



OPEN ACCESS

EDITED BY

Quan Zou,
University of Electronic Science and
Technology of China, China

REVIEWED BY

Tongjun Gu,
University of Florida, United States
Yunyun Zhou,
Children's Hospital of Philadelphia,
United States

*CORRESPONDENCE

Zhongming Zhao,
✉ zhongming.zhao@uth.tmc.edu

SPECIALTY SECTION

This article was submitted to
Computational Genomics,
a section of the journal
Frontiers in Genetics

RECEIVED 13 November 2022

ACCEPTED 15 December 2022

PUBLISHED 04 January 2023

CITATION

Das HS, Das A, Neog A, Mallik S, Bora K
and Zhao Z (2023), Breast cancer
detection: Shallow convolutional neural
network against deep convolutional
neural networks based approach.
Front. Genet. 13:1097207.
doi: 10.3389/fgene.2022.1097207

COPYRIGHT

© 2023 Das, Das, Neog, Mallik, Bora and
Zhao. This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Breast cancer detection: Shallow convolutional neural network against deep convolutional neural networks based approach

Himanish Shekhar Das¹, Akalpita Das², Anupal Neog³,
Saurav Mallik^{4,5,6}, Kangkana Bora¹ and Zhongming Zhao^{4,7*}

¹Department of Computer Science and Information Technology, Cotton University, Guwahati, India, ²Department of Computer Science and Engineering, GIMT Guwahati, Guwahati, India, ³Department of AI and Machine Learning COE, IQVIA, Bengaluru, Karnataka, India, ⁴Center for Precision Health, School of Biomedical Informatics, The University of Texas Health Science Center at Houston, Houston, TX, United States, ⁵Department of Environmental Health, Harvard T. H. Chan School of Public Health, Boston, MA, United States, ⁶Department of Pharmacology and Toxicology, University of Arizona, Tucson, AZ, United States, ⁷Department of Pathology and Laboratory Medicine, McGovern Medical School, The University of Texas Health Science Center at Houston, Houston, TX, United States

Introduction: Of all the cancers that afflict women, breast cancer (BC) has the second-highest mortality rate, and it is also believed to be the primary cause of the high death rate. Breast cancer is the most common cancer that affects women globally. There are two types of breast tumors: benign (less harmful and unlikely to become breast cancer) and malignant (which are very dangerous and might result in aberrant cells that could result in cancer).

Methods: To find breast abnormalities like masses and micro-calcifications, competent and educated radiologists often examine mammographic images. This study focuses on computer-aided diagnosis to help radiologists make more precise diagnoses of breast cancer. This study aims to compare and examine the performance of the proposed shallow convolutional neural network architecture having different specifications against pre-trained deep convolutional neural network architectures trained on mammography images. Mammogram images are pre-processed in this study's initial attempt to carry out the automatic identification of BC. Thereafter, three different types of shallow convolutional neural networks with representational differences are then fed with the resulting data. In the second method, transfer learning via fine-tuning is used to feed the same collection of images into pre-trained convolutional neural networks VGG19, ResNet50, MobileNet-v2, Inception-v3, Xception, and Inception-ResNet-v2.

Results: In our experiment with two datasets, the accuracy for the CBIS-DDSM and INbreast datasets are 80.4%, 89.2%, and 87.8%, 95.1% respectively.

Discussion: It can be concluded from the experimental findings that the deep network-based approach with precise tuning outperforms all other state-of-the-art techniques in experiments on both datasets.

KEYWORDS

breast cancer, medical imaging, deep learning, convolutional neural networks, transfer learning

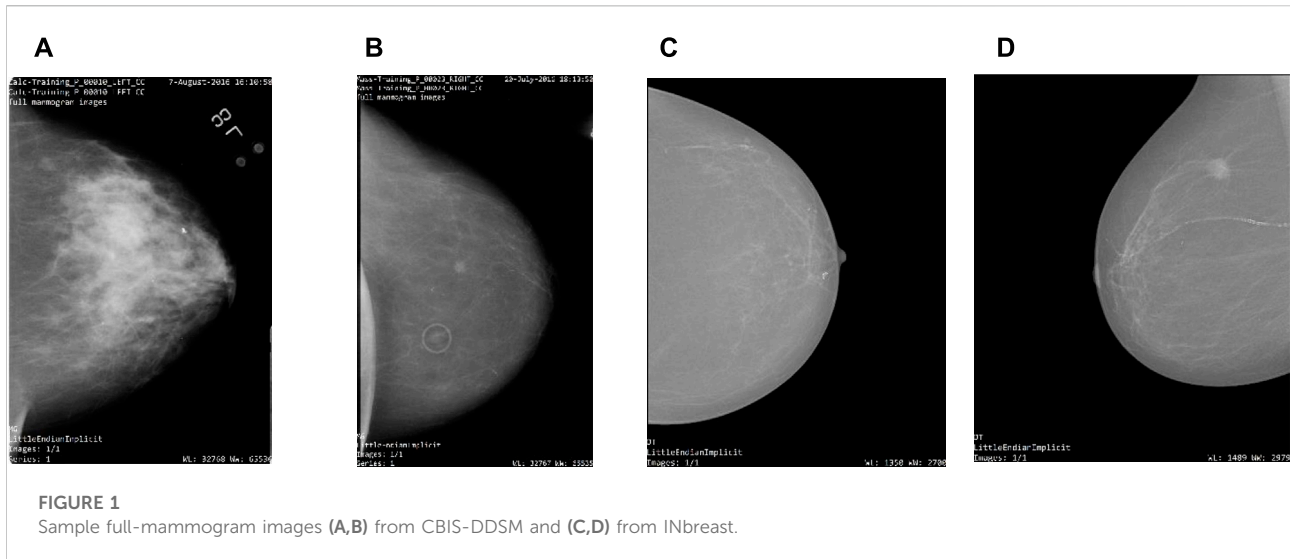
1 Introduction

Cancer affects several body organs, including the lungs, pancreas, blood, and breasts, making it more than simply a typical sickness. These cancer forms are similar in some ways, yet they differ in their modes of development and dissemination. Breast cancer affects women mostly and is thought to be the second-deadliest illness in women, according to the International Agency for Research on Cancer (IARC) study released by the World Health Organization (WHO) in 2012 (Boyle and Levin, 2008). According to the Indian Council of Medical Research (ICMR), 2018 in India, breast cancer accounts for the biggest percentage of malignancies and is the second most prevalent disease overall that affects women, with 87,090 fatalities on average (Mathur et al., 2020). With an anticipated 2.3 million new cases, or 11.7% of all cancer cases, it now overtook lung cancer as the most common kind of cancer worldwide in 2020 (Sung et al., 2021). Cancer develops in the human body as a result of the tumor cells' aberrant development and invasion of the surrounding tissues. Tumors are often divided into benign and malignant categories. In contrast to malignant tumor cells, which are thought to be cancerous, benign tumor cells are not carcinogenic. The benign tumor's cells proliferate only in that area of the body and are unable to invade nearby tissues to spread. On the other hand, malignant tumor cells can develop out of control, infiltrate nearby tissue, and eventually spread to different areas of the body (Prusty et al., 2022). To diagnose breast cancer, a variety of screening techniques are utilized, although mammography is by far the most effective. In mammography, several perspectives, such as the craniocaudal (CC) and medial-lateral oblique (MLO), are utilized to better comprehend the breast abnormalities that are present. Radiologists employ any of these MLO or CC views to examine breast lesion indicators such as masses and micro-calcifications to differentiate between benign, normal, and malignant classifications of breast abnormalities (Bick, 2014). It takes a lot of time and knowledge from a radiologist with extensive training and experience to interpret mammographic images (Hubbard et al., 2011). These problems have led to an increase in the need for computer-aided diagnosis and detection (CAD) technologies, which automate medical image processing (Das et al., 2022).

