. Baseline classifiers that have been trained on input data and can make predictions are the building block of every ensemble. An aggregate forecast is then derived from the individual techniques that can be used to enhance the machine learning procedure (Putten and Bamford 2023). Table 6 depicts the strengths and weakness of pretrained models.

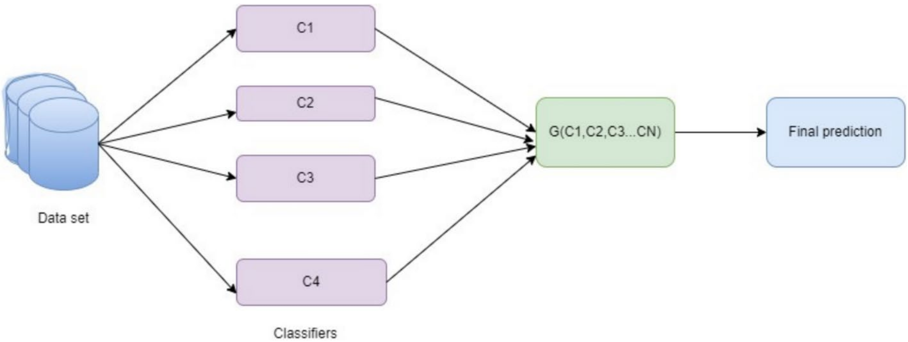


Fig. 5 Ensemble learning

Table 6 Strength and weakness of pretrained models

Model	Model variant	Merits	Demerits
LeNet (1998)	LeNet-5	Successful milestones for optical character task	LeNet-5 type structure is not enough to achieve high recognition capability
AlexNet (2012)	AlexNet	Features are not lost much Relu does not limit output	This model is swallow and struggles to learn features from image It takes more time to get high accuracy than future model Normal distribution weight initialization cannot solve gradient vanishing
VGGNet (2014)	VGG 16 VGG 19	VSS's receptive fields are much smaller than AlexNet	VGG required more memory and parameter Cost is high
GoogleNet (2015)	Inception V3	GoogleNet is faster than VGG pre-trained GoogleNet is 96mb and inception V3 is 92 MB	More parameters induce over fitting Parameters explosion on Inception layer
Xception (2016)	Xception	The accuracy is higher than inception model	Each layer's feature maps are spliced with the previous layer and duplicated
DenseNet (2017)	DenseNet-121 DenseNet 169 DenseNet 201	Efficiency decreasing the gradient disappearance problem	Each layer within the system is interconnected with other layers, resulting in potential duplication of data across these connected layers
MobileNet (2017)	MobileNet V1 MobileNet V2	Fewer parameters better classification accuracy than others	MobileNet is much smaller in size than others
ShuffleNet (2017)	ShuffleNet	Reduced the cost of computation while retaining accuracy	Property of ShuffleNet prevents communication between channel groups and degrades representation
EfficientNet (2019)	EfficientNet-b0 EfficientNet-b3 EfficientNet-b7	Efficient net achieved greater accuracy and efficiency with fewer parameters Increasing the number of chances will enhance overall capacity	More data transfer as a result of numerous channels On hardware accelerators EfficientNet perform poorly

6 Reviewed application of breast density classification method

This section covers three reviewed applications of breast density using Convolution Neural Network (CNN), Transfer Learning (TL), and Ensemble Learning (EL).

6.1 Reviewed application of CNN based classification of breast density method

Breast density was classified using a deep convolution neural network trained on the BIRADS database. According to Wu, the model was trained and evaluated using nearly 200 k labeled pictures from a clinically realistic dataset of screening mammography from four different perspectives. The dataset included 19,939 grouped into class-0, 85,665 class-1, 83,852 class-2, and 11,723 class-3 images. Due to the model's ability to utilize a large and varied clinically relevant dataset of high-resolution images. It is able to accomplish the task as well as human experts (Wu et al. 2018; Lizzi et al. 2019).

According to the research, the CNN can detect masses and distinguish dense and non-dense tissue in any type of breast tissue. The method demonstrated classification accuracies of 95.6% and 97.72% for non-dense and dense tissue respectively. The model faces an issue in effectively recognizing mass and non-mass categories since it uses the same CNN structure for both density categorization and identifying masses in breasts with varied densities (Bandeira Diniz et al. 2018).

Benitez presented a hybrid algorithm UNET architecture based on CNN and identify-based clustering using K-means clustering. 384 mammograms served as the training data in this model and manually segmented by an expert. In order to segment the regions of interest, this method integrated CNN based segmentation of fibro glandular tissue with a clustering algorithm. This model underwent 8064 iteration of training and the intersection over union showed accuracy of 93%. The network produces binary classification as its output (Benitez et al. 2022).

Multi view DL technique for BIRADS density assessment of mammograms were produced by Nguyen. To predict BIRADS and density scores, the gathered characteristics are subsequently consolidated and inputted into a light gradient boosting machine (Light GBM) classifier. Two benchmark datasets were used in the experiments. The outcome showed that the F1 score margin on clinical dataset and DDSM dataset were + 5% + 10% respectively. These results showed that how important multi view information fusion for improving the precision of breast cancer risk prediction (Nguyen et al. 2022).

Li constructed Deep CNN that effectively used to estimate the amount of breast tissue in full-field digital mammography (FDM). One million picture patches were created and used in the process of training DCNN. The (PD) Percent density was determined using (PMD) Probability map of breast density) as a starting point by dividing the dense region of the breast. Backpropagation and mini-batch stochastic gradient descent (mSGD) were used in each training cycle to optimise the parameters. Furthermore, feature-based learning obtained DC = 0.620 and r = 0.75 whereas DCNN obtained DC = 0.76 and r = 0.94. These findings showed that the prospective value of DCNN is more trustworthy for automated categorization of breast density and risk prediction (Li et al. 2021b).

Mohammed investigated breast density classifiers on deep learning to effectively differentiate between scattered-dense and heterogeneously dense categories of breast density. Further, intended to provide a potential automated tool to aid radiologists in the process of designating a BIRADS category. A convolution neural network-based model was trained using 22,000 mammogram images as training samples to determine how effectively is distinguished the breast density groups. The classifier's efficacy was evaluated using receiver operating characteristics (ROC) curves and the area under the curve (AUC). The obtained AUC was 0.9421 and the accuracy rose steadily as the training sample size increased (Mohamed et al. 2017a).

Saffari suggested an automatic method for determining mammographic breast density in accordance with the ACR BIRADS. Over 20,000 mammography images were used to train a deep CNN, which was then tested using an augmented dataset. In accordance with the fatty and dense category, MLO projections obtained 99% and CC projections obtained 96%. Mohammed used a CNN, a type of deep learning architecture in order to build a breast density classifier for two class. Six-fold cross validation was used for CNN training and validation. The AUC of breast density classifier was determined and the obtained AUC is 0.95 when using mammography images from two views MLO and CC (Mohamed et al. 2017b) Table 7 gives the summary of CNN without pretrained model for breast density classification.

Table 7 Summary of CNN without pretrained model for breast density classification

Authors	Database	Purpose	CNN Based Method	Class Obtained	Augmentation	Cross Validation	Accuracy	AUC
Diniz et al. (2018) (Bandeira Diniz et al. 2018)	DDSM	Breast density and breast mass classification	CNN from scratch	2 Class dense and non-dense	No	No	94.8	–
Wu et al. (2018)	Clinical Dataset	Breast density classification	CNN from scratch	2 Class dense and non-dense	No	No	82.5	69.4
Mohamed et al. (2017a)	ImageNet	Breast density classification	CNN from scratch	2 Heterogeneous dense and scattered dense	No	Yes	–	0.94
Nguyen et al. (2022)	CBIS-DDSM	Breast density classification	CNN from scratch	2 Class dense and non-dense	Yes	No	–	
Li et al. (2021b)	DDSM	Breast density classification	CNN from scratch	2 Class dense and non-dense	No	Yes	–	0.91
Mohammed et al. (2017b)	Clinical Dataset	Breast density classification	CNN from scratch	2 Class dense and non-dense	No	Yes	82.5	0.95
Duffy et al. (2018)	Clinical Dataset	Breast density classification	CNN from scratch	2 Class dense and non-dense	Yes	Yes	99-MLO 96-CC	–
Shi et al. (2019)		Breast density classification	CNN from scratch	2 Class dense and non-dense	Yes	Yes	94.6	
Nithya and Santhi (2021) (2021)	MIAS Dataset	Breast density classification	CNN from scratch	2 Class dense and non-dense	Yes	Yes	98.5	–

6.2 Breast density classification based on transfer learning

A radiomics approach to mammographic density categorization using dilation and attention guided learning has been reported (Li et al. 2018). This technique involves extensive training from beginning to end as well as it includes advanced pre-processing of mammographic image. Both the clinical and publicly available dataset were used to verify and test the model. Furthermore, it was shown that multi-view mammogram images such as CC and MLO views were significant in improving the accuracy of breast density classification problems. Overall, this model was 88.7% accurate.

Yi developed a model using ResNet 50. This model developed to analyse breast density, determine breast laterality and classify two-dimensional mammographic image. Totally 3034 two dimensional mammographic images used as a training data obtained from the DDSM database. AUC obtained before augmentation was 0.75. The data augmentation technique eventually raising the AUC to 0.93. In addition, it is possible to use DCNN to automatic semantic labelling of 2D mammograms, and this can be done even with relatively limited dataset. However, automated breast density classification is more complex and thus required large dataset (Yi et al. 2019).

Lehman conducted a study and investigated the clinical implementation of deep learning model for evaluating breast density among patients who were undergoing screening digital mammography. Furthermore, the methodology of deep learning can be employed to asses' breast density, without imposing limitations based on previous surgical interventions or other breast related procedures. Moreover, DL has the ability to address issues with the present regulations and assist physicians in order to give accurate information and optimal utilization of additional screening resources (Lehman et al. 2019).

Zhao introduced a novel approach called the Bilateral View Adaptive Spatial and Channel Attention Network (BASCNET) which utilises ResNet-50 as the underlying architecture. The primary objective of this network is to achieve fully automated breast density classification. By incorporating data from both the left and right breast it dynamically captures distinctive features in terms of spatial and channel dimensions. The DDSM and INBREAST datasets were used for the training and validation process. The achieved accuracy was 85.10% and 90.5% (Zhao et al. 2021).

Gandomkar examined the breast density classification into various categories such as fatty or dense based on BIRADS categorization. A network architecture known as inception V3 was utilised to train a dataset consisting of 3813 images sourced from nine distinct mammography devices and three different vendors. A segmentation process was employed to isolate breast tissue from the background and pectoral muscle in the MLO view. The bounding box of the breast was utilised for the purpose of cropping and resizing the input image of the network. The aforementioned technique had a Cohen's Kappa coefficient of 0.775 and an accuracy rate of 83.33 percent (Gandomkar et al. 2019).

