

Article

DeepBreastCancerNet: A Novel Deep Learning Model for Breast Cancer Detection Using Ultrasound Images

Asaf Raza ^{1,2}, Naeem Ullah ¹, Javed Ali Khan ³ , Muhammad Assam ⁴, Antonella Guzzo ¹ 
and Hanan Aljuaid ^{5,*} 

¹ Department of Computer Engineering, Modeling Electronics and Systems Engineering, University of Calabria, 87036 Rende Cosenza, Italy

² Department of Software Engineering, University of Engineering and Technology, Taxila 47050, Pakistan

³ Department of Software Engineering, University of Science and Technology, Bannu 28100, Pakistan

⁴ Department of Computer and Technology, Zhejiang University, Hangzhou 310027, China

⁵ Computer Sciences Department, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University (PNU), P.O. Box 84428, Riyadh 11671, Saudi Arabia

* Correspondence: Haaljuaid@pnu.edu.sa

Abstract: Breast cancer causes hundreds of women's deaths each year. The manual detection of breast cancer is time-consuming, complicated, and prone to inaccuracy. For Breast Cancer (BC) detection, several imaging methods are explored. However, sometimes misidentification leads to unnecessary treatment and diagnosis. Therefore, accurate detection of BC can save many people from unnecessary surgery and biopsy. Due to recent developments in the industry, deep learning's (DL) performance in processing medical images has significantly improved. Deep Learning techniques successfully identify BC from ultrasound images due to their superior prediction ability. Transfer learning reuses knowledge representations from public models built on large-scale datasets. However, sometimes Transfer Learning leads to the problem of overfitting. The key idea of this research is to propose an efficient and robust deep-learning model for breast cancer detection and classification. Therefore, this paper presents a novel DeepBraestCancerNet DL model for breast cancer detection and classification. The proposed framework has 24 layers, including six convolutional layers, nine inception modules, and one fully connected layer. Also, the architecture uses the clipped ReLu activation function, the leaky ReLu activation function, batch normalization and cross-channel normalization as its two normalization operations. We observed that the proposed model reached the highest classification accuracy of 99.35%. We also compared the performance of the proposed DeepBraestCancerNet approach with several existing DL models, and the experiment results showed that the proposed model outperformed the state-of-the-art. Furthermore, we validated the proposed model using another standard, publically available dataset. The proposed DeepBraestCancerNet model reached the highest accuracy of 99.63%.

Keywords: medical image classification; disease detection; deep learning; breast cancer; convolutional neural network

check for
updates

Citation: Raza, A.; Ullah, N.; Khan, J.A.; Assam, M.; Guzzo, A.; Aljuaid, H. DeepBreastCancerNet: A Novel Deep Learning Model for Breast Cancer Detection Using Ultrasound Images. *Appl. Sci.* **2023**, *13*, 2082. <https://doi.org/10.3390/app13042082>

Academic Editor: Jan Egger

Received: 23 December 2022

Revised: 29 January 2023

Accepted: 30 January 2023

Published: 6 February 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Cancer, which is defined as the uncontrolled and disruptive growth of abnormal body cells, is one of the major causes of death globally. The two primary types of cancer are malignant and benign, based on benignity and malignancy. The benign cancer cell is often non-cancerous and develops extremely slowly. Malignant tumors, on the other hand, grow quickly, are dangerous, and spread to other parts of the body via blood flow [1]. Among various cancers, lungs, bones, blood, brain, liver, and breast Cancer (BC) are commonly found in women. World Health Organization reports show that around 2.1 million women have life-threatening BC [2].

Patients diagnosed with BC and a tumor less than 10 mm in size have a 98% probability of surviving the disease. The tumor size at the diagnosis is highly associated with BC survival, and 70% of BC cases are diagnosed when the tumor is 30 mm or larger. [3]. Hence, breast tumor size deeply impacts the survival of a patient and its detection. Various imaging techniques investigated the breast tumor include X-ray [4] CT scan and ultrasound [5,6]. However, these methods are still unreliable, and sometimes misclassification leads to preventable surgery of a benign tumor that is non-cancerous [7]. Early identification and treatment are part of the current approach to BC. In the United States, this strategy has a 10-year survival rate of 85%. Survival is substantially correlated with the disease stage upon diagnosis, with a 10-year survival rate of 98% for patients with stages 0 and I cancer and a 10-year survival rate of 65% for those with stage III cancer. Therefore, detecting BC immediately and accurately is crucial for avoiding needless surgery and improving patient survival [8]. Separating normal and abnormal cells is challenging due to the size, shape, and location of tumor cells. Therefore, the medical image segmentation technique identifies abnormalities [9].

In the recent past, DL has played a tremendous role in different tasks of medical image processing, such as detection [10], classification [11] Segmentation [12]. A Convolutional neural network (CNN) requires an extensive data set for the training process. However, there is a limitation of BC data; therefore, we will use a TL technique to address the problem of limited data will increase the detection accuracy.

In CNN, the convolutional layers are used for feature extraction, pooling layers are utilized to reduce the feature and map and, finally, a fully connected layer classifies these features. More importantly, in convolutional neural networks, the initial layer extracts low-level features; however, as the number of layers increases, it extracts high-level features. Therefore, deep convolutional layers of CNN architecture extract more detailed and deep features. The inspiration to apply the pretrain DL model is that these models are trained on a huge dataset, i.e., ImageNet [13]. Thus, when pre-train models are trained for a new task such as classification and detection, the training accuracy increases and computational time is reduced. The Transfer Learning (TL) methodology is commonly used in different detection and classification research problems.

In recent years, a number of DL algorithms have been applied to identify BC, producing cutting-edge results in a BC detection applications. Despite the extensive spectrum of research on the classification and identification of BC, there is still interest in creating high-accuracy automated methods for BC detection and classification. According to our research, BC does not have a lot of cutting-edge literature when it comes to AI-driven technologies that utilize image methods. The low accuracy of BC detection is one of the shortcomings of the BC detection study that is currently published. The majority of research used datasets with fewer images (small datasets). There is a lack of training data, the model is not completely generalizable, and the training samples may have been overfit. To identify and rate BC, the majority of research employ traditional ML and transfer learning methods. However, the biggest problem with classical ML (such as support vector machine (SVM)) is the lengthy training period for big datasets [14]. The most problematic shortcomings in transfer learning systems, however, are negative transfer and generalization. The concept that pre-trained classification techniques are usually trained on the ImageNet database, which contains images irrelevant to medical imaging, is one of its disadvantages. As a result, developing efficient CADs to rapidly and accurately distinguish BC from ultrasound remains challenging. The DeepbreastcancerNet DL model is recommended to address these limitations. It makes use of filter-based feature extraction, which can aid in achieving excellent classification performance. Convolutional layer, clipped ReLu, and Leaky ReLu activation functions were used to develop the proposed model, which extracts the most essential and detailed features from the chest BC images. The proposed framework has 24 layers, including six convolutional layers, nine inception modules, and one fully connected layer. Through the use of a max-pooling operation, the design may reduce a number of weight characteristics. The suggested model, which

is a novel method for BC detection and classification, includes batch and cross channel normalizing procedures as well as both clipped Relu and leaky ReLu activation functions. Furthermore, we adopted the strategy of TL for BC detection using nine pre-train DL model i.e., ResNet101 [15], ResNet-50 [16], ResNet-18 [17,18], GoogLeNet [19], ShuffleNet [19], AlexNet [20], SqueezeNet [21]. XceptionNet [22] The frozen weights of pretrain models are transformed to the target cancer dataset and the final softmax layer is replaced by the new one, as we have three classes. Therefore, the new softmax is adjusted three.

The main contribution of this research paper is as follows.

1. We proposed a DeepBreastCancerNet deep learning model for breast cancer detection and classification.
2. We use data augmentation to increase model performance and avoid the problem of overfitting.

The rest of the paper is organized as follows. Section 2 provides a literature review, Section 3 materials and method, Section 4 is about experiments and results, and Section 5 has a conclusion and future work.