Deep learning has made significant strides in research over the past 10 years (Das et al., 2021), and the subject of healthcare is no exception (Das et al., 2020). In the research on breast cancer, which produced encouraging findings as well, several deep architectures (Hamidinekoo et al., 2018) are investigated and effectively applied. The use of computer-aided diagnostics to diagnose breast cancer is still being studied by a sizable number of researchers. In a hybrid Convolutional Neural Network (CNN) approach presented by Arevalo et al. (2016), handmade image-based features are learned using supervised learning techniques. Huynh et al. (2016) used transfer learning to the pre-trained AlexNet model for the classification of mammographic tumors without fine-tuning, and the Support Vector Machine (SVM) technique is utilized for classification at the back end. Recent advances in deep learning technology have the potential to improve the standard of treatment in the healthcare sector (Carneiro et al., 2015). Fine-tuned ImageNet, a pre-trained CNN, to discriminate between masses and micro-calcification. The Breast Imaging-Reporting and Data System (BI-RADS) score is particularly helpful for identifying the kind of breast cancer. In their study, the authors used the BI-RADS score to distinguish between various forms of breast cancer. For the classification of mammographic images, Levy and Jain (2016) used transfer learning on two pre-trained models, such as AlexNet and GoogleNet. AlexNet was found to have the greatest results when they compared the two networks. Ting et al. (2019) created a brand-new network named Convolutional Neural Network Improvement for Breast Cancer Classification network (CNNI-BCC). The proposed model was created from scratch and trained for BC classification. The experiment used the MIAS database, and the model was provided with the area of interest (ROIs) identified using a one-shot detector. After training with all CNN, (Rampun et al., 2018) used the three top-performing model predictions in an ensemble and modified version of the AlexNet model. Tsochatzidis et al. (2019) conducted a thorough analysis of several designs, including Inception-BN (v2) from scratch, GoogleNet, ResNet50, ResNet101, ResNet152, AlexNet, VGG16, and VGG19. Both the Digital Database for Screening Mammography (DDSM-400) and Curated Breast Imaging-Digital Database for Screening Mammography (CBIS-DDSM) datasets might use some fine-tuning. Arora et al. (2020) proposed a high

TABLE 1 Mammography datasets for breast cancer.

Dataset	Type	No. of images	View	Format	Classes
CBIS-DDSM	Digital mammogram (DM)	10,239	MLO/CC	DICOM	Benign and malignant
INbreast	Digital mammogram (DM)	410	MLO/CC	DICOM	Benign and malignant



ensemble transfer learning model to differentiate between benign and malignant tumors. A neural network classifier is then used to do auto-feature extraction. In the literature, several deep learning models have been proposed. To identify and classify benign and malignant lesions included in digital mammography images, authors (Duggento et al., 2019) employed the CNN model. Deep Convolution Neural Network and AlexNet were used by (Ragab et al., 2019) for feature extraction and classification of mammographic images from the DDSM dataset. Another deep CNN was used by Ting et al. (2019) to classify BC-lesion from the MIAS dataset. BreastNet is another CNN-based deep learning model that has been put up by authors (Toğaçar et al., 2020) and beats AlexNet, VGG-16, and VGG-19. The authors (Wang et al., 2020) have also developed bilateral residual GANs (BR-GANs), which are based on the cycle GAN idea, for the job of segmenting mammograms for the INbreast dataset. Mohiyuddin et al. (2022) have suggested modified YOLOv5 for the identification and categorization of breast tumors of the DDSM dataset. Al-Antari et al. (2018) are using AlexNet, a second CNN-based classifier, in addition to YOLO for multiclass classification for the INbreast dataset. U-Net is renowned for its effectiveness in the segmentation of medical images. For the segmentation challenge, Zeiser et al. (2020) applied an architecture based on U-Net to a variety of publicly and privately accessible datasets. The segmentation of masses on mammography images is the focus of their work. Authors (Abdelhafiz et al., 2020) have also presented the Vanilla U-Net, another form of U-Net, for segmenting mass in mammography images obtained from three separate publicly accessible datasets. For breast density segmentation, Saffari et al. (2020) suggested a conditional form of the GAN (cGAN) and U-Net-based cGAN-UNET model. Later, CNN was applied for classification purposes for the INbreast dataset. The efficiency of transfer learning for breast cancer categorization was demonstrated by Rahman et al. (2020). For

the DDSM dataset, the authors tested the transfer learning-based models. Similar transfer learning principles were applied by Saber et al. (2021) for the categorization of breast cancer using a variety of pre-trained deep learning models, including Inception V3, ResNet50, Visual Geometry Group networks (VGG)-19, VGG-16, and Inception-V2. Another method for classifying breast cancer using multi-DCNNs was proposed by Ragab et al. (2021), who similarly employed the idea of transfer learning for feature extraction and SVM for final classification on the MIAS-DDSM datasets. Lotter et al.'s annotation-efficient deep learning strategy for mammograms and digital breast tomosynthesis image-based breast cancer diagnosis has been proposed by Lotter et al. (2021). Malebary and Hashmi. (2021) suggested an ensemble-based technique for classifying breast masses. In the proposed study, they have employed, pre-trained CNN and RNN-LSTM-based deep learning models to extract both the low-level and high-level features. Finally, the classification was completed by combining the random forest approach with high gradient boosting. Deep learning techniques based on attention have also proven effective in classifying images. Sun et al. (2020) presented the attention-guided deep learning network known as AU-Net, which segments breast mass using an attention-guided upsampling block. Multi-scale attention-based network MSANet was created by Xu et al. (2021) for the categorization of mammograms for the DDSM

TABLE 2 Distribution of data for full-mammogram images.

Dataset	Type	Train	Test	Total
CBIS-DDSM	Full mammogram	2138	566	2704
INbreast	Full mammogram	328	82	410
Grand total		2466	648	3114

dataset. Su et al. (2022) introduced the YOLO-LOGO segmentation model, which combines the YOLO and LOGO architectures with a deep learning technique based on transformers for the identification and segmentation of breast masses for DDSM-INbreast datasets.

The Curated Breast Imaging Subset of the Digital Database for Screening Mammography (CBIS-DDSM) and INbreast datasets are used to train several convolutional neural network architectures in this paper's effort to construct an automated system employing mammographic images. Investigations are carried out in this study to classify breast cancer into benign and malignant categories. In the first technique, three different types of shallow convolutional neural networks have been trained from scratch; in the second approach, various pre-trained convolutional neural networks have also been tested using transfer learning *via* fine-tuning. Convolutional neural network variations such as VGG19, ResNet50, MobileNet-v2, Inception-v3, Xception, and Inception-ResNet-v2 respectively, have each been taken into consideration for this work. The accuracy of various breast cancer datasets for different algorithms as well as with various proposed strategies is the main focus of prior research studies. However, the major focus of this work is on how various deep-learning models behave for complete mammography images. The following are this paper's key contributions:

- To compare and analyze the performance of shallow convolutional networks against deep convolutional neural networks.
- To analyze the usefulness of various proposed models while working with only full-mammogram images.
- Impact of transfer learning approach on the pre-trained convolutional neural networks using fine-tuning and regularization techniques.

The organization of the paper is as follows: Section 2 explains the database along with the proposed mechanism. Section 3 explains the experimental results of the proposed approaches, followed by the discussion in Section 4. Finally, the conclusion of the paper is presented in Section 5.

2 Materials and methods

This section explains the databases that are used for this study. It also briefly describes the proposed model along with the comparative approach of the shallow network against deep convolutional pre-trained networks with their set parameters.

2.1 Database description

2.1.1 CBIS-DDSM

It is a section of the DDSM (Lee et al., 2016) database, which contains 6775 studies altogether. The qualified and experienced

radiologists who choose the mammographic images from DDSM represent them in an updated and standardized version of CBIS-DDSM. After lossless decompression, all of the images are transferred to the DICOM format. The segmented region of interest (ROI) for training data is also included in the database along with information on pathologic diagnosis. While the information may be generally categorized based on the sorts of anomalies, such as bulk and calcification, it can also be divided into subclasses based on malignant and benign tumors.

2.1.2 INbreast

The INbreast dataset contains 410 different digital mammographic images from 115 patients. The mammography images were interpreted by experienced radiologists, and following the analysis, the lesions identified by the scans were given a standard score known as BI-RADS (Orel et al., 1999). The six BI-RADS scores represent the various stages of abnormalities that can be found in the breasts; score 0 denotes an inconclusive examination; score 1 suggests no findings; score 2 indicates benign; score 3 denotes probably benign findings; score 4 denotes suspicious findings; score 5 ensures a high probability of malignancy, and score 6 denotes breast cancer. The INbreast dataset is not accessible to the general public, although it may be requested from (INbreast Dataset, 2012).

The comprehensive description of both datasets can be found in Table 1. Table 2 gives insights into the distribution of data for full-mammogram images that have been taken into consideration for the proposed work. Figure 1 shows examples of some sample images from both datasets.