Wu proposed a methodology for classifying breast density by grouping breast tissue into two categories dense and non-dense. The three-layer CNN utilised all four perspectives of the mammogram as its input. The findings indicated that the super classes achieved an accuracy rate of 82.5%. Additionally, the four-class density classification demonstrated a macro average AUC of 0.934. Specifically, class-0 achieved an AUC of 0.971, class-1 achieved an AUC of 0.859, AUC of 0.905 in class-2 and AUC 1 in class-3.

Kate proposed technique to examined breast tissue density classification using VGG 16 and Inception V3 model. Further, image pre-processing technique was used to extract the foreground image as well as to improve the image quality by reducing noise appear in that image. The Inception V3 model obtained 97.98% and VGG 16 got 91.92% using DDSM dataset (Kate and Shukla 2022).

Ma created multi path DCNN to classify digital mammography image into one of four BIRADS category. Around 2068 mammogram instances were used for the breast density based on BIRADS and obtained the accuracy of 80% for the dense category and 89% for the non-dense category. Xue modelled breast density assessment as a machine intelligence problem that automatically extracts features of image and dynamically improves density classification accuracy in clinical environments. Variety of deep learning networks are investigated in order to extract the image features automatically. Transfer learning is used to retain the pre-trained model that were already trained using clinical 2D digital mammography images. As part of the human–machine gaming process, a comprehensive reinforcement network is implemented. In this study, mammograms were pre-processed using CNN models (Xue et al. 2020).

Lin proposed the CNN incorporated with AlexNet, DenseNet, and ShuffleNet. In this study, breast density and breast mass such as benign and malignant were combined into a single model. This was done so that the researchers could identify the difference between breast density and the two forms of breast mass. A comparison was made between the pre and post-data augmentation accuracies of the three models. Before applying data augmentation techniques, the accuracy of AlexNet for both the training and testing sets was recorded as 40.47%. DenseNet achieved an accuracy of 90%, while ShuffleNet achieved an accuracy of 96.48%. Additionally, ShuffleNet's accuracy was measured to be 38.57%. Following the application of data augmentation techniques, the training and testing accuracies of AlexNet were observed to be 99.35% and 95.46% respectively. Similarly, DenseNet exhibited accuracies of 99.91% and 97.84% for its respective sets (Lin et al. 2021).

Trivizakis employed the DenseNet model to categorise the three levels of breast density, resulting in an accuracy rate of 73.9%. In their study, Mohammed et al. were able to achieve an area under curve (AUC) value of 0.95 when employing an enhanced iteration of the AlexNet model for the purpose of classifying breast density (Trivizakis et al. 2019). Table 8 Summarizes the previous work on mammogram breast density analysis with pretrained model.

Pretrained models are employed in the categorization of breast density to automatically evaluate the density of breasts based on mammography images (Li et al. 2018). It is crucial to consider breast density since it significantly increases the risk of breast cancer.

Table 8 Summary of work included in mammogram breast density with pretrained model

Author	Database	CNN Method	Augmentation	Cross Validation	Accuracy (%)	AUC
Li et al. (2018)	INbreast and clinical database	ResNet50	Yes	Yes	88.7 and 70	97.2
Lehman et al. (2019)	INbreast	ResNet-18	Yes	No	95	94
Gandomkar et al. (2019)	INbreast	Inception-V3	Yes	Yes	63.9	82.1
Mohamed et al. (2017b)	INbreast	AlexNet	Yes	Yes	59.6	82
Zhao et al. (2021)	DDSM and INbreast	ResNet	No	–	–	–
Kate and Shukla (2022)	DDSM	Inception-V3 VGG16	Yes	Yes	97.98 91.92	–
Rigaud et al. (2022)	DDSM	EfficientNetB0	No	No		–
Yi et al. (2019)	DDSM	ResNet50	No	No	68	93
Busaleh et al. (2022)	DDSM and INbreast	ResNet50 DenseNet EfficientNetB0	Yes	No	91.36 90.89 89.23	99.51 97.44

Therefore, it is vital to accurately evaluate breast density in order to make well-informed decisions on screening and treatment. Historically, radiologists have traditionally evaluated breast density by visually inspecting mammography pictures. Nevertheless, this method is based on personal opinion and may need a significant amount of time (Yi et al. 2019). Pretrained models provide a more impartial and automated method for evaluating breast density. Figure 6 presents a comparison of the performance of various pretrained models on a task of classifying breast density (Lehman et al. 2019). The models under consideration for comparison encompass AlexNet, ResNet, Inception, EfficientNet, Vgg16, DenseNet, and MobileNet. In general, pretrained models have the capacity to fundamentally transform the method of evaluating breast density. nevertheless, it is crucial to confront the obstacles linked to this technology prior to its widespread implementation in clinical practice.

6.3 Breast density classification based on ensemble learning

The predictions from various neural network models are combined through ensemble learning to lower prediction variance and generalisation error. Ensemble learning techniques can be categorised into various learning algorithm or models and training data (Rigaud et al. 2022). It integrates the data obtained from the several model to provide precise and effective decision. The ensemble strategies can be broken down into three distinct categories namely bagging, stacking, and boosting (Busaleh et al. 2022). Following are the few studies focused on classification of mammogram breast density based on ensemble techniques. Table 9 Outlines the work included in ensemble learning.

Kumar introduced and described a model for classification to predict breast density. The classification model has two-stage process. The first stage is made up of single four class classifier, while another stage is made up of an ensemble of six class classifiers. The study made use of 480 mammograms that were taken from the DDSM dataset. 90.80% accuracy has been attained for the categorization (Kumar et al. 2017). The deep learning system that detects breast density with noisy labels regularization was presented by Justaniah. Female participants of 1395 were taken from a multi-centre and assessed by three highly qualified radiologists. These mammograms were classified according to BIRADS categories (Haque et al. 2018). The dataset was spliced into training data, validation data and testing

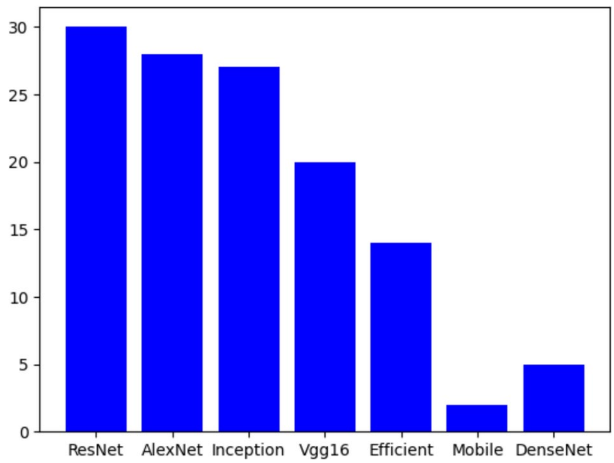


Fig. 6 Quantitative performance analysis of pretrained models in breast density classification method

Table 9 Summary of work included in ensemble learning

Author	Sub architecture	Ensemble learning approach	Database	Purpose	Augmentation	Cross validation	Accuracy (%)	Class obtained
Kumar et al. (2017)	Stage 1-Five Neural network stage 2 6 Neural network	Model averaging	DDSM	BIRADS Classification	No	No	90.2	Stage 1:4 Class Stage 2:2 Class
Haque et al. (2018)	Multilayer perception (MLP) Support Vector Machine	Model averaging	MIAS	BIRADS Classification	No	Yes	92.6	4 Class
Justaniah et al. (2022)	VGG-19 ResNet DenseNet 121 Wide ResNet 50 EfficientNet B1	Model averaging	ImageNet	BIRADS Classification	No	Yes	84.6	4 Class
Azour and Boukerche (2022)	CNN EfficientNet	Soft voting	CBIS-DDSM	Pathology Classification	Yes	Yes	96.5	2 Class

data. Training data includes 892 mammograms, testing and validation includes 279 and 224 respectively. The accuracy and kappa indices of the ensemble model that were found to be 0.85 and 0.71 respectively (Justaniah et al. 2022).

Haque developed an automated system capable of analysing mammograms and effectively differentiating breast density. This system utilises a two-dimensional discrete cosine transform (D-DCT) and a principal component analysis (PCA) extract a minimal feature set from the mammogram image (Azour and Boukerche 2022). The extracted features are subsequently utilised as input for three classifiers, namely, multi-layer perceptron (MLP), support vector machine (SVM), and K nearest neighbour (KNN). The combination of the output from multiple classifiers was subjected to a majority vote, resulting in a significantly high level of classification performance.

7 Summary on TL, CNN and ensemble models

Transfer learning can significantly reduce the amount of labelled data needed for training, as it leverages pre-trained models on large datasets. This approach can lead to faster model development and improved accuracy in breast cancer detection. Additionally, it allows for the adaptation of advanced neural network architectures fine-tuned for image recognition tasks. However, transfer learning may not always capture the specific nuances of medical imaging data, as pre-trained models are often based on general datasets that differ from mammograms. This can result in suboptimal performance if the source and target domains are not sufficiently similar. Moreover, there may be a reliance on potentially outdated or biased datasets, which could affect the accuracy and reliability of predictions.

Ensemble modelling can enhance breast cancer detection in mammogram images by combining the predictions of multiple models, which leads to more robust and accurate results. This approach mitigates the weaknesses of individual models by leveraging their diverse strengths, reducing the likelihood of false positives and negatives. Furthermore, ensemble methods can improve generalization, making the system more reliable across different patient populations and imaging conditions.

Implementing ensemble models in breast cancer detection can be computationally expensive, requiring significant resources for training and maintaining multiple models. Additionally, combining predictions from multiple models can introduce complexity in model interpretation, making it harder for clinicians to understand and trust the results. Finally, the integration of ensemble models into existing clinical workflows may face resistance due to increased complexity and the need for specialized expertise.

7.1 Contemporary model

Over the past decade, a variety of deep-learning models have been proposed to improve model prediction. Several techniques can be integrated with deep learning models to improve performance, such as self-attention, transformer models, semi supervised learning, interpretable techniques, and uncertainty estimation. This section introduces some of the recently proposed pioneering breast cancer prediction models that have recently been proposed.

A novel Computer-Aided Diagnosis (CAD) framework leveraging Artificial Intelligence (AI) techniques has been developed to detect and classify breast cancer images (Al-Tam et al. 2024). The framework is meticulously structured into two pipelines (Stage 1 and Stage 2) and was trained and evaluated using various multimodal ultrasound and mammogram datasets. The framework explores the potential of combining cutting-edge AI techniques and the ViT-based Resnet50 model to create an innovative CAD model for detecting and classifying breast cancer. Grad-CAM is used to visualize the model's predictions for better understanding. Grad-CAM plays a crucial role in this framework by providing visual explanations for the model's predictions. It highlights the specific regions in the breast cancer images that the model focused on when making its classification, thereby offering insights into the decision-making process. This transparency helps radiologists verify the model's accuracy and understand its reasoning, ultimately increasing trust in AI-assisted diagnoses.