2. Literature Review

S. A. Mohammed et al. [23] balanced two different imbalanced BC datasets by resampling and removing the missing attributes. They used three machine learning classifiers (Naïve Bayes, MO, and J48) to classify the balanced data. Classification results demonstrated that balancing data enhanced the classification results. M. Tahmoorei et al. [24] used different machine learning classifiers such as SVM, KNN, ANN (Artificial neural network), and a Decision tree for BC detection. They attained an accuracy of 98.8% accuracy with SVM. Hussain et al. [25] extracted different image features using different image descriptors such as texture (SIFT), (EFDs), morphological entropy and SVM, decision tree, and Bayesian as classifiers to discriminate normal mammography from BC. Goyal et al. [26] Used machine learning for BC classification. They applied principal component analysis for dimensionality reduction in features extraction and utilized the classifier of CNN for BC dataset features classification. Jain et al. [27] applied five different machine learning classifiers, namely SVM, DT, KNN, logistic regression, and random forest for BC detection. KNN and logistic regression in 10-fold cross-validation attained an accuracy of 96.52%. Twenty-eight different combination architectures for BC classification were proposed by Zerouaoui and Idri [28]. These hybrid architectures combined seven different DL models along with feature extraction and four different classifiers (DT, SVM, KNN, with MLP) for BC classification. On the FNCA and BreakHis datasets, the authors match the total performance of several architectural approaches. On the FNCA dataset, the classification accuracy achieved by the Densenet-201 combine with the MLP classifier was 99%.

Alanazi et al. CNN-based system for automatic BC detection and compare the results with Machine learning algorithms. The experimental displayed that the proposed system performed 9% better than machine learning classifiers and reached 87% accuracy. Togaçar et al. [29] proposed custom CNN model for BC ultrasound image detection. The proposed consist of only one convolution layer with 20 filters, which achieved 100% accuracy while outperforming eight different pre-train CNN models. Ting et al. [30] presented CNN for BC detection (CNNI-BCC). This framework uses a supervised DL neural network to do the classification. The experimental findings demonstrated that CNNI-BCC performed better than previously conducted research and achieved an accuracy of 89.47% Nahid et al. [31] Presented A novel deep neural model for BC classification utilizing Histopathological images, comprised of a clustering algorithm and CNN. The model is based on CNN and LSTM. At the classification layer of the model, Softmax and SVM are both used. The proposed model achieved 91% accuracy. Ragb et al. [32] presented their own proposed CNN model for the diagnosis of BC in their study. In addition, they employed a TL technique to combine nine different pre-trained DL models to classify two BC datasets. The suggested model obtained the finest accuracy of 99.5%.

Bevilacqua et al. [33] used MR images chosen for training and testing. They employed ANN to classify and detect BC after extracting data and analyzing it. However, When the Machine Learning model was applied to enhance the ANN, the maximum accuracy was increased to 100%. Khan et al. [34] Proposed a framework for detecting and classifying BC TL approaches were adopted by the authors as opposed to GoogLeNet, VGG, and ResNet for feature extraction. Following that, a fully connected layer is provided with the combined extracted characteristics for classification. The unique framework that was suggested has a classification accuracy of 97.52%. Tang et al. [35] provided a detailed overview of recent advanced development in computer-aided detection or diagnosis of BC detection. Joseph et al. [36] Extracted handcrafted features such as textures, color, and shape from BC Histopathological images; classified these features with the deep neural network using softmax. Furthermore, the author applied data augmentation to address the problem of overfitting. Experimental results demonstrated that the projected approach accomplished an accuracy of 97.87% for 40x. Sharma and Kumar [37] compared handcrafted features with the proposed XceptionNet model for features extraction for BC classification. The XceptionNet used as feature extractor and SVM as a classifier outperformed the handcrafted feature and achieved a classification accuracy of 96.25% for a 40x level of magnification. Inan et al. [38] proposed a unified automatic framework for breast tumor segmentation and detection based on segmentation leading to classification. They utilized SLIC and K Means++ for tumor segmentation and VGG16, VGG19, DenseNet121, and ResNet50 for breast tumor classification. Furthermore, SLIC, UNET, and VGG16 outclassed all other combined combinations. Jabarani et al. [39] proposed a new hybrid approach for BC detection. On images from the dataset, the researchers used an adaptive median filter to reduce noise and some preprocessing approaches to remove the background. They also modify the segmentation-related parameter values for K Means and the Gaussian mixture framework. The accuracy of the suggested hybrid model was 95.5%.

Saber et al. [40] presented the TL approach, and six performance elevation matrices are used in the DL model suggested for automatically detecting BC. They deployed five pre-train DL models for feature extraction and classification. These models were the Inception-V3 ResNet, the VGG19 ResNet, the VGG16 ResNet, and the Inception-V3. It has been shown that the VGG16 model is suitable for BC detection, with an accuracy of 98.96 percent. A CNN-based hybrid model for BC classification was proposed by Eroglu et al. in [41]. The Mobilenetv2, Resnet50, Alexnet, and Resnet50 neural networks extracted information from BC ultrasound images. The cooperation of the mRMR technique is used to link the extracted characteristics together in order to increase accuracy. The support vector machine (SVM) is then implemented to categorize these features. The suggested hybrid model attained an accuracy of classification of 95.6%. Convolutional Neural Network (CNN) and Transfer Learning are used in Chowdhury et al. [42] suggested method to improve diagnosis by boosting the efficacy and accuracy of early breast cancer detection. Instead of creating the entire model from scratch, the thought process entailed leveraging a pre-trained model that already had certain weights given. This research primarily focuses on a Transfer Learning Model built on ResNet101 and used with the ImageNet dataset. The proposed framework gave us a 99.58% accuracy rate.

Moon et al. [43] In their research, they intended to detect BC from ultrasound. In their research, they used two different data sets. They combined multiple CNN algorithms using the image fusion approach. In their initial attempt, they claimed to have attained an accuracy of 91.10%. They obtained 94.62% in their first dataset and 94.62% in their second Xiao et al. [44] Examined InceptionV3, ResNet50, and XceptionNet and presented a basic model for identifying breast malignancies from the breast Neural Computing and Applications 123 ultrasound images dataset, consisting of three convolutional layers. There were 2058 images in the dataset, including benign 1370, while malignant 688 samples. According to their research, InceptionV3 had the highest accuracy (85.13%). Jean el al. [45] the offered approach is composed of a few sequential steps. The breast ultrasound data are first enhanced, and then a DarkNet-53 deep learning model is used to retrain the data. The

features were then retrieved from the pooling layer, and the two alternative optimization techniques, such as the reformed BGWO and the reformed DE, were used to choose the best feature. The last step involves fusing the chosen features using a proposed technique, which is then classified using machine learning techniques.

3. Materials and Method

3.1. Dataset

In this study, we utilized two publicly available BC datasets. Both datasets are being evaluated for the given reasons: (1) to enhance the dataset size for training purposes to minimize overfitting and bias, and (2) to include three classes (malignant, normal, and benign). The model's effectiveness will also be enhanced by integrating the datasets. The BUSI dataset of BC ultrasound images for experimentation. The dataset is already used in the literature. The data was gathered from 600 women ranging in age from 25 to 75 in 2018. The dataset in [46] consists of 780 images, including 437 benign, 210 malignant and 133 normal ultrasound images (<https://doi.org/10.1016/j.dib.2019.104863> (19 January 2023)). The sample ultrasound image from the dataset is shown in Figure 1. All images are 500×500 pixels in size and in PNG format. The dataset in [47] contains 250 BC images, 100 benign and 150 malignant (<https://data.mendeley.com/datasets/wmy84gzngw/1> (19 January 2023)). The details of BC datasets are presented in Table 1.