2.2 Proposed methodology

Deep learning architectures' main contribution is their capacity to autonomously extract low-level to high-level properties (Das and Roy, 2021). CNNs are the best models for identifying detailed properties in images. CNN may learn feature representation automatically as opposed to manually constructed features. Various requirements have been the subject of an extensive investigation in this research. All mammogram images first went through a pre-processing stage in which they were converted from DICOM images to the portable network graphics (PNG) format. The DICOM format is being converted to PNG to prevent loss in image quality. After that, the pixel values between "0" and "1" are normalized to ensure that the higher pixel values have no impact on the investigation's findings. Thereafter, we have to adjust the labels for the binary classification problem, so that "0" corresponds to "benign" and "1" maps to "malignant." Then the mammographic images from both datasets are resized to 224 × 224 image size; split the training data are into "training" and "validation" subsets; build Keras generators for training and validation data. The paper consists of two different approaches. A very small CNN with only two

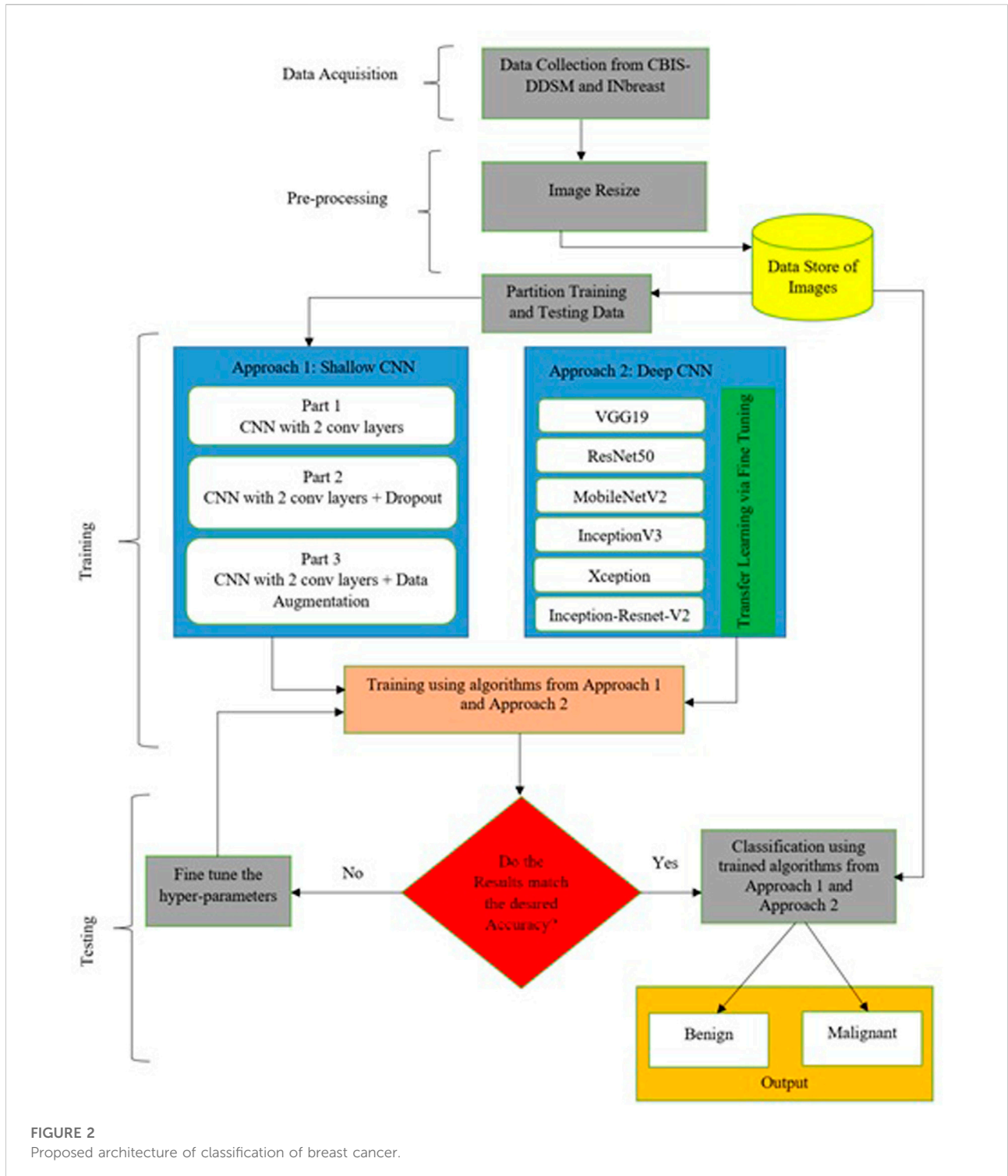


FIGURE 2 Proposed architecture of classification of breast cancer.

convolutional layers has been used for simulation in the first part of the first approach. A dropout layer has been added in the second part to mitigate hard overfitting, and in the third part, data augmentation, another regularization technique, is applied to the previous model to further mitigate overfitting. The second

method involves testing a variety of pre-trained CNN networks using a fine-tuning strategy, including VGG19, ResNet50, MobileNet-v2, Inception-v3, Xception, and Inception-ResNet-v2 respectively. Figure 2 provides a concise summary of the suggested methods.

2.3 Classification model

Finding the approximate size of the model is the first stage. A model with too many parameters may learn slowly and overfit, whereas a network that is too tiny will not be able to generalize well. Starting with a tiny, naïve model and progressively increasing its size until it begins overfitting while learning is a useful technique to determine the right size. At that point, the model is adaptable enough to suit the training data and may be generalizable to additional data with the right training. Of course, the model may be improved later on to produce better performances by adding new layers, changing the existing ones using regularization techniques, or adjusting the hyperparameters. As mentioned, the study consists of two different approaches. Two convolutional layers interspersed with max-pooling make up a very tiny CNN used in the first part of the first approach. The output is produced at the very end by a single neuron with sigmoid activation following a completely connected layer (binary classification). The loss function is the binary cross-entropy, which is particularly suitable for this kind of problem (binary classification). The optimizer is RMSprop, an adaptive optimization algorithm that is considered quite efficient. Since the validation set is where the model often performs at its best, we keep an eye on how the loss changes there as well throughout training and store the relevant model weights when that loss is minimal. The top objective right now is to lessen the harsh overfitting that was shown in the previous part since it keeps the network from discovering a useful broad set of weights. The second part is identical to the first, with the exception that there is a dropout layer added after the last fully-connected block. Dropout is a potent regularization method that's frequently used to cut down on overfitting. Every iteration during the training phase involves disregarding a randomly selected portion of the neurons from the preceding layer; this drives the network to identify redundant and alternative representations for the inputs, ultimately improving performance. The dropout rate is set at .5, which is experimentally found to be the ideal number. Data augmentation is used to further regularize the prior model before modifying the architectural arrangement to reduce overfitting. When using data augmentation, specified changes are applied to the original data to produce a greater variety of samples for the network to train on. Image flipping, shifting, rotation, scaling, distortion, and noise injection are common approaches. The Keras image generator has to be reinitialized with the appropriate inputs in order to take use of data augmentation. We apply the augmentation techniques using the following settings: Flipping (horizontal and vertical), Rotation (angle between 0° and 180°), Shear (10 deg), and Scale (.2).

2.3.1 VGG19

VGG19: At the Oxford Robotics Institute, the Visual Geometry Group Network (VGG) was created based on the convolutional neural network architecture (Hemdan et al., 2020),

and (Simonyan and Zisserman, 2014) unveiled it. The ImageNet data cluster has shown particularly strong performance from VGGNet. To calibrate 138 million weight parameters, this network underwent more than 370,000 iterations of training on more than 1 million pictures in 1000 classes. Particularly, in the Large-Scale Visual Recognition Challenge, a worldwide image recognition competition, in 2014 (Simonyan and Zisserman, 2014; Sukegawa et al., 2020), VGG19 took first place in a classification and localization competition. VGG11, VGG13, VGG16, and VGG19 are some of the transfer learning networks that make up the VGG. These network architectures all share the trait of having several convolution-layer modules coupled to three complete connection layers (Wan et al., 2021). VGG19 is composed of five components. There are two convolutional layers and one pooling layer in the first and second building blocks, respectively. The third and fourth blocks each contain one pooling layer and four convolutional layers. There are four convolutional layers in the final block. There are also 3×3 filters (Arshad et al., 2022).

2.3.2 ResNet50

The most crucial component of ResNet-50 is the residual building block (RBB). The foundation of RBB is the concept of employing shortcut connections to skip whole convolutional layer blocks. In order to avoid the vanishing/exploding gradients problem, these shortcuts help optimize trainable parameters in error backpropagation, which can help to build deeper CNN structures to enhance overall performance for fault detection. Convolutional layers, batch normalizations, the Rectified Linear Unit (ReLU) activation function, and one shortcut make up the RBB (Wen et al., 2020).

2.3.3 MobileNetV2

A widely used CNN-based model for classifying images is called MobileNetV2. The key benefit of adopting the MobileNet architecture is that the model requires far less computing work than the traditional CNN model, making it appropriate for use with mobile devices and personal computers with limited processing power. The MobileNet model is a convolution layer-based simplified structure that can be used to distinguish between the finer details that depend on two controllable characteristics that switch between the parameter's accuracy and latency efficiently. The MobileNet approach has the benefit of shrinking the network (Srinivasu et al., 2021).

2.3.4 InceptionV3

It takes at least a few days to train the deep neural network model Inception-v3 (Szegedy et al., 2016), which is highly challenging for us to train directly on a low-configured machine. We can use transfer learning to retrain Inception's final Layer for new categories using lessons from Tensorflow (Abadi et al., 2016). We employ the Inception-v3 (Szegedy et al., 2016) model's last layer, which is removed while keeping the

parameters from the layer before it. The last layer is then retrained. In the final layer, there are as many output nodes as there are categories in the dataset. For instance, the final layer in the original Inception-v3 model contains 1,000 output nodes since the ImageNet dataset has 1,000 classes. During the fine-tuning stage, the model has been tuned as per the number of classes as well as the number of tuned layers.