A Yolo-based model for breast cancer detection was proposed (Prinzi et al. 2024). A gradient-free Eigen-CAM method is used to highlight all suspicious ROIs, including incorrect predictions, allowing us to integrate our model into a clinical decision support system. This enhancement allows clinicians to visualize which areas of the image the model focuses on when making predictions, thus providing insights into its decision-making process. By highlighting the regions of interest, Eigen-CAM helps in understanding the rationale behind both correct and incorrect predictions. This transparency increases trust and reliability in the model's outputs, essential for its integration into clinical workflows.

A quantum spinal network is proposed to detect breast cancer using mammogram images (MG) (Sathish et al. 2024). The system is designed using a Deep Quantum Neural Network (DQNN) and SpinalNet, and it achieves 90.3% accuracy, 90.9% True Negative Rate (TNR), and 90% True Positive Rate (TPR). An NLM filter is applied to mammogram images to detect breast cancer, Edge Attention-SegNet is applied to extract features, and Q-SpinalNet is introduced by incorporating DQNN and SpinalNet. Edge Attention-SegNet focuses on enhancing the key structural details in mammogram images by emphasizing the edges and contours, which are critical for accurate feature extraction. This allows for more precise detection of abnormalities by highlighting the relevant areas that could indicate the presence of cancer. By doing so, the model can better differentiate between benign and malignant tissues, improving the overall diagnostic accuracy. Using this model, we achieve maximum accuracy, TNR, and TPR values of 90.3%, 90.9%, and 90% respectively.

CNN-based mammogram image enhancement techniques have been proposed and shown promising results in improving breast cancer prediction accuracy (Singla et al. 2020). By leveraging convolutional neural networks, these methods can enhance critical features on mammograms, facilitating more reliable and early detection of cancerous tissues. This makes CNN-based approaches a valuable tool in the ongoing efforts to improve diagnostic outcomes in breast cancer screening.

The author proposes a hybrid model to improve representativeness in breast cancer detection (Kalpana and Selvy 2024). Transfer learning and probabilistic principal component analysis are used for feature extraction and classification. The hyperparameters of these models are optimized using firefly binary grey optimization and metaheuristic moth flame lion, respectively. Metaheuristic optimization techniques, such as firefly binary grey optimization and metaheuristic moth flame lion, help enhance model performance by efficiently searching for optimal hyperparameter settings. These techniques explore a wide solution space to identify configurations that improve accuracy and reduce error rates. By optimizing hyperparameters, the models can better capture complex patterns within the data, leading to more robust and reliable predictions.

Researcher (Chai et al. 2024) proposed an uncertainty-based interpretable deep neural network for breast cancer outcome prediction called Uncertainty-based Integrated Semisupervised Net UISNet to overcome the interpretability challenge. An innovative multitask deep neural network called UISNet was proposed for predicting breast cancer outcomes. In addition to incorporating prior biological pathway knowledge, UISNet utilizes patient heterogeneity information to improve prediction. A model identified 20 genes as associated with breast cancer, of which 11 have been proven to be associated with breast cancer in previous studies.

The longitudinal trajectory of mammographic breast density is one method of understanding a woman's breast cancer risk over time. Using data from a large Swedish mammography cohort, (Illipse et al. 2023) fitted three joint models (cumulative, current value and slope, and current value association structures). Across all models, MD trajectory was associated with BC risk, but models with cumulative association structures and with current value and slope association structures performed better.

Mammographic breast density is well-established and cross-sectional international data suggest that it is a strong risk factor for breast cancer (Jiang et al. 2023). However, retrospective data from Korea show that a change in density over time is associated with a change in breast cancer diagnosis risk. We hypothesized that there is a difference in the rate of change in breast density in women who develop breast cancer.

To predict breast cancer risk, authors (Karaman et al. 2024) proposed the Longitudinal Mammogram Risk (LoMaR) model, which combines a transformer architecture with convolutional feature extraction. Even using just the most recent mammogram, it achieves state-of-the-art prediction results. With the help of longitudinal mammograms, we extend a state-of-the-art machine-learning model to predict future breast cancer risks. The results of our study show that predicting future breast cancer risk with longer histories is more accurate.

7.2 Interpretable models

Interpretability is crucial because it allows clinicians to understand the reasoning behind a model's predictions, which can enhance trust in its recommendations. This transparency is essential for making informed decisions about patient care and for identifying potential errors or biases in the model. Additionally, interpretable models can facilitate better communication between healthcare professionals and patients, improving overall treatment outcomes. Techniques such as feature visualization, SHAP values, and LIME can be employed to enhance the interpretability of these models. Additionally, integrating attention mechanisms and providing clear visual explanations can help clinicians understand the model's decision-making process. Regular consultations with domain experts can further refine these interpretability methods to align with clinical needs. Achieving model interpretability in these contexts presents several challenges, including the complexity of deep learning models, which often function as "black boxes." Furthermore, the need to balance accuracy with interpretability can limit the model's performance, as more interpretable models may sacrifice precision. Additionally, ensuring that the interpretability methods are clinically relevant and understandable to medical professionals adds another layer of complexity.

A hybrid explainable deep model for cardiac prediction is proposed by Wani et al. 2024a, which uses CNN and the Light Gradient Boosting method to effectively learn representational features. The integration process involves initially employing the CNN to extract high-level features from the cardiac data. These features are then fed into the Light Gradient Boosting framework, which refines the model's predictions by leveraging its efficient gradient-based optimization. This combined approach enhances the model's accuracy and interpretability, leading to improved prediction performance in cardiac analysis. The model uses one of the most widely used XAI techniques, SHAP, to provide comprehensive and detailed explanations. SHAP scores are computed by calculating the contribution of each feature to the prediction made by the model. This is done by considering all possible combinations of features and using Shapley values from cooperative game theory to fairly distribute the prediction among the features. The result is a clear and interpretable insight into how each feature influences the model's output. The explanations are provided in the form of several graphs, which will assist medical practitioners in improving their diagnostic abilities. A similar explainable hybrid framework is proposed using CNN and XGBoost for lung cancer detection (Wani et al. 2024b). One of the main advantages of using SHAP scores is that they provide a consistent and unified measure of feature importance, which helps in understanding the model's decision-making process. Additionally, SHAP scores enhance transparency and trust in machine learning models by making complex models interpretable to stakeholders who may not have technical expertise. Furthermore, they allow for the identification of potential biases or errors in the model, leading to more robust and fair predictions.

In response to the rapid expansion of the Internet of Medical Things (IoMT) and the increasing prevalence of automation, end-users have become increasingly apprehensive about their trust. In (Wani et al. 2024d), author reviewed different explainable AI Artificial Intelligence (AI) frameworks that explain the Internet of Medical Things have emerged as essential tools for addressing trust concerns. A similar hybrid model using CNN and Light Gradient Boosting is proposed for breast cancer prediction (Yan et al. 2024). Model predictability is improved by these enhanced representational features. SHAP scores are used to understand model prediction and feature importance in model output.

An interpretable AI system (Yan et al. 2024) that provides an overall cancer risk assessment from multimodal US images could enhance patient outcomes and reduce unnecessary biopsies. Deep learning frameworks have previously been used in medical imaging, and their superiority over hand-crafted features has been demonstrated. However, deep learning's black-box nature has made it difficult to build trust among human experts. As part of the work, the author proposes the use of multimodal US images to generate an interpretable AI system based on domain knowledge.

An interpretable AI system using multimodal ultrasound images is presented for breast cancer classification (Klanecek et al. 2024). A domain-based interpretable AI system that predicts cancer risk from multimodal US images could improve patients' outcomes. Despite their superiority over hand-crafted features, deep learning frameworks lack the trust of human experts due to their black-box nature. This understandable MUP-Net was comparable to popular black-box models and outperformed junior radiologists while remaining competitive with senior radiologists. In the case of inaccurate malignant probabilities, explainable features can help readers stay on course. A domain expert supervises the AI learning process, making sure that output features are explainable. As part of clinical practice, learned contribution scores are calculated.

A unique methodological approach was proposed to determine how early the BCR model can identify morphological changes associated with oncogenic processes (Yan et al. 2024). Different attribution methods such as Class Activation Map (CAM), Grad-CAM, Integrated Gradients, Guided backpropagation, and Input x Gradients are used to visualize tumor region growth. Results showed that the model relies more on the signal from the breast with cancer in patients where breast cancer is already screen-detected, but less on the breast without cancer.

7.3 Model hyperparameter optimizations

Proper hyperparameter optimization can significantly improve the performance and accuracy of a learning model. It ensures that the model is well-tuned to the specific dataset, leading to better generalization and reduced overfitting. Additionally, it can also enhance the model's efficiency by reducing training time and resource consumption. Common methods for hyperparameter optimization include grid search, random search, and Bayesian optimization (Eid and Abualigah 2024). Grid search involves exhaustive searching over a specified parameter grid, while random search randomly samples from the parameter space. Bayesian optimization uses probabilistic models to predict which hyperparameters are likely to lead to better performance, allowing for more efficient exploration. Metaheuristic algorithms, such as genetic algorithms and particle swarm optimization, are also employed for hyperparameter optimization. These algorithms mimic natural processes, like evolution and swarm behavior, to efficiently explore the hyperparameter space (Yassen et al. 2024). They are particularly useful in navigating complex, high-dimensional spaces where traditional methods may be less effective.

In breast cancer detection, bio-inspired deep model hyperparameter optimization can play a crucial role in enhancing diagnostic accuracy. Techniques such as genetic algorithms can be used to fine-tune deep learning models, ensuring that they are optimized for the intricate patterns in medical imaging data (Wani et al. 2024c). This approach not only improves the detection rates but also reduces false positives, leading to more reliable and early diagnosis of breast cancer. Despite its benefits, hyperparameter optimization faces several challenges. The search space can be vast and complex, making it computationally expensive and time-consuming to explore thoroughly. Additionally, the optimal hyperparameters can vary significantly between different datasets and models, requiring a tailored approach for each scenario. Moreover, there is often a trade-off between the depth of exploration and the available computational resources, which can limit the effectiveness of the optimization process (Wani et al. 2024c).

8 Data set and performance evaluation

In this section, the resources or data sources as well as the evaluation measures to measure the performance of the model, that research scientist employ in the process of generating and reviewing the studies that are being examined are described.

8.1 Mammogram dataset

Mammography datasets are increasingly needed by researchers to create diagnostic system to design, test and evaluate automatic breast cancer in (CAD) Computer Aided Detection System. Few mammography datasets are publicly available that can be use by the researchers in the creation of breast cancer prediction tools, whereas some of the mammogram databases are private. Despite the fact, that there have been some recent studies at mammography classification, employing confidential mammogram records has been a common approach. To provide enhanced insights for researchers lacking access to proprietary datasets, we choose to concentrate on the datasets that are available for the public (Thiagarajan and Ganesan 2012).

8.1.1 Mammographic Image Analysis Society

One of the earliest and oldest databases is MIAS database. It is confidential dataset compiled by the research organization in the U.K. There are a total of 161 samples and 322 images included, including those of benign, malignant and normal mammogram. Annotation images in the form of circles surrounding the ROI are included in the dataset (Alsolami et al. 2021).