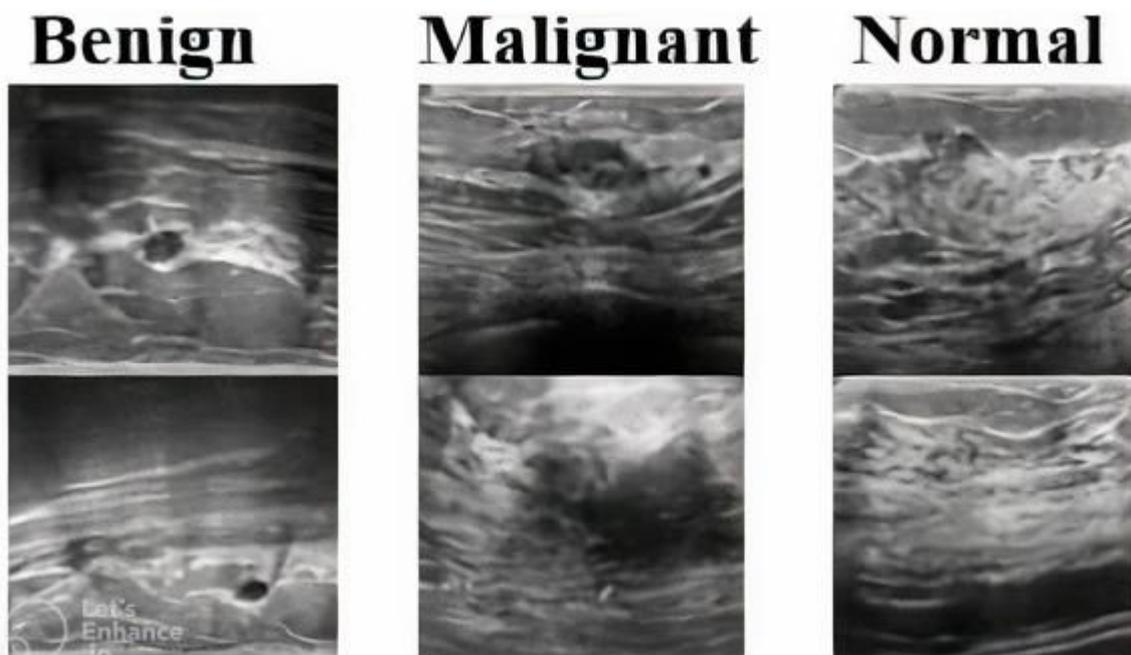


Figure 1. BC ultrasound images from dataset.

Table 1. Detail of BC dataset.

DATASETS	BENIGN	MALIGNANT	NORMAL	TOTAL
DATASET1	437	210	133	780
DATASET2	100	150	-	250
TOTAL	537	360	133	1030

3.2. Methodology

3.2.1. Motivation

For Breast cancer detection, DL algorithms may be an effective method. Similar issues have been addressed using DL techniques, including the classification of skin cancer [48], the classification of Parkinson's disease and other brain disorders, and the diagnosis of pneumonia using chest radiograph images. In [29] the authors proposed custom CNN

model for BC ultrasound image detection. The proposed consist of only one convolution layer with 20 filters. Similarly, in [31] the authors presented CNN for BC detection (CNNI-BCC). This framework uses a supervised DL neural network to do the classification. The experimental findings demonstrated that CNNI-BCC performed better than previously conducted research and achieved an accuracy of 89.47%. Furthermore, in [32] the authors presented their own proposed CNN model for the diagnosis of BC in their study. In addition, they employed a TL technique to combine nine different pre-trained DL models to classify two BC datasets. The suggested model obtained the finest accuracy of 99.5%. Subsequently, motivated by the effectiveness of DL-based architecture in detecting breast cancer using certain images. In this study, we suggested a DeepbreastcancerNet model for detecting and classifying breast cancer. The objective of this study is to suggest a DL model that can more accurately and successfully detect breast cancer from ultrasound images. The depth and input image resolution of the proposed model are based on the following information: It is generally accepted that a deeper DL-based model improves the network's classification performance by capturing more intricate and important deep features [49,50]. In addition, depth scaling was widely used by various DL-based models to increase their accuracy. However, when the network's depth increases, computational complexity rises as well, therefore accuracy may not always be improved. In order to achieve better performance, DL-based models gather more detailed features from high-resolution input photos. DL models can detect images with resolutions ranging from 224×224 to M and 299×299 , however, models with greater resolution often perform better. Similar to that, the suggested model has 24 layers and is capable of processing images with a 224 by 224 pixel resolution. The DeepbreastCancerNet architecture and input image size are chosen based on the needs of the current computational infrastructure.

In the suggested architecture, we utilized both cross-channel and batch normalization (BN) layers. Cross-channel normalization layer because it enhances generalization and lowers top-1 and top-5 error rates. Batch normalization reduces internal covariate shifting, which speeds up model training to address this issue, we used a leaky ReLu in the DeepbreastcancerNet strategy that was recommended [51]. To resolve this problem, we employed a leaky ReLu in the DeepbreastcancerNet strategy that was suggested. When the unit is inactive, the Leaky ReLu activation method permits a small (non-zero) gradient. As a result, it keeps learning instead of coming to a stop or hitting a brick wall. Additionally, Max pooling is used in our suggested study since it keeps the feature maps' most salient features, producing statistical basis features. Additionally, Max pooling is used in our proposed study as it retains the most significant features of the feature maps, producing precise classification features. For down sampling, max-pooling is employed with a filter size of 3×3 , At the conclusion of the structure, a global average pooling is utilized to combine each feature map into a single value.

The suggested architecture seems to do an excellent work of classifying and recognizing ultrasound images of breast cancer. It uses the benefits of batch normalization, cross channel normalization, leaky ReLu activation function, and ReLu activation function to enhance the performance of the proposed deepbreastcancer detecting and categorizing breast cancer. Additionally, the suggested model is deep enough to capture more intricate and crucial aspects and employs a high quality image of 256×256 . The suggested approach's architecture is described below.

3.2.2. Preprocessing and Data Augmentation

The image normalization is an essential phase before it is given to the CNN model. Therefore, both the dataset images are resized according to the input size of different models. DL algorithms have renowned for consuming much data. DL algorithms require a large quantity of data to train for high accuracy. This high demand for data makes it difficult for them to be used in regions or domains where data is hard to obtain. Although there is sufficient data in every area to train a DL system, we used the data augmentation strategy in this work to fill in the gaps. The training set's images were randomly translated

up to thirty pixels vertically and horizontally, and rotated at an angle between -30 and 30 degrees. To generate more images, they translated the images at random between $[0.9$ and $1.1]$. This aided in training a DL algorithm on a minimal quantity of data to produce reasonable results without overfitting the data’s majority class.

3.2.3. Deep Learning

Machine learning techniques are similar to artificial intelligence approaches that approximate how people obtain knowledge. To perform the classification task, traditional machine learning approaches involve a series of processes, including preprocessing steps including feature extraction, careful feature selection process, and the learning and classification stage. The efficacy of these approaches is highly reliant on the characteristics chosen, which may or may not be the best features for class discrimination. In contrast to standard machine learning approaches, DL allows for the automatic learning of feature sets for various tasks [52]. DL, which incorporates forecasting and statistics, is a crucial part of data science. A deep neural network used to analyze visual data is called a convolutional neural network. CNN is a DL algorithm that takes an input image and gives different parts of the image different weights, enabling it to distinguish between them. CNN categorizes and identifies images because of its high accuracy [53,54].

3.2.4. DeepBreastCancerNet Architecture

In this paper we proposed DeepBraestNet model for BC detection and classification. Table 2 provides more details on the proposed framework’s design. Figure 2 depicts the abstract perspective of the proposed DeepBreastCancerNet technique, which consists of three main components. The suggested architecture is more complex than conventional CNN. The DeepBraestNet of 24 layers, i.e., six convolutional layers and nine inception modules. The presented model’s first layer serves as the input layer, allowing 224×224 input images for processing. The architecture has a total of 24 layers. Four batch Normalization Layer and max-pooling layers. two normalization layers, one fully connected layer, two activation functions, i.e., (57 leaky Relu and three Clipped Relu) and finally a linear layer with softmax activation in the output.