2.3.5 Xception

Numerous structured models utilizing CNN have been developed as a result of the widespread use of CNNs in computer vision. LeNet-style models (LeCun et al., 1995) were first introduced in 1995, after which several models were developed for use in classification and recognition issues. Inception is one such example. Szegedy et al. developed the Inception architecture, also known as Inception-v1 (Szegedy et al., 2015), in 2014. Inception-v2, Inception-v3 (Szegedy et al., 2016), and Inception-ResNet (Szegedy et al., 2017) were later updated. An explanation of the Inception modules may be found in the Xception network (Chollet, 2017) that was employed in this investigation. “Extreme inception” is also where the word Xception originates. A brief overview of Inception will be conducted initially to better understand the Xception design. The item that has to be recognized in object recognition or picture classification may appear little or huge depending on the image. To put it another way, the size of the item may vary across images. It might be challenging to choose the right filter size for the convolution process due to the different object sizes. For an object that seems enormous in the images, a high filter size should be recommended, while a small filter size should be used for little things. The problems caused by objects of different sizes can be resolved, according to the Inception design, by placing several filters of varying sizes at the input. Additionally, it recommends forwarding this module’s output to yet another inception module.

2.3.6 Inception-ResNet-V2

Szegedy’s Inception-Resnet (Szegedy et al., 2017) design combines the Resnet and Inception network backbone systems. The Inception module is a network with good local topology, which enables it to conduct simultaneous convolution or pooling operations on the input image. It does not limit itself to a single convolution kernel, but rather employs all convolution kernels of various sizes simultaneously, merging the output of each convolution to create a more detailed feature map. Benefiting from it may result in improved visual representation (Wu et al., 2017). Kaiming He, submitted Resnet (He et al., 2016), a residual neural network architecture of 152 layers, in the ImageNet competition. The neural network’s shortcut architecture was introduced by him.

2.4 Transfer learning

Typically, a large dataset is needed to train the CNN (at least thousands of samples if not available in millions). Due to the restricted time and effort of professionals to provide labeled sample datasets on medical images, it is challenging to use CNN trained from scratch. Methods based on the CNN typically overfit and are unable to extract the image features in good quality when the training dataset is short, as is the situation in this area of medical image analysis. Transfer learning is a technique that allows a CNN to be initially trained on a large-scale labeled image dataset to learn standardized image properties before being used to obtain comparable features from a smaller dataset. It has already been effectively used in many image-processing applications and clinical studies for diseases. For all assessments and comparisons, Table 3 shows the fine tweaking done with the pre-trained models.

3 Results and discussion

To detect BC from mammography images, the authors used two distinct methods. The comparison and analysis of the performance of shallow CNN and deep CNN models that have already been trained is the main objective. On the CBIS-DDSM and INbreast datasets, we put the suggested ways into practice and evaluated them to determine how well they performed in comparison to the existing methodologies. These models were produced using the Tensor Flow backend and Keras deep learning framework provided by Google Colab.

3.1 Performance metrics for evaluation of classification task

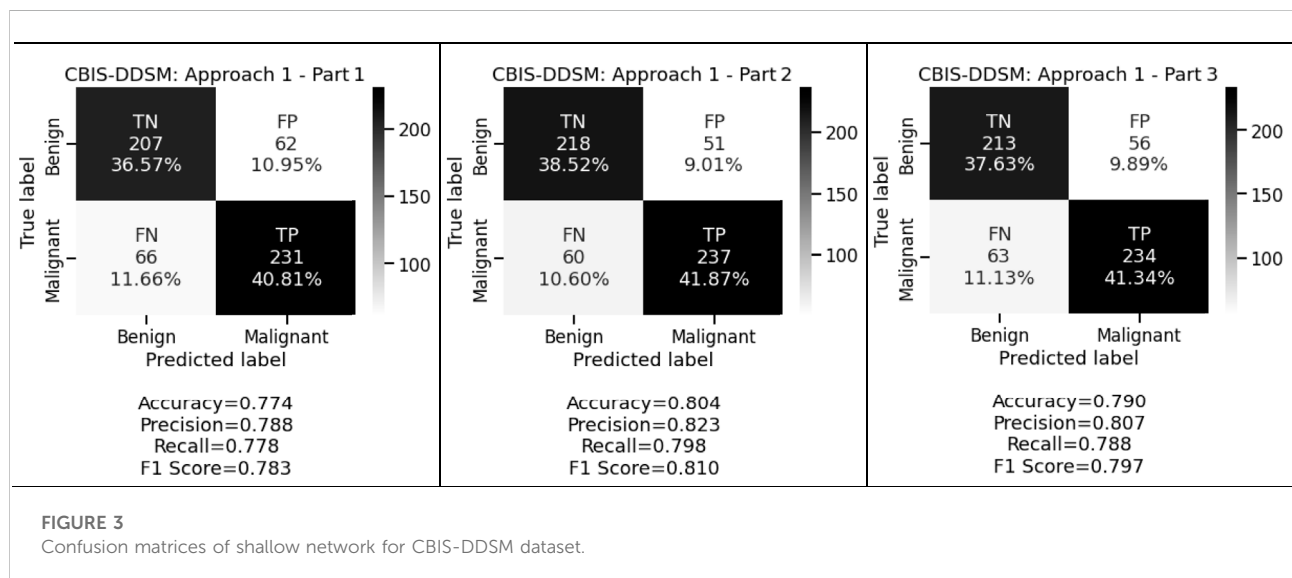
A classification model’s performance assessment parameters are based on the model’s accurate and wrong predictions of test records. For the test dataset for all classes, the confusion matrix provides information on how predicted values compare to real values that may be seen. The confusion matrix includes the four measurements true positive (TP), false positive (FP), true negative (TN), and false negative (FN). These four metrics can be used to assess effective parameters for comparing various categorization systems. Table 4 explains the most used performance metrics based on the confusion matrix. When assessing the effectiveness of our proposed approaches in this study, accuracy, precision, recall, and F1 score were taken into account.

TABLE 3 Pre-trained CNNs with fine-tuned parameters.

Models	Number of layers added	Number of nodes added	Dropout rate (%)
VGG19 Simonyan and Zisserman, (2014)	3	2048, 1024, 1024	0.5%
ResNet50 Akiba et al. (2017)	2	1024, 512	0.5%
MobileNet-v2 Howard et al. (2017)	2	2048, 1024	0.5%
Inception-v3 Xia et al. (2017)	1	1024	0.6%
Xception Chollet., 2017	2	1024, 512	0.5%
Inception-ResNet-v2 Szegedy et al. (2017)	1	1024	0.7%

TABLE 4 Performance measures for the evaluation.

Measures	Formula	Description
Accuracy Urban et al. (2018), Zhang et al. (2016), Bandyopadhyay et al. (2013), Bedrikovetski et al. (2021)	$\frac{ TP + TN }{ TP + TN + FP + FN }$	The ratio of the number of correct predictions with respect to total observations
Precision (Urban et al. (2018), Zhang et al. (2016), Bandyopadhyay et al. (2013), Bedrikovetski et al. (2021)	$\frac{ TP }{ TP + FP }$	The ratio of the number of correct positive predictions with respect to total positive predictions
Recall/Sensitivity (Urban et al. (2018), Zhang et al. (2016), Bandyopadhyay et al. (2013), Bedrikovetski et al. (2021)	$\frac{ TP }{ TP + FN }$	The ratio of the number of correct positive predictions with respect to actual positive observations
F1 score/Dice-coefficient (Urban et al. (2018), Zhang et al. (2016), Bandyopadhyay et al. (2013), Bedrikovetski et al. (2021)	$2 \times \frac{\text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}}$	F1 score is the harmonic mean of both precision and recall



3.2 Evaluation of classifier performance

Figure 3 and Figure 4 display the performance of the shallow convolutional neural network in various scenarios on the CBIS-DDSM and INbreast datasets. For breast cancer detection, our proposed comparison reflects that experiment

on the shallow network in the first aspect where only a two convolution layer-based model is proposed which obtained 77.4%, 78.8%, and 77.8% accuracy, precision, and recall for the CBIS-DDSM dataset, and 84.1%, 86.4%, and 84.4% accuracy, precision, and recall for INbreast dataset respectively; for the second aspect where to reduce the impact of overfitting,