8.1.2 Digital Database for Screening Mammography (DDSM)

The Institution referred to as the University of South Florida established DDSM dataset which was publicly released in 1999. The dataset comprises mammography images accompanied by pertinent information, including the patient's age, mammogram screening date, the nature of the anomaly detected, and the breast density. The dataset containing the highest number of mammograms which consist of 2620 instances, with each instance comprising four distinct views. The cases have been systematically categorized into 43 volumes, each volume containing images that are classified as normal, malignant or benign (Hath 2018).

8.1.3 Curated Breast Imaging Subset of DDSM (CBIS-DDSM)

The DDSM dataset has undergone recent updates, resulting in the inclusion of numerous updated images. The revised iteration of the Digital Database for Screening Mammography (DDSM) dataset is referred as the curated Breast Imaging Subset of the Digital Databse for Screening mammography (CBIS DDSM) dataset. The primary aim of this dataset is to improve the classification results in order to address the health-related issues. Region of Interest (ROI) annotation is updated by the CBIS-DDSM, which also assess segmentation and specialized approaches. The dataset is used to train and evaluate any breast cancer detection model, The dataset comprises over 1000 images which have been categorised into two distinct groups: calcification and mass (Tahmoush and Samet 2008).

8.1.4 INBreast

The breast research group released a publicly accessible mammography dataset, which is commonly referred to INBreast dataset and initially released to the public in 2010. The dataset was acquired from the Breast Cancer Institute at porto Hospital of St john, encompassing all available sources. It included a total of115 cases that were formatted in DICOM, each of which contained 90 images in two different views (CC and MLO). According to BIRADS categorical, the images of dense mammogram, calcification and normal tissue were included in the INBreast dataset. It is no longer possible to locate the dataset (Moreira et al. 2012).

8.1.5 Breast Cancer Digital Repository (BCDR)

Breast cancer resource is currently being employed in the field of mammography classification. The dataset comprises a total of 736 lesions obtained from 344 patients who have undergone biopsy and received positive result by the radiologist. The dataset including MLO and CC views, that has enabled lesion counter coordinate collection (Li and Nishikawa 2015).

8.1.6 Other datasets

There are many factors considered as the Risk factor of breast cancer. One among the risk factor associated with culture and society. Therefore, the provision of local and public mammography datasets is imperative for facilitating global research efforts at the detection and classification of breast abnormalities in women. The Abdul Aziz University Breast Cancer Mammogram Dataset KAU-BCMD) was one of the first datasets to be established in Saudi Arabia. The dataset comprises a total of 1416 cases, each of which includes the information of the BIRADS category. the dataset consists of 5662 images, which includes two views of the breast. The MIRACLE dataset is another publicly available dataset that contain 204 mammogram images from 196 cases or patients. The Magic 5 Italian dataset was compiled and stored the information received from a wide variety of medical facilities. It contains 967 cases altogether, depending on the type of pathology. The LLNL dataset has a total of 197 images, both of which have been saved in the format known as Image Cytometry Standard format (ICS). Patient records and biopsy results are also included in the dataset. The IRAM dataset is a unique dataset that combines the contents of several other datasets (Becker 2005; Antoniou et al. 2009; Tangaro et al. 2008; Karssemeijer et al. 2012). Table 10 gives the concise overview of the dataset.

Breast cancer prediction mammogram datasets from different demographics affect the generalizability of the model because the data may not represent the entire population. This can lead to biased results and limit the model's ability to predict breast cancer in diverse groups accurately. Therefore, it is essential to ensure that training datasets are diverse and inclusive. One approach is to perform subgroup analysis, which involves evaluating the model's performance on different demographic groups separately. Another method is to use fairness metrics that assess how equitably the model performs across these groups. Additionally, cross-validation techniques can be employed to ensure the model's robustness

Table 10 Concise overview of the dataset

Dataset	Year of Commence	Views	Classification Classes	Image Type	Image Format	No of Patients	Overall, Image	Description
DDSM	1999	4	Normal malignant and benign	SFM	JPEG	2620	10,480	Non-Standard compression files Non-Standard format
BCDR	2012	2	Malignant and benign	FFDM	DICOM	736	736	Lesion location available Different resolution image standard format
MIAS	2015	1	Normal and abnormal	SFM	PGM	161	322	Low-resolution image Only one view is available
INBreast	2017	2	Normal, malignant and benign	FFDM	DICOM	419	419	Limited size standard format
CBIS-DDSM	2017	2	Mass and calcification	SFM	DICOM	1566	10,239	The standard version of DDSM image pathology information
KAU-BCMD	2020	3	BI-RADS-5 categories	SFM	DICOM	1416	5662	Annotated images are available

and accuracy across diverse populations. Another method to increase dataset diversity is to collect data from multiple geographical regions and include participants from various ethnic backgrounds. Additionally, collaborations with international research institutions can help acquire datasets that reflect different lifestyle factors and genetic predispositions. Implementing these strategies will enhance breast cancer prediction models' robustness and fairness. However, implementing these strategies comes with challenges such as data privacy concerns, which can hinder data sharing across institutions and regions. Moreover, logistical and financial constraints may limit the ability to collect and integrate data from diverse populations. Additionally, ensuring that the datasets are representative of all demographics requires ongoing collaboration and coordination between researchers worldwide.

High-quality datasets are crucial for accurate model outcomes, as they ensure the data is representative and free from bias. Without this, the model's predictions can be skewed and unreliable. Ensuring diverse and comprehensive data collection can significantly enhance the model's ability to generalize across different populations.

8.2 Evaluation Criteria

The popularity and durability of standard works are contingent upon their quality, necessitating the use of evaluation criteria such as accuracy, precision, recall, and F1 measure to gauge prospective results. Accuracy measure is a metric that quantifies the proportion of correctly categorized occurrences, specifically indicating the accuracy in identifying persons with mass lesions and benign disease patients (Tajerian et al. 2021) with the use of a confusion matrix. The metric of sensitivity quantifies the accuracy of correctly identifying positive cases, specifically referring to the proportion of patients with cancerous pathology among those who have malignant abnormalities (Qin et al. 2022). The measure of specificity pertains to the proper identification of individuals with abnormal mammograms (Dewangan et al. 2022). The percentage of accurately classified positive predictions is shown by precision. The metric employed in the study emphasizes the influence of precision and recall on memory by using harmonic means (Bouzar-Benlabiod et al. 2023). The confusion matrix is a graphical representation that offers a complete evaluation of performance results in classification procedures, considering both actual and expected outcomes. (Rajpal et al. 2021). Cross-validation is a widely used data analysis approach that involves the resampling of data through the partitioning of the dataset into K folds. The process is then iterated K times to extract concealed information. (Singh and Singh 2022).

$$ACC = \frac{(TP + TN)}{(Tp + TN + FP + FN)}$$

$$S_n = \frac{(TP)}{(TP + FN)}$$

$$S_p = \frac{(TN)}{(TN + FN)}$$

$$P_r = \frac{(TP)}{(TP + FP)}$$

$$F_1\text{-score} = \frac{(2 \times Precision \times recall)}{(Precision + recall)}$$

True Positive (TP): Both the current and predicted classes exhibit positive characteristics. False Positive (FP): Misclassification arises when the observed class is positive, but the predicted class is negative. True Negative (TN): Both the desired class and the true class exhibit negative values. False Negative (FN): The imagined result is positive, and the final outcome is negative. Table 11 depicts the confusion matrix.

9 Discussion

Several concerns were identified during the surveys, which added to our progression to the closing remarks.

9.1 Effect of the dense breast on breast cancer

As the correlation between dense breast and breast cancer becomes more apparent. It is imperative that women should be aware of the connection between breast density and cancer. According to the study by Albeshan, breast density is the only significant risk factor linked to a diagnosis of breast cancer (Albeshan et al. 2019). Women who fall into BIRADS categories C and D showed a significant decrease in mammography sensitivity and a linear relationship was identified between mammography sensitivity and automated volumetric breast density. According to a recent review by Santiago, breast cancer prevention and early detection practices can benefit from an understanding of breast density (Santiago-Rivas et al. 2017). From Table 3 it is observed that mammogram density is a major risk, and it affects the sensitivity of mammograms leads to a false positive rate. Moreover, when employing a completely automated breast density measurement, it has been found that women who grouped under the category of extremely dense breast had a four times higher risk of developing cancer than other women.

9.2 Recent trends in breast density classification

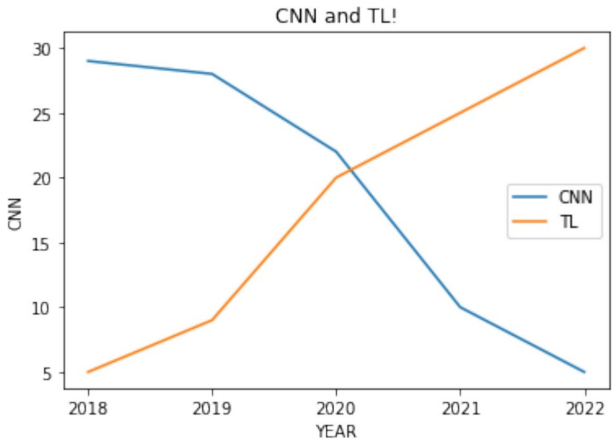
In order to classify breast density, researchers using a deep learning architecture explored the CNN. Initially, CNN began to train the model from scratch, because of the time complexity and the large amount of data required to process. The approach has progressed to transfer learning via the pre-trained model and ensemble model. Moreover, the use of traditional CNN has been reduced slowly in the past 5 years (Kamal et al. 2023). In Contrast, the usage of transfer learning along with a pre-trained model has elevated over the last 5 years for the task of mammogram breast density classification which is illustrated in Fig. 7.

Figure 8 depicts the researcher's utilization of various models for the classification of breast density regularly. Over 60% of the existing research is dedicated to the exploration

Table 11 Confusion matrix

	Predicted		
		Negative	Positive
Actual	Negative	TP	FP
	Positive	FP	TP

Fig. 7 Usage of CNN and TL over the years



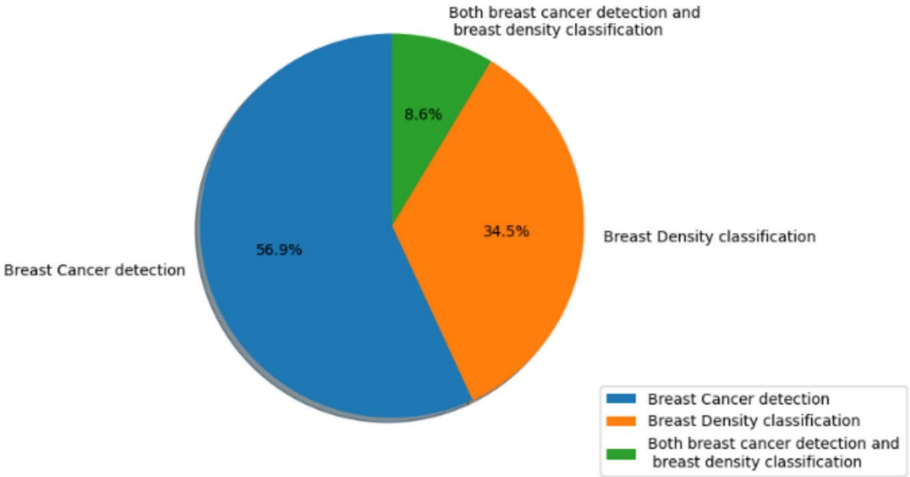


Fig. 8 Comparing the popularity of objectivity among researchers

of transfer learning and its potential applications in addressing challenges related to mammography density classification. Accordingly, Fig. 9 illustrates a significant proportion of researchers primarily focus on the detection of abnormalities, neglecting the analysis of mammogram density. In contrast, a notable proportion of researchers, specifically around 9%, effectively used both mammogram density and breast cancer diagnosis in their studies (Arya and Saha 2022). Along with the pre-trained and CNN model, the most highly used dataset by the researchers for the estimation of mammograms is illustrated in Fig. 10 and this helps to improve the classification model performance. Figure 11 also depicts the commonly employed evaluation criteria for breast density classification.