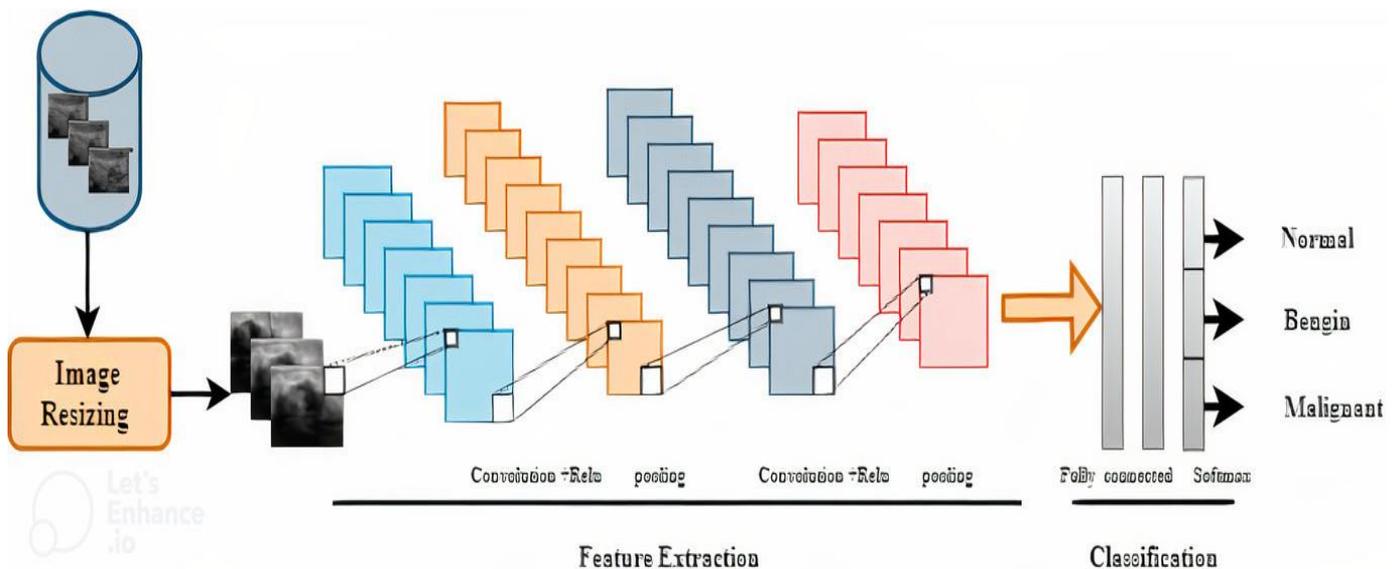


Figure 2. The abstract perspective of the proposed DeepBreastCancerNet technique.

Table 2. Characteristics of the proposed DeepBreastCancerNet model.

S.No	Type	No of Filter	Filter Size
1	Input layer		
2	Convolution (Leaky Relu + Cross channel normalization)	64	7×7
3	maxpooling		3×3
4	Convolution 2 (Leaky Relu + Cross channel normalization)	192	3×3
5	Max pooling		3×3
6	Inception 1	256	
7	Inception 2	480	
8	Max pooling		3×3
8	Inception 3	256	
9	Inception 4	280	1×1
10	Inception 5	512	
11	Inception 6	512	
12	Inception 7	512	
13	Max pooling	512	3×3
14	Inception 8	528	
15	Inception 8	832	
16	Convolution (Clipped Relu + Batch normalization)	960	1×1
17	Group Convolution (Clipped Relu + Batch normalization)	960	3×3
18	Convolution (Batch normalization)	320	1×1
19	Convolution (Clipped Relu + Batch normalization)	1280	1×1
20	Global Average Pooling		
21	Fully Connected Layer		
22	Softmax		
23	Classification		

The first convolution layer adopted a filter (patch) size of 7×7 , which substantially cut the size of the image. The second convolution layer used the 1×1 convolution block, which is the result of high dimensional data, and had a depth of two. The leaky Relu, max pooling, and cross-channel sectional normalization layers come after the first two convolution layers. Furthermore, each inception module also includes six convolutional layers, of which four are utilized for dimension reduction, as well as one max-pooling layer. All fully connected layers employ the Leaky ReLU activation function. The DeepBraestNet inception module also offers a variety of convolution kernels, including convolution kernels of 1×1 , 3×3 , and 5×5 , which extract features at various scales, going from the most delicate features to the core features. For the purpose of computing the features, the greater convolution kernel spans a wider region. Similar to that, the 1×1 convolution kernel provides more information while requiring less processing.

Convolution sub layers with a maximum pooling of three are used in inception modules with one, three, and five layers. Operation block doing arbitrary parallel operations. These inception blocks take data from the layers above and do any number of concurrent operations on it. Prior to parallel convolution procedures, the 1×1 convolution technique is used to reduce calculation loss. However, inside the inception module, the 1×1 convolution sub layer is positioned after the for max pooling layer. Different characteristics are computed in each branch of the Inception layer using the information from the preceding layer. Next, CNN is concatenated with each output. This model overcomes the overfitting issue by employing inception modules rather than fully connected layers.

The last 4 convolution layers (Including one group convolution layers) adopted a filter (patch) size of 1×1 and 3×3 respectively. Moreover, these convolution layers are followed by global average pooling, Batch normalization and clipped ReLu Activation Function which increase the accuracy and expressiveness of the proposed model. Equations (1) and (2) illustrate how activation functions are mathematically calculated. In Equation (3), the pooling procedure is described. In the context of Breast detection

and classification the output of the last FC layer is sent as an input to 3-way Softmax (tri-class classification).

Leaky ReLU:

$$f(x) = \begin{cases} x, & x \geq 0 \\ scale * x, & x < 0 \end{cases} \tag{1}$$

Clipped ReLU:

$$f(x) = \begin{cases} 0, & x < 0 \\ x & 0 \leq x < ceiling \\ ceiling, & x \geq ceiling \end{cases} \tag{2}$$

$$S = w2 * h2 * d2 \tag{3}$$

3.3. Transfer Learning

Deep CNNs techniques are still often employed in current research. For a number of detecting issues, they provide innovative solutions. Lack of training data is a common problem when utilizing deep CNN models, which require a large amount of data to execute well. Additionally, gathering a large dataset takes time and is currently underway. As a result, the small dataset issue has now been addressed using the TL technique [54]. This strategy is exceptionally successful when there is a lack of training data. TL (TL) is when CNN models are trained on large datasets and then fine-tuned to train on a smaller desired dataset. Because the pre-trained model already understands all of the basic features, so training time is greatly reduced compared to training from scratch or without the TL approach. Compared to learning from scratch, memory and computing resources will be minimized. Many pre-trained algorithms are trained on big datasets, such as the ImageNet dataset [55], with over 15 million images divided into over 22,000 categories. The TL approaches heavily rely on these pre-trained algorithms. However, the pertained model utilized in this study was trained using the BC ultrasound images dataset, not the ImageNet dataset. In computer vision problems, TL is the most extensively utilized strategy. Nine different CNN models are fine-tuned to train on the target BC dataset, as shown in Figure 3. We fine-tuned eight different pre-trained CNN models for BC detection and classification in the research. The complete information about the number of parameter and layers of various CNN architectures are shown in Table 3.

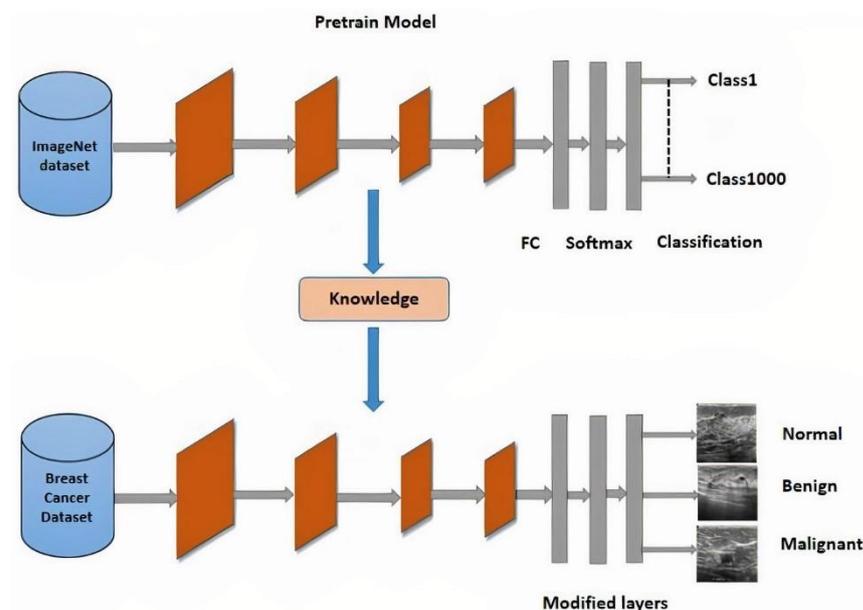


Figure 3. TL of Pretrain models.