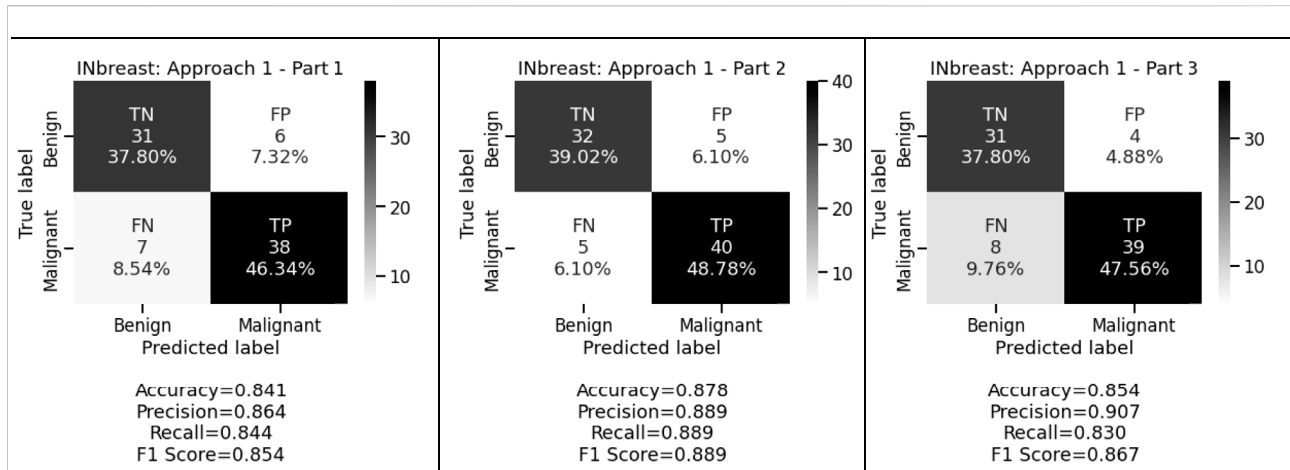


FIGURE 4
Confusion matrices of shallow network for INbreast dataset.

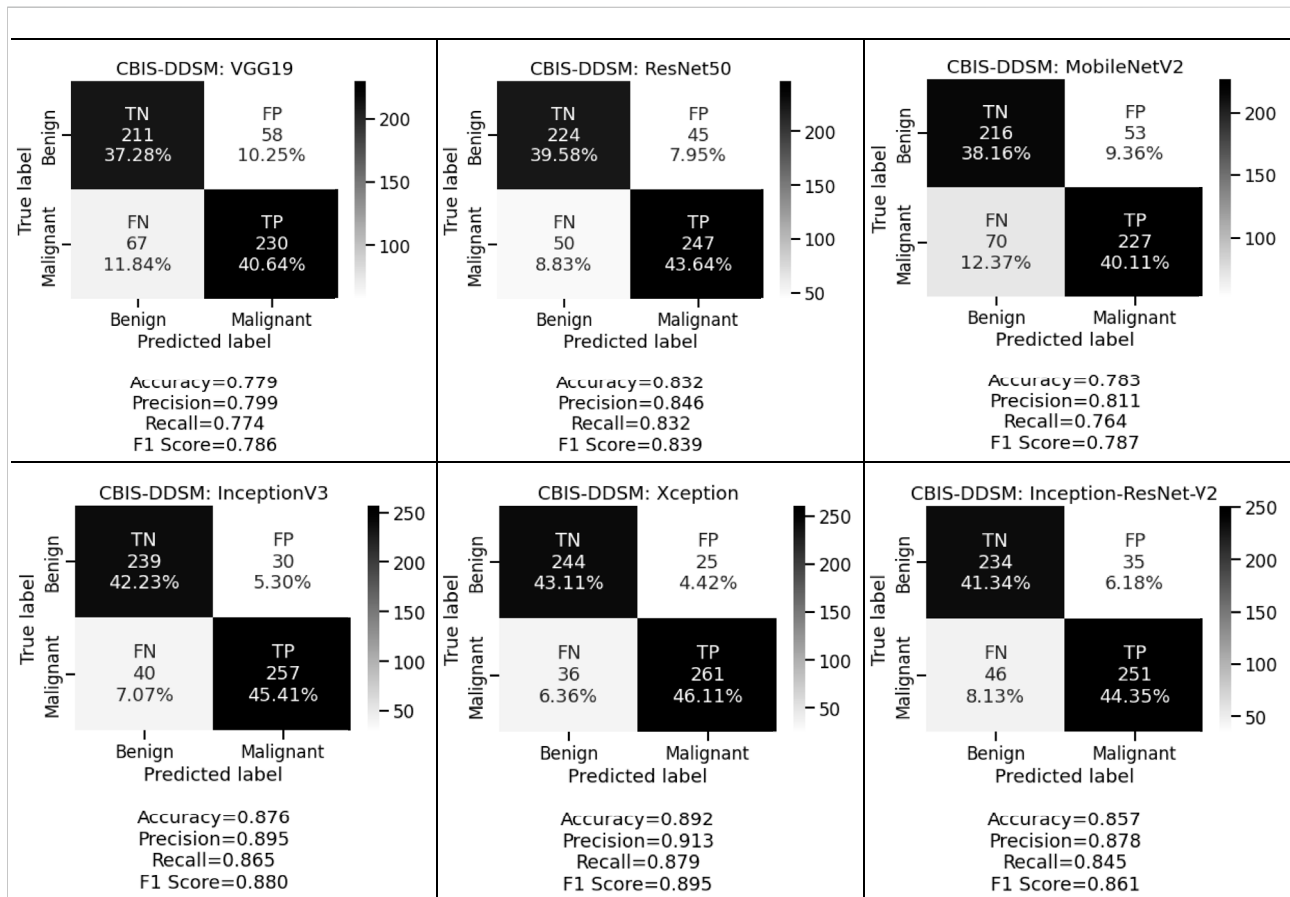


FIGURE 5
Confusion matrices of pre-trained CNNs for the CBIS-DDSM dataset.

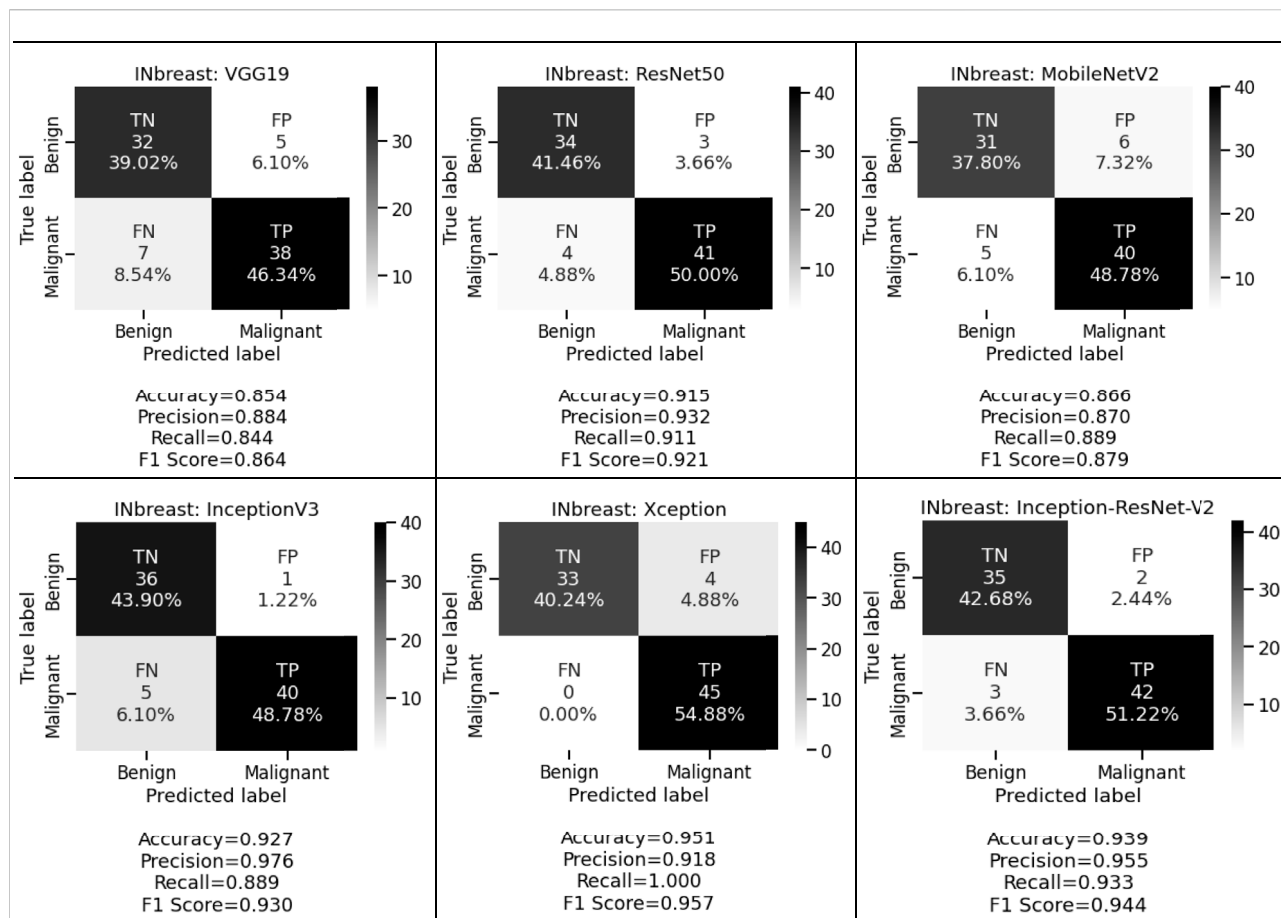


FIGURE 6
Confusion matrices of pre-trained CNNs for INbreast dataset.

dropout strategy has been taken into account has obtained 80.4%, 82.3%, and 79.8% accuracy, precision, and recall for CBIS-DDSM dataset, and 87.8%, 88.9%, and 88.9% accuracy, precision, and recall for INbreast dataset respectively; and for the third aspect where again to reduce the effect of overfitting due to less number of images, augmentation technique has been applied which resulted into 79.0%, 80.7%, and 78.8% accuracy, precision, and recall for CBIS-DDSM dataset, and 85.4%, 90.7%, and 83.0% accuracy, precision, and recall for INbreast dataset respectively.