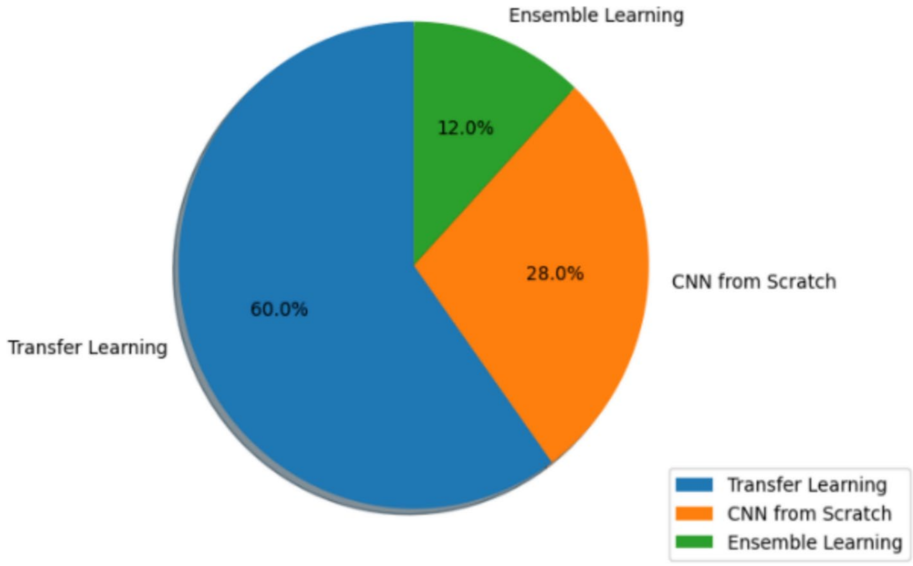
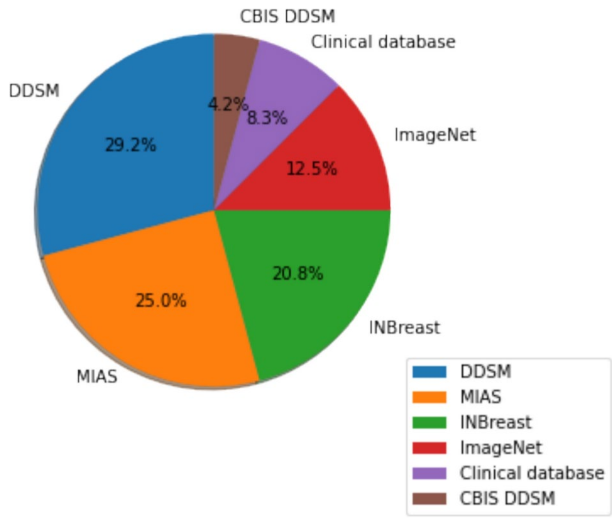


Fig. 9 Popularity in various techniques

Fig. 10 Distribution of dataset



10 Challenges and research direction in mammogram breast density analysis

Deep learning models have demonstrated remarkable precision in predicting breast density from mammography images, thereby streamlining the data processing procedure. One of the primary obstacles and potential avenues for the advancement in the analysis of mammography breast density relies on the models that predict breast cancer at the early stage.

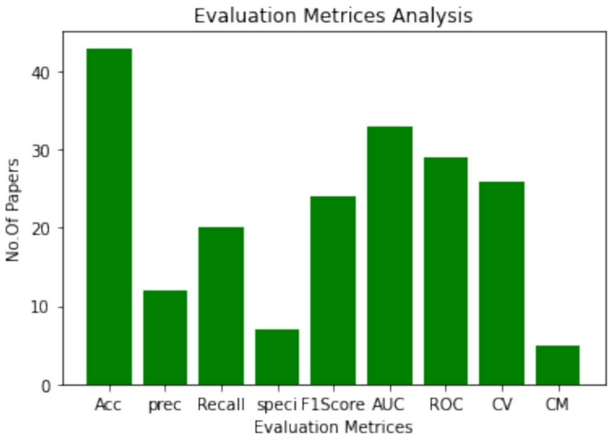


Fig. 11 Analysis of evaluation metrics

10.1 Major challenges in mammogram breast density classification

The deep learning model has numerous benefits in the realm of medical imaging. Moreover, their implementation by categorization of mammography images needs to cross out the obstacles mentioned below:

- Early prediction of breast cancer the potentially reduces the patient's risk thus increasing the overall lifespan of the patients.
- The mammography breast density measurement indicates its inadequacy in accuracy. identifying the presence of breast cancer (Gade et al. 2023).
- Due to the masking effect of mammogram images false positive rate is increased.
- Deep learning models used in the process of predicting breast cancer need to train the model with a greater number of mammographic images. But practically it is not feasible to collect the number of mammographic images.
- The choice of selection of an appropriate deep learning model itself is vital because the selection demonstrates its significant improvements in predicting the accuracy and effectiveness of breast density analysis.

11 Research direction

This section highlights the potential possibilities for future research in the application that uses deep learning techniques for mammography breast density analysis.

- The analysis and detection of mammogram breast density should be conducted prior to the identification of breast cancer, thus deploying multitasking techniques for breast cancer classification (Sehgal et al. 2022).
- Utilization of advanced deep learning models may enhance classification accuracy by exploring concatenation or hybrid methodologies.
- Cooperative and collaborative work between radiologists and scientists needs to be built for the investigation of multivariate hybrid data to carry out the subsequent analysis.

- Further, this survey paves directions for the deployment of an ensemble learning model with pre-trained model to classify breast density.
- Furthermore, the need of a new technology or tool may help the patient to understand the analysis of the breast density in a much better way by offering better guidance to increase their survival rate (Priyadarshani and Singh 2023).
- In addition, it is imperative that an early prediction of breast cancer incorporates mammogram breast density as an initial stage. This approach has the potential to decrease the false prediction and improve overall classification accuracy.

12 Conclusion

This study provides a comprehensive assessment on the current trends and limitations in utilising deep learning models for the interpretation of mammography breast density. A comprehensive evaluation is conducted on the latest techniques and models, such as CNN, TL, EL, and other, to assess their effectiveness and constraints. In addition, this paper examines several data sources and evaluation metrics used in this field. The study identifies significant research gaps and future objectives in the field of breast density classification and breast cancer detection. These areas include early prediction, model selection, and masking effect. Addressing these gaps is crucial to improve the accuracy, and reliability of breast density classification and breast cancer detection. This survey is expected to be a great resource and guide for scholars and practitioners who are interested in this important and promising field of research.

Author contributions T. Joshva Devadas: Conceptualization, review, interpretation and supervision. Hemalatha K: Methodology, interpretation, drafting and review. J. Naveenkumar: Methodology, interpretation, drafting and review. Jeba Prasanna Idas: Conceptualization, data curation and drafting.

Funding Open access funding provided by Vellore Institute of Technology.

Data availability No datasets were generated or analysed during the current study.

Declarations

Competing interests The authors declare no competing interests.

References

Abdelhafiz D, Yang C, Ammar R, Nabavi S (2019) Deep convolutional neural networks for mammography: advances, challenges and applications. BMC Bioinform. https://doi.org/10.1186/s12859-019-2823-4

Abdelrahman L, Al Ghamdi M, Collado-Mesa F, Abdel-Mottaleb M (2021) Convolutional neural networks for breast cancer detection in mammography: a survey. Comput Biol Med 131:104248. https://doi.org/10.1016/j.compbiomed.2021.104248

Abd-Elsalam NM, Fawzi SA, Kandil AH (2020) Comparing different pre-trained models based on transfer learning technique in classifying mammogram masses. In: 2020 30th international conference on computer theory and applications (ICCTA), December 1. https://doi.org/10.1109/ICCTA52020.2020.9477663

Advani SM, Zhu W, Demb J, Sprague BL, Onega T, Henderson LM et al (2021) Association of breast density with breast cancer risk among women aged 65 years or older by age group and body mass index. JAMA Netw Open 4(8):e2122810. https://doi.org/10.1001/jamanetworkopen.2021.22810

Albeshan SM, Hossain SZ, Mackey MG, Peat JK, Al Tahan FM, Brennan PC (2019) Preliminary investigation of mammographic density among women in Riyadh: association with breast cancer risk factors and implications for screening practices. Clin Imaging 54:138–147. https://doi.org/10.1016/j.clinimag.2019.01.002

Ali KAS, Fateh SM (2022) Mammographic breast density status in women aged more than 40 years in Sulaimaniyah, Iraq: a cross-sectional study. J Int Med Res 50(12):030006052211397–030006052211397. https://doi.org/10.1177/03000605221139712

Allison E, Edirimanne S, Matthews J, Fuller SJ (2022) breast cancer survival outcomes and tumor-associated macrophage markers: a systematic review and meta-analysis. Oncol Ther. https://doi.org/10.1007/s40487-022-00214-3

Aloufi AS, AlNaeem A, Almousa AS, Hashem A, Malik M, Altahan F et al (2022) Mammographic breast density and breast cancer risk in the Saudi population: a case–control study using visual and automated methods. Brit Inst Radiol. https://doi.org/10.1259/bjr.20211197

Alsolami AS, Shalash W, Alsaggaf W, Ashoor S, Refaat H, Elmogy M (2021) King Abdulaziz University Breast Cancer Mammogram Dataset (KAU-BCMD). Data 6(11):111. https://doi.org/10.3390/data6110111

Al-Tam RM, Al-Hejri AM, Alshamrani SS, Al-antari MA, Narangale SM (2024) Multimodal breast cancer hybrid explainable computer-aided diagnosis using medical mammograms and ultrasound Images. Biocybern Biomed Eng 44(3):731–758. https://doi.org/10.1016/j.bbe.2024.08.007

Antoniou ZC, Giannakopoulou GP, Andreadis II, Nikita KS, Ligomenides PA, Spyrou GM (2009) A web-accessible mammographic image database dedicated to combined training and evaluation of radiologists and machines. In: 2009 9th International conference on information technology and applications in biomedicine, November 1. https://doi.org/10.1109/ITAB.2009.5394465