Table 3. Presents different layers and parameters along with size of various CNN architectures.

Model	No of Layers	Parameters	Size
AlexNet	8	60	227 Mb
ResNet-101	101	44	171 MB
ResNet-50	50	25.6	96 MB
ShuffleNet	50	1.4	6 MB
ResNet-18	18	11.7	44 MB
GoogLeNet	22	7	27 MB

3.3.1. XceptionNet

Inception has been replaced by the Xception network. Extreme inception is referred to as XceptionNet. Traditional convolution layers are replaced with depth-wise separable convolution layers in the XceptionNet network. While XceptionNet includes mapping spatial and cross-channel correlations, CNN feature maps allow for the decoupling of spatial and cross-channel correlations. The core Inception architecture was outlived by XceptionNet. The XceptionNet model's 36 convolution layers may be divided into 14 distinct modules. Each layer still has a continuous residual link after the first and last layers have been removed. An input image is translated into determine the likelihood within each output channel in order to get cross-channel correlations in the input image. The following approach uses depth-wise 11 convolution.. Instead of using 3D maps, relationships may be displayed.

3.3.2. GoogLeNet

GoogLeNet this network contains 144 layers. The inception blocks are formed up of four branches. Rectified linear activation is used in all convolutions, including in the Inception modules. This module relies on a series of very small convolutions to decrease the cost of parameters. The architecture's 22 layers decreased the number of parameters from 60 million to only 4 million.

3.3.3. ShuffleNet

The ShuffleNet model contains 50 layers and requires a 224×224 input image for processing. In order to have high-level features, we built activations on the most recent global average pooling layer. 544 deep features are offered by ShuffleNet for image representation. The pre-trained version of ShuffleNet, which can be used to categories new jobs, was trained using several images from the ImageNet database.

3.3.4. AlexNet

AlexNet, there are 11 layers in the network. The large number of layers in the network aids feature extraction. Furthermore, the variety of parameters has a detrimental impact on performance. The convolution layer is AlexNet's initial layer. Following the maximum pooling and normalizing layers comes the convolution layer. The softmax layer is the final layer in the classification process.

3.3.5. ResNet-18

ResNet-18 was developed in 2018. ResNet consists of 71 layers. The numerous convolutions are followed by batch normalization and ReLU, while the residual connection takes advantage of an extra layer with two inputs. Average pooling, a fully connected layer, and softmax are the last layers.

3.3.6. SqueezeNet

The SqueezeNet model comprises 18 layers and requires a 227×227 input image to process. With 50 times fewer parameters, SqueezeNet achieved greater accuracy on

ImageNet. We utilized activations on the final fully connected layer to retrieve high-level features from the images. The Squeezenet model interprets the input image as 1000 deep features.

3.4. Hyperparameters Setting

Hyperparameters are important factors that must be specified before any model can be trained since they affect the learning process and are the model's primary component. There are several approaches to determining the criteria. We divide BC ultrasound images into training and testing using 70% and 30% ratios. Batch size is the number of training samples counted in a single forward and backwards pass. More memory space is required as the batch size grows.

The learning rate is a hyper-parameter that regulates how much the weights of our network are adjusted in relation to the loss gradient. We go down the slope more slowly the lower the value is. While adopting a low learning rate may be a good idea to ensure that we do not miss any local minima, it may also prolong the time it takes for us to converge, particularly if we become trapped on a plateau region. The word "epoch" is applied to the number of passes the machine learning algorithm has made across the full training dataset. Generally, datasets are organized into batches (especially when the amount of data is very large). One batch being run through the model is referred to by some as an iteration, which is a vague usage of the term. Overfitting is a crucial problem when training a neural network using sample data. When a neural network model is trained using more epochs than necessary, the training model largely learns patterns that are unique to the sample data. This prevents the model from functioning successfully on a new dataset. This model performs well on the training set (sample data), however it performs poorly on the test set. In other words, by overfitting to the training data, the model loses its ability to generalize. In a variant of the gradient descent process known as mini-batch gradient descent, the training dataset is divided into smaller batches that are then utilized to compute model error and update model coefficients. The variance of the gradient can be further reduced by implementations by summing the gradient over the mini-batch. Mini-batch gradient descent aims to strike a compromise between batch gradient descent's efficiency and stochastic gradient descent's resilience. It is the deep learning application of gradient descent that is utilized the most frequently.

To find the best hyperparameters (which provide high accuracy and little error) for the suggested DL model, we used a grid search strategy DL models that have already been trained by TL were trained using stochastic gradient descent (SGD). We used a 10-image minibatch size and a 0.001 learning rate. Additionally, to avoid for overfitting, each DL model was trained for 100 epochs before being used in the TL tests for identifying and classifying different forms of BC. We used a computer with an Intel (R) Core (TM) i5-5200U processor and 8 GB of RAM for all of our experiments. We employed MATLAB R2020a for the implementation. Table 4 shows the optimum parameters utilized in the categorization experiment.

Table 4. Parameters used for training.

Parameters	Values Given
Learning rate	0.001
Validation frequency	50
Optimization algorithm	SGDM
Shuffle	Every epoch
Iterations per epoch	7
Maximum epochs	100
Verbose	false
Activation function	Leaky ReLu + ReLu

4. Results and Discussion

4.1. Performance Evolution

To evaluate the performance of the models utilized in this study, we used accuracy, precision, recall and F1-score. Which may be calculated as follows:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (4)$$

$$precision = \frac{TP}{TP + FP} \quad (5)$$

$$Re\ call = \frac{FN}{FN + TP} \quad (6)$$

$$F1 - Score = 2 * \frac{(precision \bullet Re\ call)}{(precision + Re\ call)} \quad (7)$$

True Positive (TP): The positive data anticipated properly is estimated. The diagonal is the most important value.

True Negative (TN): Negative data was assessed to be negative. Except for the row and column of the associated class, it is the sum of all values in the confusion matrix.

False Positive (FP): Negative data was assessed to be positive. For each class, it is the total of all values in the relevant column except TP.

False Negative (FN): Positive data is assessed to be negative. For each class, it is the total of all values in the relevant row except TP.

Performance Evaluation of DeepBreastCancerNet

In this experiment, chest radiograph images are used to test the proposed DL model's detection (three-class classification) performance. The dataset was split into training and testing sets for this experiment, with 30% of the data being used for model testing and 70% being utilized for training. More specifically, we employed all 1030 BC ultrasound pictures (537 benign, 360 malignant, and 133 normal), of which 376 were used for training, 252 were used for testing, and 95 were of benign, malignant, and normal persons. The parameters presented in Table 2 are used in the training set to practice the proposed framework's BC detection and classification. The proposed DL model underwent 700 total iterations, averaging iterations per epoch throughout the course of the 7 epochs. The suggested DeepBreastCancerNet achieved the maximum classification accuracy, precision, recall, and F1-score values 99.35%, 99.60%, 99.66% and 99.60% at epoch 100, respectively, demonstrating the effectiveness of our method for detecting BC. Additionally, we've shown accuracy and loss in Figure 3 to demonstrate how well the proposed methodology performs throughout training and testing. How successfully the framework can predict the dataset is shown by the loss function. The loss and accuracy of our model after epoch 55 roughly remain the same, demonstrating that it is still more accurate at forecasting BC at lower epochs than 100. Figure 4 depicts the proposed DL approach's training and validation procedure as well as the confusion matrix for the testing phase of the Breast cancer detection framework for DL detection. The suggested Model classified incorrectly just 2 images. Figure 5. The proposed DL framework's training and validation loss (black line shows testing loss whereas red line shows training loss).

Furthermore, a detailed training parameters ablation study was conducted to demonstrate the validity of the proposed DL approach we have trained the proposed model with 20% of the data being used for model testing and 80% being utilized for training. The proposed DL model achieved the maximum classification accuracy, precision, recall, and F1-score values 98.50%, 98%, 98.50% and 98% respectively. Moreover, we used different training parameters we used a 10-image minibatch size and a 0.01 learning rate. Additionally, DL model was trained for 50 epochs before being for identifying and classifying different forms of BC. The suggested DeepBreastCancerNet achieved the maximum classification accuracy, precision, recall, and F1-score values 97.50%, 97%, 97.50% and 97%

respectively. Hence, it's been noticed that decreasing the number of epochs and change the data split effect the performance of the proposed model.