Figure 5 and Figure 6 display the performance of the pre-trained deep convolutional neural networks on the CBIS-DDSM and INbreast datasets. We have also analyzed the other aspect where pre-trained deep convolutional neural networks have been executed with the help of transfer learning *via* fine-tuning on VGG19, ResNet50, MobileNetV2, InceptionV3, Xception, and Inception-ResNet-V2 respectively. From the above-mentioned experimental analysis on the CBIS-DDSM dataset, we obtained 77.9%, 79.9%, and 77.4% accuracy, precision, and recall for VGG19; 83.2%, 84.6%, and 83.2% accuracy, precision, and

recall for ResNet50; 78.3%, 81.1%, and 76.4% accuracy, precision, and recall for MobileNetV2; 87.6%, 89.5%, and 86.5% accuracy, precision, and recall for InceptionV3; 89.2%, 91.3%, and 87.9% accuracy, precision, and recall for Xception, and 85.7%, 87.8%, and 84.5% accuracy, precision, and recall for Inception-ResNet-V2 respectively.

When the same fine-tuning-based experiment is performed on the INbreast dataset, we obtained 85.4%, 88.4%, and 84.4% accuracy, precision, and recall for VGG19; 91.5%, 93.2%, and 91.1% accuracy, precision, and recall for ResNet50; 86.6%, 87.0%, and 88.9% accuracy, precision, and recall for MobileNetV2; 92.7%, 97.6%, and 88.9% accuracy, precision, and recall for InceptionV3; 95.1%, 91.8%, and 100.0% accuracy, precision, and recall for Xception, and 93.9%, 95.5%, and 93.3% accuracy, precision, and recall for Inception-ResNet-V2 respectively. These observed results indicated that the proposed deep CNN with fine-tuning-based approach performed better than the shallow classifier in various conditions for both classification tasks.

TABLE 5 Comparison of the proposed approach with the state-of-the-art techniques.

Dataset	Author	Year	Accuracy (%)
CBIS-DDSM	Shams et al. (2018)	2018	75
CBIS-DDSM	Tsochatzidis et al. (2019)	2019	89
CBIS-DDSM	Falconi et al. (2020)	2020	84.4
CBIS-DDSM	Ansar et al. (2020)	2020	74.5
CBIS-DDSM	Zhang et al. (2020)	2020	87.05
CBIS-DDSM	Proposed (Xception classifier with fine-tuning)	2022	89.2
INbreast	Dhungel et al. (2017)	2017	90
INbreast	Carneiro et al. (2017)	2017	90
INbreast	Shi et al. (2019)	2019	83.9
INbreast	Zhang et al. (2020)	2020	87.93
INbreast	El Houby and Yassin. (2021)	2021	93.04
INbreast	Proposed (Xception classifier with fine-tuning)	2022	95.1

Bold values are the significant results achieved for the proposed work.

We want to keep both the FP and FN low since they are essential to achieving the work's purpose. First, our suggested system warns that patients with malignant cases who are mistakenly classified as benign instances may result in misclassification marked as a false negative (FN). Another possibility is that people who are diagnosed with cancer but do not have it might result in a false positive (FP). Both FNs and FPs have a large effect on misclassification, which results in incorrect diagnosis and poses health risks to people. We prioritized FP and FN equally by taking into account the F1 score and other performance assessment metrics. We observed that in the first approach we obtained 78.3%, 81.0%, and 79.7% for CBIS-DDSM and 85.4%, 88.9%, and 86.7% for the INbreast dataset respectively. Whereas in the second approach for the CBIS-DDSM dataset, we obtained 78.6%, 83.9%, 78.7%, 88.0%, 89.5% and 86.1% for VGG19, ResNet50, MobileNetV2, InceptionV3, Xception, and Inception-ResNet-V2 respectively. For the INbreast dataset, we obtained 86.4%, 92.1%, 87.9%, 93.0%, 95.7% and 94.4% for VGG19, ResNet50, MobileNetV2, InceptionV3, Xception, and Inception-ResNet-V2 respectively. Our proposed fine-tuning method for the Xception model gives the highest F1 score of 89.5% and 95.7% for breast cancer detection. Experimental results in terms of precision, recall, and F1 score conveys that the proposed model has a negligible rate of misclassification.

3.3 Discussion

According to the findings, the fine-tuned Xception classifier is a reliable model for the detection of breast

cancer. The difference in the performances amongst the models because of the selected hyperparameters for each model which are learning rate, number of layers added after the last max-pooling layer of each model, number of hidden units for each layer added in fully-connected layers, and dropout rate respectively. The fine-tuned parameters are described in [Section 2.2](#), [Section 2.3](#), and [Table 3](#) for reference. The main element influencing the outcome of the performances for the CBIS-DDSM and INbreast datasets may be due to the sharp fall in dimensions. The INbreast dataset has higher-quality mammograms as compared to the CBIS-DDSM dataset. The produced feature's ability to discriminate is destroyed by dimensions that are too small. When deep-learning-based features are applied, the overfitting issue also appears in the INbreast dataset. Despite the lack of training samples in the INbreast dataset, the overfitting issue is not as severe as it is in the CBIS-DDSM dataset. This may be mostly due to the excellent mammograms in the INbreast dataset. Along with that because of the fine-tuned parameters, all models performed much better for the INbreast dataset even with lesser sample size. For this work, rigorous experiments have been carried out, such as when the models were tested on the INbreast dataset after being trained on the CBIS-DDSM dataset and *vice versa*. Obtained results were not promising compared to the results obtained when trained and tested on the same dataset, but it will lead to the future scope of research. Though the suggested strategy performs better, this study does not take into account other clinical data, such as other medical problems, geographic location, etc. Future research that takes these parameters into account might enhance

computer-aided methods for early breast cancer diagnosis and tailored medicine treatment. The transfer-learned features are more reflective of the natural image properties and may not always represent the delicate qualities of medical images since the vast dataset used for transfer learning incorporates natural images. As a result, it is anticipated that transfer learning from a big dataset in the same domain would result in the creation of a breast cancer detection system that is more effective. The proposed models have been put through a performance comparison with cutting-edge methods. Table 5 compares the performance of our suggested classifiers on the CBIS-DDSM dataset as well as the INbreast dataset. These outcomes demonstrate that our suggested fine-tuned Xception classifier outperforms other classifiers.

4 Conclusion

In this article, we compared shallow convolutional neural networks to deep convolutional neural networks and introduced pre-trained models that have been fine-tuned to classify full-mammogram images as benign or malignant. ROI-based mammography images can produce better outcomes, but grinding up a sufficient degree of precision for full-mammogram images is a laborious process. Our suggested trained CNNs can pick up various information included in individual images. The improved CNN models produce a more effective image categorization method than the individual CNNs that are trained from scratch. Due to Xception's optimization of ResNet, which allows it to inherit not just ResNet's benefit of residual connection but also its capability to extract objects when occluded by occlusions using depth-wise separable convolution, it obtains the greatest performance when extracting features. When compared to the other classifiers, the comparison analysis shows that the performance of the improved Xception classifier is quite significant. As a result, among all of our suggested methods, the improved Xception classifier performs the best at detecting breast cancer, with acceptable levels for all performance metrics in the range of .87–.91 for the CBIS-DDSM dataset and .91 to 1.00 for the INbreast dataset respectively.