Arya N, Saha S (2022) Multi-modal classification for human breast cancer prognosis prediction: proposal of deep-learning based stacked ensemble model. IEEE/ACM Trans Comput Biol Bioinf 19(2):1032–1041. https://doi.org/10.1109/TCBB.2020.3018467

Azour F, Boukerche A (2022) An efficient transfer and ensemble learning based computer aided breast abnormality diagnosis system. IEEE Access. https://doi.org/10.1109/access.2022.3192857

Bandeira Diniz JO, Bandeira Diniz PH, Azevedo Valente TL, Corrêa Silva A, de Paiva AC, Gattass M (2018) Detection of mass regions in mammograms by bilateral analysis adapted to breast density using similarity indexes and convolutional neural networks. Comput Methods Programs Biomed 156:191–207. https://doi.org/10.1016/j.cmpb.2018.01.007

Barnard ME, Martheswaran T, Van Meter M, Buys SS, Curtin K, Doherty JA (2022) Body mass index and mammographic density in a multiracial and multiethnic population-based study. Cancer Epidemiol Biomark Prev. https://doi.org/10.1158/1055-9965.epi-21-1249

Becker GS (2005) The potential power of reference image databases and oncologic imaging. Am Coll Radiol 2(7):565–567. https://doi.org/10.1016/j.jacr.2005.03.019

Benitez CS, Pertuz S, Arponen O, Lääperi A-L, Rinta-Kiikka I (2022) Automatic dense tissue segmentation in digital mammography images based on fully convolutional network and intensity-based clustering. In: 2022 IEEE Colombian conference on applications of computational intelligence (ColCACI), July 1. https://doi.org/10.1109/ColCACI56938.2022.9905248

Bobo J, Hudley C, Michel C (2004) The black studies reader. Psychology Press New York

Bodewes FTH, van Asselt AA, Dorrius MD, Greuter MJW, de Bock GH (2022) Mammographic breast density and the risk of breast cancer: a systematic review and meta-analysis. Breast (Edinburgh, Scotland) 66:62–68. https://doi.org/10.1016/j.breast.2022.09.007

Bond-Smith D, Stone J (2018) Methodological challenges and updated findings from a meta-analysis of the association between mammographic density and breast cancer. Cancer Epidemiol Biomark Prev 28(1):22–31. https://doi.org/10.1158/1055-9965.epi-17-1175

Bouzar-Benlabiod L, Harrar K, Yamoun L, Khodja MY, Akhloufi MA (2023) A novel breast cancer detection architecture based on a CNN-CBR system for mammogram classification. Comput Biol Med 163:107133–107133. https://doi.org/10.1016/j.compbiomed.2023.107133

Bozinovski S (2020) Reminder of the first paper on transfer learning in neural networks, 1976. Informatica. https://doi.org/10.31449/inf.v44i3.2828

Braithwaite D, Miglioretti DL, Zhu W, Demb J, Trentham-Dietz A, Sprague B, Kerlikowske K (2018) Family history and breast cancer risk among older women in the breast cancer surveillance consortium cohort. JAMA Intern Med 178(4):494. https://doi.org/10.1001/jamainternmed.2017.8642

Busaleh M, Hussain M, Aboalsamh HA, Fazal-e-Amin, Al Sultan SA (2022) TwoViewDensityNet: two-view mammographic breast density classification based on deep convolutional neural network. Mathematics 10(23):4610. https://doi.org/10.3390/math10234610

Chai H et al (2024) An uncertainty-based interpretable deep learning framework for predicting breast cancer outcome. BMC Bioinform 25(1):1–17. https://doi.org/10.1186/s12859-024-05716-7

Choi E, Jun JK, Suh M, Jung K-W, Park B, Lee K et al (2021) Effectiveness of the Korean National Cancer Screening Program in reducing breast cancer mortality. Npj Breast Cancer. https://doi.org/10.1038/s41523-021-00295-9

Chollet F (2017) Xception: deep learning with depthwise separable convolutions. https://openaccess.thecvf.com/content_cvpr_2017/papers/Chollet_Xception_Deep_Learning_CVPR_2017_paper.pdf

Cohn BA, Terry MB (2019) Environmental influences on mammographic breast density in California: a strategy to reduce breast cancer risk. Int J Environ Res Public Health 16(23):4731. https://doi.org/10.3390/ijerph16234731

Daly AA, Rolph R, Cutress RI, Copson ER (2021) A review of modifiable risk factors in young women for the prevention of breast cancer. Breast Cancer Targets Ther. https://doi.org/10.2147/BCTT.S268401

Dayaratna S, Jackson R (2022) The conundrum of breast density; guidance for healthcare providers. Best Pract Res Clin Obstet Gynaecol. https://doi.org/10.1016/j.bpobgyn.2022.01.012

Dewangan KK, Dewangan DK, Sahu SP, Janghel R (2022) Breast cancer diagnosis in an early stage using novel deep learning with hybrid optimization technique. Multimedia Tools Appl. https://doi.org/10.1007/s11042-022-12385-2

Duffy SW, Morrish OW, Allgood PC, Black R, Gillan MG, Willsher P et al (2018) Mammographic density and breast cancer risk in breast screening assessment cases and women with a family history of breast cancer. Eur J Cancer 88:48–56. https://doi.org/10.1016/j.ejca.2017.10.022

Eid MM, Abualigah L (2024) Machine learning in public health forecasting and monitoring the Zika virus. Metaheuristic Optim Rev 01(02):1–11

Gade A, Dash DK, Mita Kumari T, Ghosh SK, Tripathy RK, Pachori RB (2023) Multiscale analysis domain interpretable deep neural network for detection of breast cancer using thermogram images. IEEE Trans Instrum Meas 72:1–13. https://doi.org/10.1109/tim.2023.3317913

Gandomkar Z, Suleiman ME, Demchig D, Brennan PC, McEntee MF (2019) BI-RADS density categorization using deep neural networks. In: Medical imaging 2019: image perception, observer performance, and technology assessment. https://doi.org/10.1117/12.2513185

Gargouri N, Mokni R, Damak A, Sellami D, Abid R (2022) An automatic breast computer-aided diagnosis scheme based on a weighted fusion of relevant features and a deep CNN classifier. IET Image Proc. https://doi.org/10.1049/ipr2.12572

Greenspan H, van Ginneken B, Summers RM (2016) Guest editorial deep learning in medical imaging: overview and future promise of an exciting new technique. IEEE Trans Med Imaging 35(5):1153–1159. https://doi.org/10.1109/tmi.2016.2553401

Hanis TM, Arifin WN, Haron J, Wan Abdul Rahman WF, Ruhaiyem NIR, Abdullah R, Musa KI (2022) Factors influencing mammographic density in Asian Women: a retrospective cohort study in the northeast region of Peninsular Malaysia. Diagnostics (Basel, Switzerland) 12(4):860. https://doi.org/10.3390/diagnostics12040860

Haque MSM, Hassan MR, BinMakhashen GM, Owaidh AH, Kamruzzaman, J (2018) Breast density classification for cancer detection using DCT-PCA feature extraction and classifier ensemble. In: Advances in intelligent systems and computing. Springer, Cham, pp 702–711. https://doi.org/10.1007/978-3-319-76348-4_68

Hath M (2018) Computer vision and pattern recognition group. www.eng.usf.edu website: http://marathon.csee.usf.edu/Mammography/database.html

Howard A, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, Andreetto M (2017) MobileNets: efficient convolutional neural networks for mobile vision applications. https://arxiv.org/pdf/1704.04861.pdf%EF%BC%89

Huang M-L, Lin T-Y (2021) Considering breast density for the classification of benign and malignant mammograms. Biomed Signal Process Control 67:102564. https://doi.org/10.1016/j.bspc.2021.102564

Illipse M, Czene K, Hall P, Humphreys K (2023) Studying the association between longitudinal mammographic density measurements and breast cancer risk: a joint modelling approach. Breast Cancer Res 25(1):1–11. https://doi.org/10.1186/s13058-023-01667-8

Jiang S, Bennett DL, Rosner BA, Colditz GA (2023) Longitudinal analysis of change in mammographic density in each breast and its association with breast cancer risk. JAMA Oncol 9(6):808–814. https://doi.org/10.1001/jamaoncol.2023.0434

Joshva Devadas T, Arumugam G (2010) Object oriented intelligent multi-agent system data cleaning architecture to clean preference based text data. Int J Comput Appl 9(8):34–44. https://doi.org/10.5120/1404-1894

Justaniah E, Aldabbagh G, Alhothali A, Abourokbah N (2022) Classifying breast density from mammogram with pretrained CNNs and weighted average ensembles. Appl Sci 12(11):5599. https://doi.org/10.3390/app12115599

Kaiser N, Fieselmann A, Ritschl L, Kappler S, Ravikumar N, Vesal S, Maier A (2019) Mammographic breast density classification using a deep neural network: assessment on the basis of inter-observer variability. In: Medical imaging 2019: image perception, observer performance, and technology assessment. https://doi.org/10.1117/12.2513420

Kalpana P, Selvy PT (2024) A novel machine learning model for breast cancer detection using mammogram images. Med Biol Eng Comput 62(7):2247–2264. https://doi.org/10.1007/s11517-024-03057-4

Kamal AM, Sakorikar T, Pal UM, Pandya HJ (2023) Engineering approaches for breast cancer diagnosis: a review. IEEE Rev Biomed Eng 16:687–705. https://doi.org/10.1109/rbme.2022.3181700

Kandel I, Castelli M (2020) Transfer learning with convolutional neural networks for diabetic retinopathy image classification: a review. Appl Sci 10(6):2021. https://doi.org/10.3390/app10062021

Karaman BK, Dodelzon K, Akar GB, Sabuncu MR (2024) Longitudinal mammogram risk prediction. In: Proceedings of medical image computing and computer assisted intervention—MICCAI 2024. https://doi.org/10.1007/978-3-031-72086-4_41

Karssemeijer N, Thijssen M, Hendriks J, van Erning L (2012) Digital mammography, 1998. Springer, Nijmegen

Kate V, Shukla P (2022) Breast tissue density classification based on gravitational search algorithm and deep learning: a novel approach. Int J Inf Technol. https://doi.org/10.1007/s41870-022-00930-z

Kehm RD, Walter EJ, Pereira A, White ML, Oskar S, Michels KB et al (2022) A comparison of various methods for measuring breast density and breast tissue composition in adolescent girls and women. Sci Rep 12(1):13547. https://doi.org/10.1038/s41598-022-17800-0

Khan A, Sohail A, Zahoora U, Qureshi A (2020) A survey of the recent architectures of deep convolutional neural networks. Artif Intell Rev 53:1–87

Kim S, Tran TXM, Song H, Ryu S, Chang Y, Park B (2022) Mammographic breast density, benign breast disease, and subsequent breast cancer risk in 3.9 million Korean women. Radiology 304(3):534–541. https://doi.org/10.1148/radiol.212727