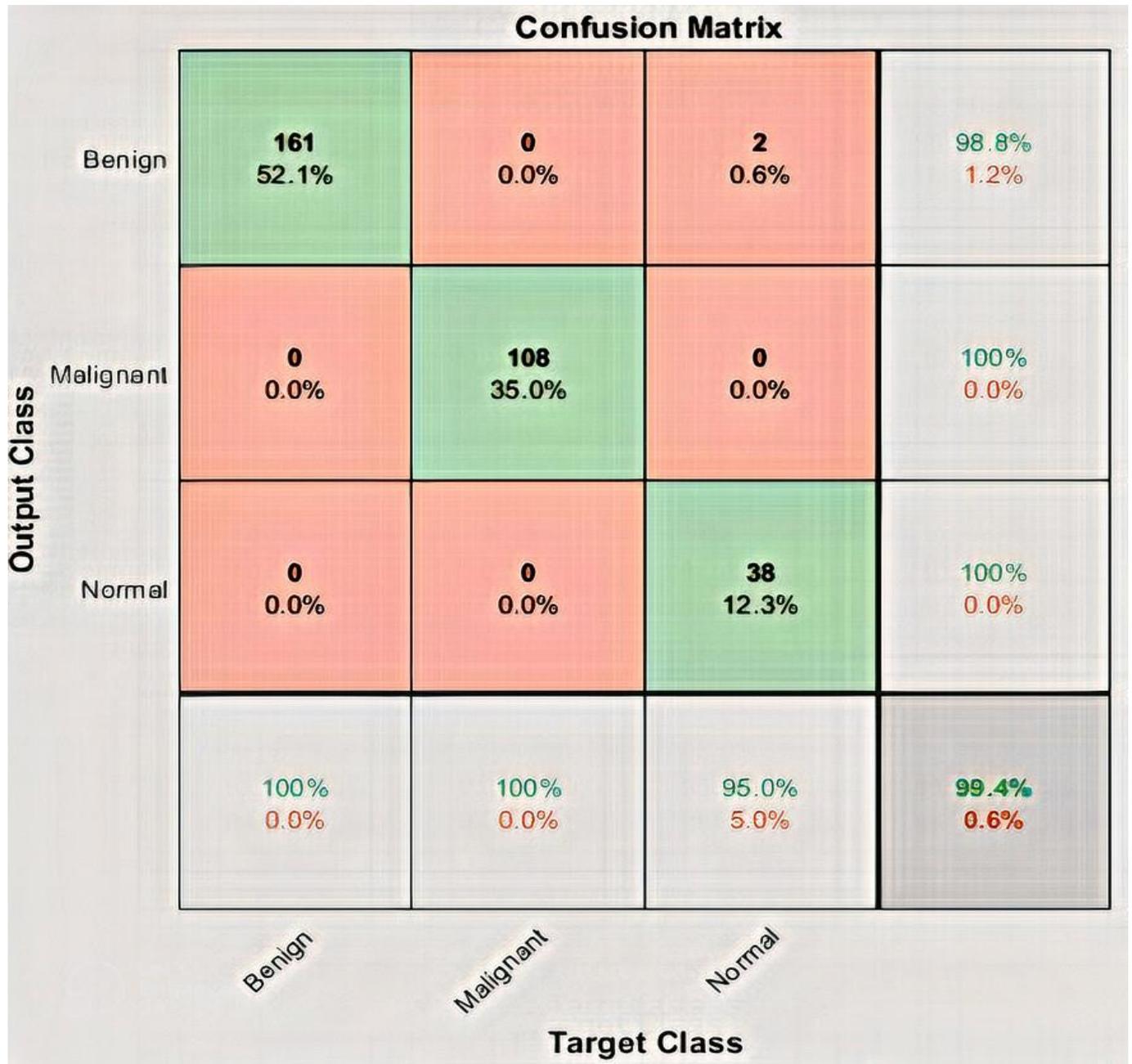


Figure 4. Confusion matrix of DeepBreastCancerNet.

This indicates how our suggested strategy can increase the accuracy of BC identification and classification from ultrasound images. These outcomes are result of the fact that our proposed approach is able to properly extract the most distinct, robust, and sophisticated deep features to represent the ultrasound image for precise and reliable identification. The recommended method requires a distinct feature extraction stage, making it simple to implement thanks to its end-to-end learning architecture.

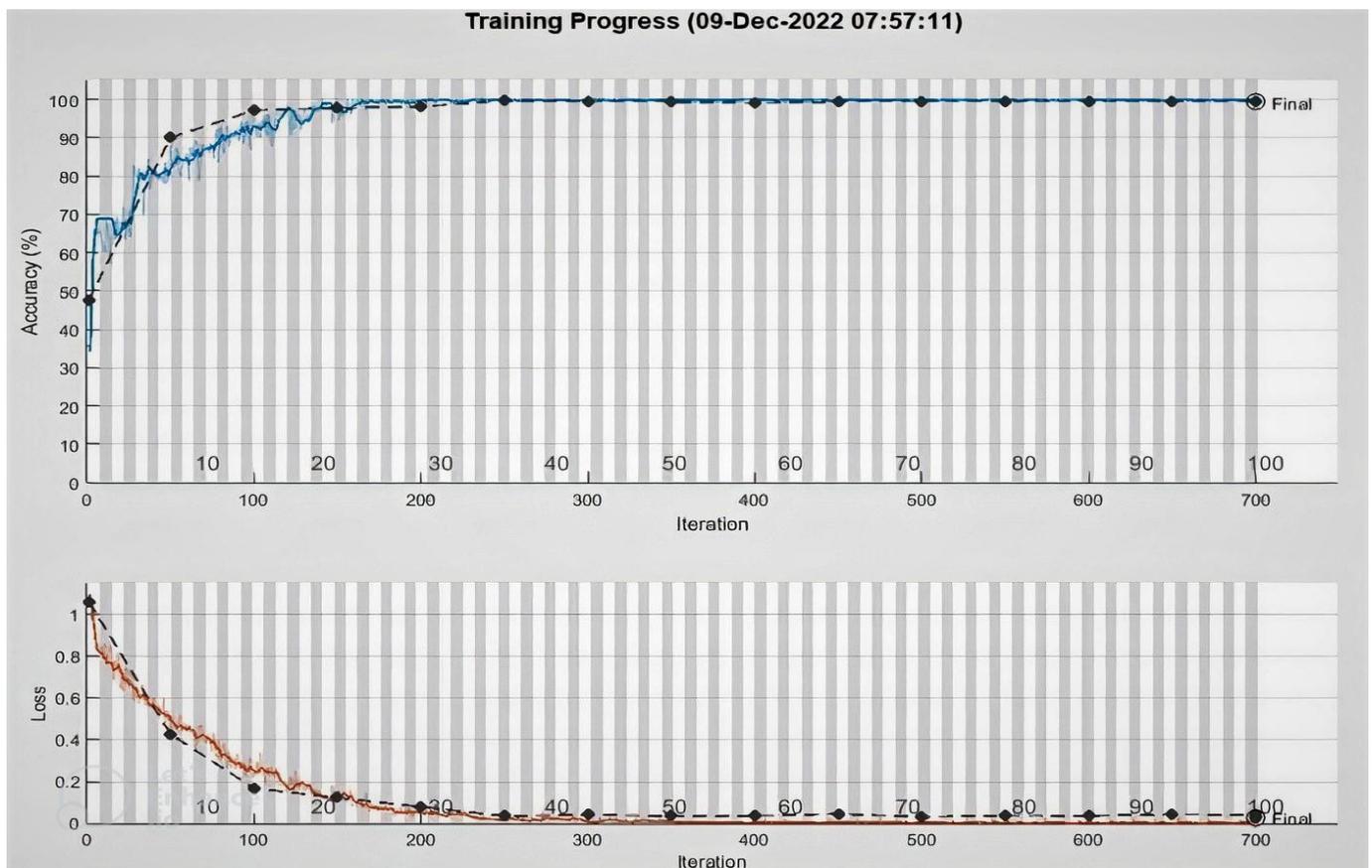


Figure 5. Accuracy and loss function graph of DeepBreastCancerNet.