References

- Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., et al. (2016). Tensorflow: Large-scale machine learning on heterogeneous distributed systems. arXiv preprint arXiv:1603.04467.
- Abdelhafiz, D., Bi, J., Ammar, R., Yang, C., and Nabavi, S. (2020). Convolutional neural network for automated mass segmentation in mammography. *BMC Bioinforma.* 21 (1), 192–219. doi:10.1186/s12859-020-3521-y
- Akiba, T., Suzuki, S., and Fukuda, K. (2017). Extremely large minibatch sgd: Training resnet-50 on imagenet in 15 minutes. arXiv preprint arXiv:1711.04325.
- Al-Antari, M. A., Al-Masni, M. A., Choi, M. T., Han, S. M., and Kim, T. S. (2018). A fully integrated computer-aided diagnosis system for digital X-ray mammograms via deep learning detection, segmentation, and classification. *Int. J. Med. Inf.* 117, 44–54. doi:10.1016/j.ijmedinf.2018.06.003
- Ansar, W., Shahid, A. R., Raza, B., and Dar, A. H. (2020). "Breast cancer detection and localization using mobilenet based transfer learning for mammograms," in International symposium on intelligent computing systems, Sharjah, United Arab Emirates, March 18 - 19, 2020, 11–21.
- Arevalo, J., González, F. A., Ramos-Pollán, R., Oliveira, J. L., and Lopez, M. A. G. (2016). Representation learning for mammography mass lesion classification with

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: CBIS-DDSM: <https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=22516629> INbreast: <https://www.kaggle.com/datasets/ramanathansp20/inbreast-dataset>.

Author contributions

HD outlined the study design, performed implementation and data collection, contributed to the data analysis, participated in result interpretation and manuscript writing. HD, AD, AN, SM, KB, and ZZ wrote and edited the manuscript. All authors read and approved the final manuscript.

Funding

ZZ was partially supported by the Cancer Prevention and Research Institute of Texas (CPRIT RP180734) and the Precision Health Chair Professorship fund. The funder had no role in the study design, data collection and analysis, decision to publish, or preparation of the entire manuscript.

Conflict of interest

Author AN was employed by IQVIA.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

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

- convolutional neural networks. *Comput. methods programs Biomed.* 127, 248–257. doi:10.1016/j.cmpb.2015.12.014
- Arora, R., Rai, P. K., and Raman, B. (2020). Deep feature-based automatic classification of mammograms. *Med. Biol. Eng. Comput.* 58 (6), 1199–1211. doi:10.1007/s11517-020-02150-8
- Arshad, H., Khan, M. A., Sharif, M. I., Yasmin, M., Tavares, J. M. R., Zhang, Y. D., et al. (2022). A multilevel paradigm for deep convolutional neural network features selection with an application to human gait recognition. *Expert Syst.* 39 (7), e12541. doi:10.1111/exsy.12541
- Bandopadhyay, S., Mallik, S., and Mukhopadhyay, A. (2013). A survey and comparative study of statistical tests for identifying differential expression from microarray data. *IEEE/ACM Trans. Comput. Biol. Bioinforma.* 11 (1), 95–115. doi:10.1109/TCBB.2013.147
- Bedrikovetski, S., Dudi-Venkata, N. N., Maicas, G., Kroon, H. M., Seow, W., Carneiro, G., et al. (2021). Artificial intelligence for the diagnosis of lymph node metastases in patients with abdominopelvic malignancy: A systematic review and meta-analysis. *Artif. Intell. Med.* 113, 102022. doi:10.1016/j.artmed.2021.102022
- Bick, U. (2014). “Mammography: How to interpret microcalcifications,” in *Diseases of the abdomen and pelvis 2014–2017* (Milano: Springer), 313–318.
- Boyle, P., and Levin, B. (2008). *World cancer report 2008*. Lyon, France: IARC Press, International Agency for Research on Cancer.
- Carneiro, G., Nascimento, J., and Bradley, A. P. (2017). Automated analysis of unregistered multi-view mammograms with deep learning. *IEEE Trans. Med. Imaging* 36 (11), 2355–2365. doi:10.1109/TMI.2017.2751523
- Carneiro, G., Nascimento, J., and Bradley, A. P. (2015). “Unregistered multiview mammogram analysis with pre-trained deep learning models,” in International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, October 5th to 9th, 2015, 652–660.
- Chollet, F. (2017). “Xception: Deep learning with depthwise separable convolutions,” in Proceedings of the IEEE conference on computer vision and pattern recognition, Honolulu, Hawaii, USA, Held 21–26 July 2017, 1251–1258.
- Das, A., Das, H. S., Choudhury, A., Neog, A., and Mazumdar, S. (2020). “Detection of Parkinson’s disease from hand-drawn images using deep transfer learning,” in *Congress on intelligent systems* (Singapore: Springer), 67–84.
- Das, A., Das, H. S., Neog, A., Bharat Reddy, B., and Swargiary, M. (2021). “Performance analysis of different machine learning classifiers in detection of Parkinson’s disease from hand-drawn images using histogram of oriented gradients,” in *Applications of artificial intelligence in engineering* (Singapore: Springer), 205–215.
- Das, H. S., Das, A., Neog, A., Mallik, S., Bora, K., and Zhao, Z. (2022). Early detection of Parkinson’s disease using fusion of discrete wavelet transformation and histograms of oriented gradients. *Mathematics* 10 (22), 4218. doi:10.3390/math10224218
- Das, H. S., and Roy, P. (2021). A CNN-BiLSTM based hybrid model for Indian language identification. *Appl. Acoust.* 182, 108274. doi:10.1016/j.apacoust.2021.108274
- Dhungel, N., Carneiro, G., and Bradley, A. P. (2017). A deep learning approach for the analysis of masses in mammograms with minimal user intervention. *Med. Image Anal.* 37, 114–128. doi:10.1016/j.media.2017.01.009
- Duggento, A., Aiello, M., Cavaliere, C., Cascella, G. L., Cascella, D., Conte, G., et al. (2019). A random initialization deep neural network for discriminating malignant breast cancer lesions. *Contrast media & Mol. imaging* 2019, 912–915. doi:10.1109/EMBC.2019.8856740
- El Houby, E. M., and Yassin, N. I. (2021). Malignant and nonmalignant classification of breast lesions in mammograms using convolutional neural networks. *Biomed. Signal Process. Control* 70, 102954. doi:10.1016/j.bspc.2021.102954
- Falconi, L. G., Perez, M., Aguilar, W. G., and Conci, A. (2020). Transfer learning and fine tuning in breast mammogram abnormalities classification on CBIS-DDSM database. *Adv. Sci. Technol. Eng. Syst. J.* 5, 154–165. doi:10.25046/aj050220
- Hamidinekoo, A., Denton, E., Rampun, A., Honnor, K., and Zwiggelaar, R. (2018). Deep learning in mammography and breast histology, an overview and future trends. *Med. Image Anal.* 47, 45–67. doi:10.1016/j.media.2018.03.006
- He, K., Zhang, X., Ren, S., and Sun, J. (2016). “Deep residual learning for image recognition,” in Proceedings of the IEEE conference on computer vision and pattern recognition, Las Vegas, NV, USA, June 27 2016 to June 30 2016, 770–778.
- Hemdan, E. E. D., Shouman, M. A., and Karar, M. E. (2020). Covidx-net: A framework of deep learning classifiers to diagnose Covid-19 in x-ray images. arXiv preprint arXiv:2003.11055.
- Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., et al. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861.
- Hubbard, R. A., Kerlikowske, K., Flowers, C. I., Yankaskas, B. C., Zhu, W., and Miglioretti, D. L. (2011). Cumulative probability of false-positive recall or biopsy recommendation after 10 years of screening mammography: A cohort study. *Ann. Intern. Med.* 155 (8), 481–492. doi:10.7326/0003-4819-155-8-2011110180-00004
- Huynh, B. Q., Li, H., and Giger, M. L. (2016). Digital mammographic tumor classification using transfer learning from deep convolutional neural networks. *J. Med. Imaging* 3 (3), 034501. doi:10.1117/1.JMI.3.3.034501
- INbreast Dataset (2012). The INbreast dataset. Available at: <https://medicalresearch.inesporto.pt/breastresearch/index.php/Get-INbreast-Database>.
- LeCun, Y., Jackel, L. D., Bottou, L., Cortes, C., Denker, J. S., Drucker, H., et al. (1995). Learning algorithms for classification: A comparison on handwritten digit recognition. *Neural Netw. Stat. Mech. perspective* 261 (276), 2.
- Lee, R. S., Gimenez, F., Hoogi, A., and Rubin, D. (2016). Curated breast imaging subset of DDSM. *cancer imaging archive* 8, 2016. doi:10.1038/sdata.2017.177
- Lévy, D., and Jain, A. (2016). Breast mass classification from mammograms using deep convolutional neural networks. arXiv preprint arXiv:1612.00542.
- Lotter, W., Diab, A. R., Haslam, B., Kim, J. G., Grisot, G., Wu, E., et al. (2021). Robust breast cancer detection in mammography and digital breast tomosynthesis using an annotation-efficient deep learning approach. *Nat. Med.* 27 (2), 244–249. doi:10.1038/s41591-020-01174-9
- Malebary, S. J., and Hashmi, A. (2021). Automated breast mass classification system using deep learning and ensemble learning in digital mammogram. *IEEE Access* 9, 55312–55328. doi:10.1109/access.2021.3071297
- Mathur, P., Sathishkumar, K., Chaturvedi, M., Das, P., Sudarshan, K. L., Santhappan, S., et al. ICMR-NCDIR-NCRP Investigator Group (2020). Cancer statistics, 2020: Report from national cancer registry programme, India. *JCO Glob. Oncol.* 6, 1063–1075. doi:10.1200/GO.20.00122
- Mohiyuddin, A., Basharat, A., Ghani, U., Abbas, S., Naeem, O. B., Rizwan, M., et al. (2022). Breast tumor detection and classification in mammogram images using modified YOLOv5 network. *Comput. Math. Methods Med.* 2022, 1359019. doi:10.1155/2022/1359019
- Orel, S. G., Kay, N., Reynolds, C., and Sullivan, D. C. (1999). BI-RADS categorization as a predictor of malignancy. *Radiology* 211 (3), 845–850. doi:10.1148/radiology.211.3.r99jn31845
- Prusty, S., Dash, S. K., and Patnaik, S. (2022). A novel transfer learning technique for detecting breast cancer mammograms using VGG16 bottleneck feature. *ECS Trans.* 107 (1), 733–746. doi:10.1149/10701.0733ecst
- Ragab, D. A., Attallah, O., Sharkas, M., Ren, J., and Marshall, S. (2021). A framework for breast cancer classification using multi-DCNNs. *Comput. Biol. Med.* 131, 104245. doi:10.1016/j.combiomed.2021.104245
- Ragab, D. A., Sharkas, M., Marshall, S., and Ren, J. (2019). Breast cancer detection using deep convolutional neural networks and support vector machines. *PeerJ* 7, e6201. doi:10.7717/peerj.6201
- Rahman, A. S. A., Belhaouari, S. B., Bouzerdoum, A., Baali, H., Alam, T., and Eldaraa, A. M. (2020). “Breast mass tumor classification using deep learning,” in 2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT), Doha, Qatar, February 2–5, 271–276.
- Rampun, A., Scotney, B. W., Morrow, P. J., and Wang, H. (2018). “Breast mass classification in mammograms using ensemble convolutional neural networks,” in 2018 IEEE 20th International Conference on e-Health Networking, Applications and Services (Healthcom), Ostrava, Czech Republic, 17–20 September 2018, 1–6.
- Saber, A., Sakr, M., Abo-Seida, O. M., Keshk, A., and Chen, H. (2021). A novel deep-learning model for automatic detection and classification of breast cancer using the transfer-learning technique. *IEEE Access* 9, 71194–71209. doi:10.1109/access.2021.3079204
- Saffari, N., Rashwan, H. A., 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. doi:10.3390/diagnostics10110988
- Shams, S., Platania, R., Zhang, J., Kim, J., Lee, K., and Park, S. J. (2018). “Deep generative breast cancer screening and diagnosis,” in International Conference on Medical Image Computing and Computer-Assisted Intervention, Granada, Spain, September 16–20, 2018, 859–867.
- Shi, P., Wu, C., Zhong, J., and 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), Qingdao, China, Aug. 23 2019 to Aug. 25 2019, 102–109.