Klanecek Z, Wang Y, Wagner T, Cockmartin L, Marshall N (2024) Longitudinal interpretability of deep learning based breast cancer risk prediction. Phys Med Biol. https://doi.org/10.1088/1361-6560/ad9db3

Kocer B (2021) The relationship between mammographic density and factors affecting breast cancer risk. Med Bull Sisli Hosp. https://doi.org/10.14744/semb.2021.30771

Kressin NR, Wormwood JB, Battaglia TA, Maschke AD, Slanetz PJ, Pankowska M, Gunn CM (2022) Women's understandings and misunderstandings of breast density and related concepts: a mixed methods study. J Women's Health. https://doi.org/10.1089/jwh.2021.0343

Krizhevsky A, Sutskever I, Hinton GE (2012) ImageNet classification with deep convolutional neural networks. Commun ACM 60(6):84–90. https://doi.org/10.1145/3065386

Kumar I, Bhadauria HS, Virmani J, Thakur S (2017) A classification framework for prediction of breast density using an ensemble of neural network classifiers. Biocybern Biomed Eng 37(1):217–228. https://doi.org/10.1016/j.bbe.2017.01.001

Kumar I, Virmani J, Bhadauria HS, Thakur S (2019) A breast tissue characterization framework using PCA and weighted score fusion of neural network classifiers. In: Classification techniques for medical image analysis and computer aided diagnosis. Academic Press, New York. https://doi.org/10.1016/b978-0-12-818004-4.00006-6

Kyanko KA, Hoag J, Busch SH, Aminawung JA, Xu X, Richman IB, Gross CP (2020) Dense breast notification laws, education, and women's awareness and knowledge of breast density: a nationally representative survey. J Gen Intern Med 35(7):1940–1945. https://doi.org/10.1007/s11606-019-05590-7

Lawrence T, Zhang L (2019) IoTNet: an efficient and accurate convolutional neural network for IoT devices. Sensors 19(24):5541. https://doi.org/10.3390/s19245541

Lee SE, Son N-H, Kim MH, Kim E-K (2022) Mammographic density assessment by artificial intelligence-based computer-assisted diagnosis: a comparison with automated volumetric assessment. J Digit Imaging 35(2):173–179. https://doi.org/10.1007/s10278-021-00555-x

Lehman CD, Yala A, Schuster T, Dontchos B, Bahl M, Swanson K, Barzilay R (2019) Mammographic breast density assessment using deep learning: clinical implementation. Radiology 290(1):52–58. https://doi.org/10.1148/radiol.2018180694

Lester SP, Kaur AS, Vegunta S (2022) Association between lifestyle changes, mammographic breast density, and breast cancer. Oncologist. https://doi.org/10.1093/oncolo/oyac084

Li Q, Nishikawa RM (2015) Computer-aided detection and diagnosis in medical imaging. Taylor & Francis, Boca Raton

Li S, Wei J, Chan H-P, Helvie MA, Roubidoux MA, Lu Y et al (2018) Computer-aided assessment of breast density: comparison of supervised deep learning and feature-based statistical learning. Phys Med Biol 63(2):025005. https://doi.org/10.1088/1361-6560/aa9f87

Li C, Xue D, Zhou X, Zhang J, Zhang H, Yao Y et al (2019) Transfer learning based classification of cervical cancer immunohistochemistry images. In: Proceedings of the third international symposium on image computing and digital medicine. https://www.semanticscholar.org/paper/Transfer-Learning-Based-Classification-of-Cervical-LiXue/df407e865a0077736089b17572f31fb80b7b921d

Li Z, Liu F, Yang W, Peng S, Zhou J (2021a) A survey of convolutional neural networks: analysis, applications, and prospects. IEEE Trans Neural Netw Learn Syst. https://doi.org/10.1109/tnnls.2021.3084827

Li C, Xu J, Liu Q, Zhou Y, Mou L, Pu Z et al (2021b) Multi-view mammographic density classification by dilated and attention-guided residual learning. IEEE/ACM Trans Comput Biol Bioinf 18(3):1003–1013. https://doi.org/10.1109/TCBB.2020.2970713

Li X, Qi Y, Lou M, Zhao W, Meng J, Zhang W, Ma Y (2022) Breast density measurement methods on mammograms: a review. Multimedia Syst 28(6):2367–2390. https://doi.org/10.1007/s00530-022-00955-1

Lin F, Sun H, Han L, Li J, Bao N, Li H et al (2021) An effective fine grading method of BI-RADS classification in mammography. Int J Comput Assist Radiol Surg 17(2):239–247. https://doi.org/10.1007/s11548-021-02541-8

Lin X, Wu S, Li L, Ouyang R, Ma J, Yi C, Tang Y (2023) Automatic mammographic breast density classification in Chinese women: clinical validation of a deep learning model. Acta Radiol. https://doi.org/10.1177/02841851231152097

Lizzi F, Laruina F, Oliva P, Retico A, Fantacci ME (2019) Residual convolutional neural networks to automatically extract significant breast density features. In: Computer analysis of images and patterns. Springer, Cham, pp 28–35. https://doi.org/10.1007/978-3-030-29930-9_3

Mai Tran TX, Kim S, Song H, Park B (2022) Family history of breast cancer, mammographic breast density and breast cancer risk: findings from a cohort study of Korean women. Breast 65:180–186. https://doi.org/10.1016/j.breast.2022.08.008

Mathur A, Taurin S (2022) What influence does mammographic density have on breast cancer occurrence? Expert Rev Anticancer Ther 22(5):445–447. https://doi.org/10.1080/14737140.2022.2065985

Mendes J, Matela N (2021) Breast cancer risk assessment: a review on mammography-based approaches. J Imaging 7(6):98. https://doi.org/10.3390/jimaging7060098

Miles RC, Lehman C, Warner E, Tuttle A, Saksena M (2019) Patient-reported breast density awareness and knowledge after breast density legislation passage. Acad Radiol 26(6):726–731. https://doi.org/10.1016/j.acra.2018.07.004

Mohamed AA, Berg WA, Peng H, Luo Y, Jankowitz RC, Wu S (2017a) A deep learning method for classifying mammographic breast density categories. Med Phys 45(1):314–321. https://doi.org/10.1002/mp.12683

Mohamed AA, Luo Y, Peng H, Jankowitz RC, Wu S (2017b) Understanding clinical mammographic breast density assessment: a deep learning perspective. J Digit Imaging 31(4):387–392. https://doi.org/10.1007/s10278-017-0022-2

Moini A, Salari E, Rashidi H, Maajani K, Abedi M, Bayani L, Alipour S (2022) Evaluation of the association of endometriosis and mammographic breast density, a cross-sectional study. BMC Womens Health. https://doi.org/10.1186/s12905-022-01663-8

Moreira IC, Amaral I, Domingues I, Cardoso A, Cardoso MJ, Cardoso JS (2012) INbreast. Acad Radiol 19(2):236–248. https://doi.org/10.1016/j.acra.2011.09.014

Murtaza Mendes J, Matela N (2021) Breast cancer risk assessment: a review on mammography-based approaches. J Imaging 7(6):98. https://doi.org/10.3390/jimaging7060098

Nazari SS, Mukherjee P (2018) An overview of mammographic density and its association with breast cancer. Breast Cancer 25(3):259–267. https://doi.org/10.1007/s12282-018-0857-5

Nguyen H, Tran S, Nguyen D, Pham H, Nguyen H (2022) A novel multi-view deep learning approach for BI-RADS and density assessment of mammograms. In: 2022 44th Annual international conference of the ieee engineering in medicine & biology society (EMBC). https://arxiv.org/pdf/2112.04490.pdf

Nithya R, Santhi B (2021) Mammogram density classification using deep convolutional neural network. J Instrum 16(01):P01019–P01019. https://doi.org/10.1088/1748-0221/16/01/p01019
Oza P, Sharma P, Patel S, Bruno A (2021) A bottom-up review of image analysis methods for suspicious region detection in mammograms. J Imaging 7(9):190. https://doi.org/10.3390/jimaging7090190
Pizzato M, Carioli G, Rosso S, Zanetti R, La Vecchia C (2022) Mammographic breast density and survival in women with invasive breast cancer. Cancer Causes Control 33(9):1207–1213. https://doi.org/10.1007/s10552-022-01590-x
Posso M, Louro J, Sánchez M, Román M, Vidal C, Sala M et al (2019) Mammographic breast density: how it affects performance indicators in screening programmes? Eur J Radiol 110:81–87. https://doi.org/10.1016/j.ejrad.2018.11.012
Prasad A, Pant S, Srivatzen S, Asokan S (2021) a non-invasive breast cancer detection system using FBG thermal sensor array: a feasibility study. IEEE Sens J 21(21):24106–24113. https://doi.org/10.1109/jsen.2021.3111684
Prinzi F, Insalaco M, Orlando A, Gaglio S, Vitabile S (2024) A Yolo-based model for breast cancer detection in mammograms. Cognit Comput 16(1):107–120. https://doi.org/10.1007/s12559-023-10189-6
Priyadarshani KN, Singh S (2023) Ultra-sensitive breast cancer cell lines detection using dual nanocavities engraved junctionless FET. IEEE Trans Nanobiosci 22(4):889–896. https://doi.org/10.1109/tnb.2023.3246106
Puliti D, Zappa M, Giorgi Rossi P, Pierpaoli E, Manneschi G, Ambrogetti D, Mantellini P (2018) Volumetric breast density and risk of advanced cancers after a negative screening episode: a cohort study. Breast Cancer Res. https://doi.org/10.1186/s13058-018-1025-8
Qin J, He Y, Ge J, Liang Y (2022) A multi-task feature fusion model for cervical cell classification. IEEE J Biomed Health Inform 26(9):4668–4678. https://doi.org/10.1109/jbhi.2022.3180989
Rafiq RB, Albert MV (2022) Transfer learning: leveraging trained models on novel tasks. In: Educational communications and technology: issues and innovations. Springer, Cham, pp 65–74. https://doi.org/10.1007/978-3-030-84729-6_4
Rajpal S, Agarwal M, Kumar V, Gupta A, Kumar N (2021) Triphasic DeepBRCA—a deep learning-based framework for identification of biomarkers for breast cancer stratification. IEEE Access 9:103347–103364. https://doi.org/10.1109/access.2021.3093616
Rampun A, Morrow PJ, Scotney BW, Wang H (2020) Breast density classification in mammograms: an investigation of encoding techniques in binary-based local patterns. Comput Biol Med 122:103842. https://doi.org/10.1016/j.compbiomed.2020.103842
Rehman K, Li J, Pei Y, Yasin A (2022) A review of machine learning techniques for the assessment of image grading in breast mammogram. Int J Mach Learn Cybern. https://doi.org/10.1007/s13042-022-01546-2
Rigaud B, Weaver OO, Dennison JB, Awais M, Anderson BM, Chiang T-YD et al (2022) Deep learning models for automated assessment of breast density using multiple mammographic image types. Cancers 14(20):5003. https://doi.org/10.3390/cancers14205003
Saffari N, Rashwan HA, Abdel-Nasser M, Kumar Singh V, Arenas M, Mangina E et al (2020) Fully automated breast density segmentation and classification using deep learning. Diagnostics 10(11):988. https://doi.org/10.3390/diagnostics10110988
Santiago-Rivas M, Benjamin S, Andrews JZ, Jandorf L (2017) Breast density awareness and knowledge, and intentions for breast cancer screening in a diverse sample of women age eligible for mammography. J Cancer Educ 34(1):90–97. https://doi.org/10.1007/s13187-017-1271-y
Sathish Kumar P, Shibu S, Mohan M, Kalaichelvi T (2024) Hybrid deep learning enabled breast cancer detection using mammogram images. Biomed Signal Process Control 95:106310. https://doi.org/10.1016/j.bspc.2024.106310
Schifferdecker KE, Tosteson ANA, Kaplan C, Kerlikowske K, Buist DSM, Henderson LM et al (2019) Knowledge and perception of breast density, screening mammography, and supplemental screening: in search of "informed." J Gen Intern Med 35(6):1654–1660. https://doi.org/10.1007/s11606-019-05560-z
Sehgal HD, Pratap Y, Kabra S (2022) Detection of breast cancer cell-MDA-MB-231 by measuring conductivity of Schottky source/drain GaN FinFET. IEEE Sens J 22(6):6108–6115. https://doi.org/10.1109/jsen.2022.3148117
Shah R (2020) Adaptive fuzzy network based transfer learning for image classification. In: 2020 IEEE international students' conference on electrical, electronics and computer science (SCEECS). https://www.semanticscholar.org/paper/Adaptive-Fuzzy-Network-based-Transfer-Learning-for-Shah/fd10c9f6ae9f20cafac87923b4d09fa61acffd76
Shamshiri AK, Alidoust M, Nokandei MH, Pasdar A, Afzaljavan F (2023) Genetic architecture of mammographic density as a risk factor for breast cancer: a systematic review. Clin Transl Oncol 25(6):1729–1747. https://doi.org/10.1007/s12094-022-03071-8