4.2. Ablation Study

The concept of ablation experiments to examine the knowledge structure represented in DNNs. A detailed ablation investigation was conducted to demonstrate the validity of the proposed DL approach. In an ablation research, a layer, activation function, etc. of the deep learning model are removed or replaced in order to evaluate how each component contributes to the representation of the overall network. To be more specific, the performance of the DeepBreastCancerNet model is evaluated without the other model branch as it is removed. The sustainability of the DeepBreastCancerNet architecture is being evaluated as part of this ablation study, which is essential for figuring out how these components impact the system's performance. An ablation analysis was performed using two experiments that involved changing numerous components of the suggested framework. These experiments confirmed the impact of the two activation functions (Leaky ReLU and clipped) on the performance of the suggested model. In the first study, the feature extraction layers solely employed the Leaky ReLU function. Moreover we removed global average pooling. Similarly, in the second study, we replaced the clipped Relu activation function in the feature extraction layers with a global average pooling layer prior to the final FC layer. The performance of the ablated models is summarized in Table 5. According to the results, each component of the suggested DL model that is changed or removed makes the system perform worse.

Table 5. Modifying the evaluation network for the ablation study.

Experiment	Activation Function	Global Average Pooling	Results/Accuracy
Ablation Study 1	Leaky ReLU	No	99%
Ablation Study 2	Clipped ReLU	Yes	99.20%
Proposed Model	LeakyReLU + Clipped ReLU	Yes	99.35

4.3. Comparison with State of the Art Pretrain Deep Learning Models

The key benefit of using TL classifiers and fine tuning is that it reduces overfitting problems, which frequently occurred in DL algorithms when testing with a smaller sample of training and testing images. For the detection and classification of BC, all TL models were trained and validated using the same TL parameters as those specified in Table 4. To identify BC, we used 1030 ultrasound pictures of BC. Each TL classifier produced acceptable results, as shown in Table 6, which summarizes the specific results of several TL algorithms in classifying BC images. Using assessment measures for accuracy, precision, recall, and f-measure, we examined and assessed the TL algorithms. The dataset was divided into training and testing sets for this experiment, with 70% of the data utilized for model training and 30% for testing. Alexnet, Resnet18, SqueezeNet, Shufflenet, Resnet50, ExceptionNet, and Proposed DL-based classification techniques were used in the proposed study. The confusion matrix in Figure 4 of the Suggested model shows that the score is high in all BC classifications. It was witnessed that the DeepBreastCancerNet achieves the highest accuracy of 99.35%, Squeezenet achieves the lowest accuracy of 70.81%, while AlexNet achieves the second-lowest accuracy. ShuffleNet and ExceptionNet achieved the same accuracy of 99.03%. In addition, the findings beat those of the pre-trained models. At the same time, the categorization results of the models are demonstrated in Table 6. The DeepBreastCancerNet model outperforms all the pre-trained models concerning the accuracy, precision, recall, and F1- measure. When the Proposed model is trained and verified, the accuracy and loss values are shown in Figure 5. The graph in Figure 5 illustrates that, as indicated in Table 5, the custom model provides exceptionally high accuracy results.

Table 6. Presents pre-trained model experimental results.

Model	Accu	Prec	Recall	F1-Score
Proposed Model	99.35%	99.60%	99.66%	99.60%
AlexNet	89%	90.3%	90.3%	90.3%
GoogLeNet	98.71%	99%	97.66%	98%
Shufflenet	99.03%	99.33%	98%	99%
ResNet-50	98.06%	98.33%	98.66%	97.33%
SqueezeNet	70.81%	73%	72%	62%
ResNet-18	99%	99%	98.33%	99%
Resnet101	99%	99%	98.33%	98.33%
ExceptionNet	99.03%	98.33%	98%	99%

4.4. Comparisons with the State of the Art Literature

We evaluated the effectiveness of the best deep neural network, DeepBreastcancerNet, with other techniques for classifying BC tumors into benign, malignant, and normal. In further detail, we contrasted the proposed work with cutting-edge DL methods [33,35,41]. Khan et al. [34] proposed a novel framework for the classification and detection of BC. The authors used TL techniques, whereas they used GoogLeNet, VGG and ResNet for features extraction. After that, the combined extracted features are given into a fully connected layer for classification. The proposed novel framework reached 97.52% classification accuracy. Tang et al. provided a detailed overview of recent advanced development in computer-aided detection or diagnosis of BC detection. Saber et al. [41] the proposed DL model for automated BC identification makes use of six alternative performance elevation matrices and the TL technique. They used five pretrain DL models, for feature extraction and classification. A 98.96% accuracy rate for BC detection was demonstrated for the VGG16 model. A thorough comparison of several strategies based on accuracy. Although accuracy is the most often used parameter in all pertinent research, it is listed as a performance parameter in Table 7. As far as we currently know, suggested DL method outperforms all recently published state-of-the-art methods. The capability of the proposed methodology to extract more stable and distinctive deep features for classification helps it get the best outcomes. Additionally, we employed a balanced combine dataset. In study [33] they

also used combine dataset, however our method outperformed them in terms of accuracy. The comparison findings showed that the proposed approach was more effective than these techniques. Additionally, categorization, which requires more complicated computing, required the usage of hand-crafted engineering. Additionally, the suggested model outperformed the stated techniques in terms of accuracy, as shown in Table 7.

Table 7. Comparison with state of the art work.

Work	Method	Accuracy
Ragb et al. [32]	DL + TL	99.5%
Khan et al. [34]	DL GoogleNet + VGG	97.52%
Saber et al. [40]	DL + TL	98.96%
Proposed Method	DCNN	99.35%

4.5. Performance Evaluation on Binary Dataset

We validated the model using another standard, publically available dataset [47] that comprises 250 BC pictures, 100 benign and 150 malignant, in order to further examine the performance and generalizability of the proposed customized framework and other pretrain models. We used 70% of around 175 of the images for training and the remaining 30% for testing. Table 6 presents the performance of all pretrain models and proposed model for BC detection and classification and shows that all the models achieved good results except AlexNet. The results of DL models were analyzed and assessed using four assessment metrics, i.e., accuracy, precision, recall and F1-Score. It can be witnessed that the ResNet-18 achieves the highest accuracy of 99.50% and the Alexnet achieves the lowest accuracy of 97.40%. While the suggested model achieves the second-highest Accuracy of 99.63%. The F1-Score provides an overall assessment of the classifier's robustness. The greater the score of F1-sore, the better the classifier's performance. The maximum F1-score value is 100% for the ResNet-50 model, while the lowest F1-score value is 97.30%percent for the AlexNet model. Table 8 represent the experimental results on binary BC dataset. Additionally in future study, By training and evaluating the suggested technique on tasks such as identifying COVID-19 from chest radiograph images, brain Tumours detection, mask removal, pest detection, and forecasting heart diseases, we want to find and replicate its performance.

Table 8. Experimental results on binary BC dataset.

Model	Accu	Prec	Recall	F1-Score
Proposed Model	99.63%	99.50%	100%	99.50%
AlexNet	97.40%	97%	97.30%	97.30%
GoogLeNet	99.26%	99.20%	98.76%	99.50%
Shufflenet	99.26%	99.20%	98.70%	99.50%
ResNet-50	99%	99%	99%	99%
SqueezeNet	99%	99%	99%	99%
ResNet-18	99.50%	99.50%	99.33%	99.30%
Resnet101	99%	99%	98.33%	98.33%
ExceptionNet	99.30%	99.30%	99%	99%

5. Conclusions

This paper presents a novel DeepBraestCancerNet deep learning model for BC detection and classification. The proposed framework has 24 layers, including six convolutional layers, nine inception modules and one fully connected layer. We observed that the proposed model reached the highest classification accuracy of 99.35%. Furthermore, we compared the performance of several Deep Learning models, and the experiment results showed that our model outperformed the others. Furthermore, we validated the proposed model using another standard, publically available dataset. The proposed model reached

the highest accuracy of 99.35%. Moreover, this paper utilizes eight pre-train DL for BC detection using the TL technique. We evaluated the performance of nine DL-based models on a standard Dataset consisting of 537 benign, 360 malignant 133 normal ultrasound images. The limited images in the publicly available BC ultrasound imaging dataset, which impact DL models' performance, are a limitation of this research. This research can be further improved by including more images in the dataset. Furthermore, we validated the proposed model using another standard, publically available binary dataset. The proposed model reached the highest accuracy of 99.63%.