- Simonyan, K., and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.
- Srinivasu, P. N., SivaSai, J. G., Ijaz, M. F., Bhoi, A. K., Kim, W., and Kang, J. J. (2021). Classification of skin disease using deep learning neural networks with MobileNet V2 and LSTM. *Sensors* 21 (8), 2852. doi:10.3390/s21082852
- Su, Y., Liu, Q., Xie, W., and Hu, P. (2022). YOLO-LOGO: A transformer-based YOLO segmentation model for breast mass detection and segmentation in digital mammograms. *Comput. Methods Programs Biomed.* 221, 106903. doi:10.1016/j.cmpb.2022.106903
- Sukegawa, S., Yoshii, K., Hara, T., Yamashita, K., Nakano, K., Yamamoto, N., et al. (2020). Deep neural networks for dental implant system classification. *Biomolecules* 10 (7), 984. doi:10.3390/biom10070984
- Sun, H., Li, C., Liu, B., Liu, Z., Wang, M., Zheng, H., et al. (2020). AUNet: Attention-guided dense-upsampling networks for breast mass segmentation in whole mammograms. *Phys. Med. Biol.* 65 (5), 055005. doi:10.1088/1361-6560/ab5745
- Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA a cancer J. Clin.* 71 (3), 209–249. doi:10.3322/caac.21660
- Szegedy, C., Ioffe, S., Vanhoucke, V., and Alemi, A. A. (2017). "Inception-v4, inception-resnet and the impact of residual connections on learning," in Thirty-first AAAI conference on artificial intelligence, Hilton San Francisco, San Francisco, California, USA, February 4–9.
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., et al. (2015). "Going deeper with convolutions," in Proceedings of the IEEE conference on computer vision and pattern recognition, Boston, Massachusetts, Held 7–12 June 2015, 1–9.
- Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., and Wojna, Z. (2016). "Rethinking the inception architecture for computer vision," in Proceedings of the IEEE conference on computer vision and pattern recognition, Las Vegas, NV, USA, June 26 2016 to July 1 2016, 2818–2826.
- Ting, F. F., Tan, Y. J., and Sim, K. S. (2019). Convolutional neural network improvement for breast cancer classification. *Expert Syst. Appl.* 120, 103–115. doi:10.1016/j.eswa.2018.11.008
- Toğaçar, M., Özkurt, K. B., Ergen, B., and Cömert, Z. (2020). BreastNet: A novel convolutional neural network model through histopathological images for the diagnosis of breast cancer. *Phys. A Stat. Mech. its Appl.* 545, 123592. doi:10.1016/j.physa.2019.123592
- Tsochatzidis, L., Costaridou, L., and Pratikakis, I. (2019). Deep learning for breast cancer diagnosis from mammograms—A comparative study. *J. Imaging* 5 (3), 37. doi:10.3390/jimaging5030037
- Urban, G., Tripathi, P., Alkayali, T., Mittal, M., Jalali, F., Karnes, W., et al. (2018). Deep learning localizes and identifies polyps in real time with 96% accuracy in screening colonoscopy. *Gastroenterology* 155 (4), 1069–1078. doi:10.1053/j.gastro.2018.06.037
- Wan, X., Zhang, X., and Liu, L. (2021). An improved VGG19 transfer learning strip steel surface defect recognition deep neural network based on few samples and imbalanced datasets. *Appl. Sci.* 11 (6), 2606. doi:10.3390/app11062606
- Wang, C. R., Zhang, F., Yu, Y., and Wang, Y. (2020). "BR-GAN: Bilateral residual generating adversarial network for mammogram classification," in International Conference on Medical Image Computing and Computer-Assisted Intervention, Lima, Peru, October 4–8, 2020, 657–666.
- Wen, L., Li, X., and Gao, L. (2020). A transfer convolutional neural network for fault diagnosis based on ResNet-50. *Neural Comput. Appl.* 32 (10), 6111–6124. doi:10.1007/s00521-019-04097-w
- Wu, C., Wen, W., Afzal, T., Zhang, Y., and Chen, Y. (2017). "A compact dnn: Approaching googlenet-level accuracy of classification and domain adaptation," in Proceedings of the IEEE conference on computer vision and pattern recognition, Honolulu, Hawaii, USA, Held 21–26 July 2017, 5668–5677.
- Xia, X., Xu, C., and Nan, B. (2017). "Inception-v3 for flower classification," in 2017 2nd international conference on image, vision and computing, Chengdu, Chin, June 2–4, 2017, 783–787. ICIVC
- Xu, C., Lou, M., Qi, Y., Wang, Y., Pi, J., and Ma, Y. (2021). Multi-scale attention-guided network for mammograms classification. *Biomed. Signal Process. Control* 68, 102730. doi:10.1016/j.bspc.2021.102730
- Zeiser, F. A., da Costa, C. A., Zonta, T., Marques, N., Roehe, A. V., Moreno, M., et al. (2020). Segmentation of masses on mammograms using data augmentation and deep learning. *J. digital imaging* 33 (4), 858–868. doi:10.1007/s10278-020-00330-4
- Zhang, H., Wu, R., Yuan, T., Jiang, Z., Huang, S., Wu, J., et al. (2020). DE-ada: A novel model for breast mass classification using cross-modal pathological semantic mining and organic integration of multi-feature fusions. *Inf. Sci.* 539, 461–486. doi:10.1016/j.ins.2020.05.080
- Zhang, R., Zheng, Y., Mak, T. W. C., Yu, R., Wong, S. H., Lau, J. Y., et al. (2016). Automatic detection and classification of colorectal polyps by transferring low-level CNN features from nonmedical domain. *IEEE J. Biomed. Health Inf.* 21 (1), 41–47. doi:10.1109/JBHI.2016.2635662