Shi P, Wu C, Zhong J, Wang H (2019) Deep learning from small dataset for BI-RADS density classification of mammography images. In: 2019 10th International conference on information technology in medicine and education (ITME). https://doi.org/10.1109/itme.2019.00034

Shrestha A, Mahmood A (2019) Review of deep learning algorithms and architectures. IEEE Access 7:53040–53065. https://doi.org/10.1109/access.2019.2912200

Simonyan K, Zisserman A (2015) Very deep convolutional networks for large-scale image recognition. In: International conference on learning representations. https://doi.org/10.48550/arXiv.1409.1556

Singh S, Singh S (2022) Dopingless negative capacitance ferroelectric TFET for breast cancer cells detection: design and sensitivity analysis. IEEE Trans Ultrason Ferroelectr Freq Control 69(3):1120–1129. https://doi.org/10.1109/tuffc.2021.3136099

Singla C, Sarangi PK, Sahoo AK, Singh PK (2020) Deep learning enhancement on mammogram images for breast cancer detection. Mater Today Proc 49:3098–3104. https://doi.org/10.1016/j.matpr.2020.10.951

Szegedy C, Ioffe S, Vanhoucke V, Alemi A (2017) Inception-v4, Inception-ResNet and the impact of residual connections on learning. In: Proceedings of the AAAI conference on artificial intelligence, vol 31(1). https://ojs.aaai.org/index.php/AAAI/article/view/11231

Tahmoush D, Samet H (2008) A web database for computer-aided detection and diagnosis of medical images. In: Digital mammography, 9th International Workshop, IWDM 200, pp 265–272. Springer, Berlin. https://doi.org/10.1007/978-3-540-70538-3_37

Tajerian MN, Pesce K, Frangella J, Quiroga E, Boietti B, Chico MJ et al (2021) Artemisia: validation of a deep learning model for automatic breast density categorization. J Med Artif Intell 4:5. https://doi.org/10.21037/jmai-20-43

Tan M, Le Q (2019) EfficientNet: rethinking model scaling for convolutional neural networks. http://proceedings.mlr.press/v97/tan19a/tan19a.pdf

Tangaro S, Bellotti R, De Carlo F, Gargano G, Lattanzio E, Monno P et al (2008) MAGIC-5: an Italian mammographic database of digitised images for research. Tech Rep 113(4):477–485. https://doi.org/10.1007/s11547-008-0282-5

Thiagarajan JD, Ganesan R (2012) Analysis of CT brain images using radial basis function neural network. Defence Sci J 62(4):212–218. https://doi.org/10.14429/dsj.62.1830

Thomaz RL, Carneiro PC, Patrocinio AC (2017) Feature extraction using convolutional neural network for classifying breast density in mammographic images. SPIE Proc 10(1117/12):2254633

Tice JA, Miglioretti DL, Li C-S, Vachon CM, Gard CC, Kerlikowske K (2015) Breast density and benign breast disease: risk assessment to identify women at high risk of breast cancer. J Clin Oncol 33(28):3137–3143. https://doi.org/10.1200/jco.2015.60.8869

Trivizakis E, Ioannidis G, Melissianos V, Papadakis G, Tsatsakis A, Spandidos D, Marias K (2019) A novel deep learning architecture outperforming “off-the-shelf” transfer learning and feature-based methods in the automated assessment of mammographic breast density. Oncol Rep. https://doi.org/10.3892/or.2019.7312

Turay T, Vladimirova T (2022) Towards performing image classification and object detection with convolutional neural networks in autonomous driving systems: a survey (December 2021). IEEE Access. https://doi.org/10.1109/access.2022.3147495

Ulagamuthalvi V, Kulanthaivel G, Balasundaram A, Kumar Sivaraman A (2022) Breast mammogram analysis and classification using deep convolution neural network. Comput Syst Sci Eng 43(1):275–289. https://doi.org/10.32604/csse.2022.023737

Van Putten LD, Bamford KE (2023) Improving systolic blood pressure prediction from remote photoplethysmography using a stacked ensemble regressor. https://openaccess.thecvf.com/content/CVPR2023W/CVPM/html/van_Putten_Improving_Systolic_Blood_Pressure_Prediction_From_Remote_Photoplethysmography_Using_a_CVPRW_2023_paper.html. Accessed 4 July 2023

Vargas-Hakim GA, Mezura-Montes E, Acosta-Mesa H-G (2021) A review on convolutional neural networks encodings for neuroevolution. IEEE Trans Evol Comput. https://doi.org/10.1109/tevc.2021.3088631

Veit A, Wilber M, Belongie S (2016) Residual networks behave like ensembles of relatively shallow networks. In: NIPS’16: proceedings of the 30th international conference on neural information processing systems. https://proceedings.neurips.cc/paper/2016/file/37bc2f75bf1bcfe8450a1a41c200364c-Paper.pdf

Verma P, Dumka A, Bhardwaj A, Kestwal MC (2021) Classifying breast density in mammographic images using wavelet-based and fine-tuned sensory neural networks. Int J Image Graph. https://doi.org/10.1142/s0219467821400040

Wanders JOP, Holland K, Karssemeijer N, Peeters PHM, Veldhuis WB, Mann RM, van Gils CH (2017) The effect of volumetric breast density on the risk of screen-detected and interval breast cancers: a cohort study. Breast Cancer Res. https://doi.org/10.1186/s13058-017-0859-9

Wani NA, Bedi J, Kumar R, Khan MA, Rida I (2024a) Synergizing fusion modelling for accurate cardiac prediction through explainable artificial intelligence. IEEE Trans Consum Electron. https://doi.org/10.1109/TCE.2024.3419814

Wani NA, Kumar R, Bedi J (2024b) DeepXplainer: An interpretable deep learning based approach for lung cancer detection using explainable artificial intelligence. Comput Methods Programs Biomed 243:107879. https://doi.org/10.1016/j.cmpb.2023.107879

Wani NA, Kumar R, Bedi J (2024c) Harnessing fusion modeling for enhanced breast cancer classification through interpretable artificial intelligence and in-depth explanations. Eng Appl Artif Intell 136:108939. https://doi.org/10.1016/j.engappai.2024.108939

Wani NA, Kumar R, Mamta, Bedi J, Rida I (2024d) Explainable AI-driven IoMT fusion: unravelling techniques, opportunities, and challenges with Explainable AI in healthcare. Inf Fusion 110:102472. https://doi.org/10.1016/j.inffus.2024.102472

Wengert GJ, Helbich TH, Kapetas P, Baltzer PA, Pinker K (2018) Density and tailored breast cancer screening: practice and prediction– an overview. Acta Radiol Open 7(9):205846011879121. https://doi.org/10.1177/2058460118791212

Wu N, Geras KJ, Shen Y, Su J, Kim SH, Kim EH et al (2018) Breast density classification with deep convolutional neural networks. In: 2018 IEEE international conference on acoustics, speech and signal processing (ICASSP). https://doi.org/10.1109/icassp.2018.8462671

Xue D, Zhou X, Li C, Yao Y, Rahaman M, Zhang J et al (2020) An application of transfer learning and ensemble learning techniques for cervical histopathology image classification. IEEE Access 8:104603–104618

Yamada D, Ohde S, Kajiura Y, Yagishita K, Nozaki F, Suzuki K et al (2022) Relationship between breast density, breast cancer subtypes, and prognosis. Clin Breast Cancer 22(6):560–566. https://doi.org/10.1016/j.clbc.2022.04.005

Yan L et al (2024) A domain knowledge-based interpretable deep learning system for improving clinical breast ultrasound diagnosis. Commun Med 4(1):1–11. https://doi.org/10.1038/s43856-024-00518-7

Yassen MA, Abdel-Fattah MG, Ismail I, El Kenawy EM (2024) An AI-based system for predicting renewable energy power output using advanced optimization algorithms. J Artif Intell Metaheuristics (JAIM) 08(1):1–8

Yi PH, Lin A, Wei J, Yu AC, Sair HI, Hui FK et al (2019) Deep-learning-based semantic labeling for 2D mammography and comparison of complexity for machine learning tasks. J Digit Imaging 32(4):565–570. https://doi.org/10.1007/s10278-019-00244-w

Yu T, Ye D (2022) The epidemiologic factors associated with breast density: a review. J Res Med Sci 27(1):53–53. https://doi.org/10.4103/jrms.jrms_962_21

Zhang J, Cai L, Pan X, Chen L, Chen M, Yan D et al (2022) Comparison and risk factors analysis of multiple breast cancer screening methods in the evaluation of breast non-mass-like lesions. BMC Med Imaging 22(1):1–11. https://doi.org/10.1186/s12880-022-00921-3

Zhao W, Wang R, Qi Y, Lou M, Wang Y, Yang Y et al (2021) BASCNet: Bilateral adaptive spatial and channel attention network for breast density classification in the mammogram. Biomed Signal Process Control 70:103073. https://doi.org/10.1016/j.bspc.2021.103073