Additionally, as the suggested model extracts increasingly in-depth, accurate, and discriminatory features, testing with the dataset reveals that there are very few images of malignant BC and a considerable number of images of normal Breast. Therefore, an effective segmentation approach should be applied to the breast cancer data before classifying the breast ultrasound images into two classes, namely malignant and normal. The segmented images may then be used by the proposed algorithm to accurately detect and identify normal and cancerous images. Furthermore, future research might be designed to answer clinically relevant inquiries. Moreover, In future we will explore models like ViT for BC detection. The success of improved DL algorithms could benefit radiologists and oncologists in accurately detecting BC from MRI and CT scans. However, the findings reported in this paper can assist professionals in making the right decision for their models, obviating the necessity for a thorough search. Using deep neural networks solves various problems connected with model training. It allows us to create effective models for BC diagnosis, which helps in early detection and care.

Author Contributions: A.R. and N.U. developed the method; A.R., N.U., M.A. and J.A.K. performed the experiments and analysis, and J.A.K., A.G. and H.A. wrote the paper. All authors have read and agreed to the published version of the manuscript.

Funding: Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2023R54), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Vikas, C.; Pal, S.; Tiwari, B.B. Prediction of benign and malignant BC using data mining techniques. *J. Algorithms Comput. Technol.* **2018**, *12*, 119–126.
2. Halim, A.; Ashraf, A.; Andrew, A.M.; Yasin, M.N.M.; Rahman, M.A.A.; Jusoh, M.; Veeraperumal, V.; Rahim, H.A.; Illahi, U.; Karim, M.K.A.; et al. Existing and Emerging Breast Cancer Detection Technologies and Its Challenges: A Review. *Appl. Sci.* **2021**, *11*, 10753. [[CrossRef](#)]
3. Roslidar; Rahman, A.; Muharar, R.; Syahputra, M.R.; Arnia, F.; Syukri, M.; Pradhan, B.; Munadi, K. A review on recent progress in thermal imaging and DL approaches for BC detection. *IEEE Access* **2020**, *8*, 116176–116194. [[CrossRef](#)]
4. Singh, S.P.; Urooj, S.; Lay-Ekuakille, A. BC detection using PCPCET and ADEWNN: A geometric invariant approach to medical X-ray image sensors. *IEEE Sens. J.* **2016**, *16*, 4847–4855. [[CrossRef](#)]
5. Guo, R.; Lu, G.; Qin, B.; Fei, B. Ultrasound imaging technologies for BC detection and management: A review. *Ultrasound Med. Biol.* **2018**, *44*, 37–70. [[CrossRef](#)]
6. Nover, A.B.; Jagtap, S.; Anjum, W.; Yegingil, H.; Shih, W.Y.; Shih, W.-H.; Brooks, A.D. Modern BC detection: A technological review. *Int. J. Biomed. Imaging* **2009**, *2009*, 902326. [[CrossRef](#)]
7. Naeem, U.; Ali Khan, J.; El-Sappagh, S.; El-Rashidy, N.; Khan, M.S. A Holistic Approach to Identify and Classify COVID-19 from Chest Radiographs, ECG, and CT-Scan Images Using ShuffleNet Convolutional Neural Network. *Diagnostics* **2023**, *13*, 162.
8. Naeem, U.; Raza, A.; Ali Khan, J.; Ali Khan, A. An Effective Approach for Automatic COVID-19 Detection from Multiple Image Sources Using Shufflenet Convolutional Neural Network (CNN). *Res. Sq.* **2022**.
9. Chahal, P.K.; Pandey, S.; Goel, S. A survey on brain tumor detection techniques for MR images. *Multimed. Tools Appl.* **2020**, *79*, 21771–21814. [[CrossRef](#)]

10. Zhou, T.; Qi, S.; Wang, W.; Shen, J.; Zhu, S.-C. Cascaded parsing of human-object interaction recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *44*, 2827–2840. [[CrossRef](#)]
11. Syeda, I.H.; Alam, M.M.; Illahi, U.; Su, M.M. Advance control strategies using image processing, UAV and AI in agriculture: A review. *World J. Eng.* **2021**, *18*, 579–589. [[CrossRef](#)]
12. Byra, M.; Jarosik, P.; Szubert, A.; Galperin, M.; Ojeda-Fournier, H.; Olson, L.; O’Boyle, M.; Comstock, C.; Andre, M. Breast mass segmentation in ultrasound with selective kernel U-Net convolutional neural network. *Biomed. Signal Process. Control* **2020**, *61*, 102027. [[CrossRef](#)] [[PubMed](#)]
13. Naeem, U.; Ali Khan, J.; Alharbi, L.A.; Raza, A.; Khan, W.; Ahmad, I. An Efficient Approach for Crops Pests Recognition and Classification Based on Novel DeepPestNet Deep Learning Model. *IEEE Access* **2022**, *10*, 73019–73032.
14. Asaf, R.; Ayub, H.; Ali Khan, J.; Ahmad, I.; Salama, A.S.; Daradkeh, Y.I.; Javeed, D.; Ur Rehman, A.; Hamam, H. A hybrid deep learning-based approach for brain tumor classification. *Electronics* **2022**, *11*, 1146.
15. Gao, H.; Lin, M. Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017.
16. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
17. Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J.; Wojna, Z. Rethinking the inception architecture for computer vision. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 2818–2826.
18. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 1–9.
19. Zhang, X.; Zhou, X.; Lin, M.; Sun, J. ShuffleNet: An Extremely Efficient Convolutional Neural Network for Mobile Devices. *arXiv* **2017**, arXiv:1707.01083v2.
20. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. In Proceedings of the 26th Annual Conference on Neural Information Processing Systems NIPS, Lake Tahoe, NV, USA, 3–6 December 2012; pp. 1106–1114.
21. Iandola, A.; Bernstein, D. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5 MB model size. *arXiv* **2016**, arXiv:1602.07360.
22. Chollet, F. Xception: DL with depthwise separable convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 1251–1258.
23. Mohammed, S.A.; Darrab, S.; Noaman, S.A.; Saake, G. Analysis of BC detection using different machine learning techniques. In Proceedings of the International Conference on Data Mining and Big Data, Singapore, 11 July 2020; pp. 108–117.
24. Alanazi, S.A.; Kamruzzaman, M.M.; Islam Sarker, M.N.; Alruwaili, M.; Alhwaiti, Y.; Alshammari, N.; Siddiqi, M.H. Boosting BC detection using convolutional neural network. *J. Healthc. Eng.* **2021**, *2021*, 5528622. [[CrossRef](#)] [[PubMed](#)]
25. Hussain, L.; Aziz, W.; Saeed, S.; Rathore, S.; Rafique, M. Automated BC detection using machine learning techniques by extracting different feature extracting strategies. In Proceedings of the 2018 17th IEEE International Conference on Trust, Security and Privacy in Computing and Communications/12th IEEE International Conference on Big Data Science and Engineering (TrustCom/BigDataSE), New York, NY, USA, 1–3 August 2018; pp. 327–331.
26. Goyal, N.; Trivedi, M.C. BC classification and identification using machine learning approaches. *Mater. Today Proc.* **2020**. [[CrossRef](#)]
27. Jain, T.; Verma, V.K.; Agarwal, M.; Yadav, A.; Jain, A. Supervised machine learning approach for the prediction of BC. In Proceedings of the 2020 International Conference on System, Computation, Automation and Networking (ICSCAN), Pondicherry, India, 3–4 July 2020; pp. 1–6.
28. Zerouaoui, H.; Idri, A. Deep hybrid architectures for binary classification of medical BC images. *Biomed. Signal Process. Control* **2022**, *71*, 103226. [[CrossRef](#)]
29. Toğaçar, M.; Özkurt, K.B.; Ergen, B.; Cömert, Z. BreastNet: A novel convolutional neural network model through histopathological images for the diagnosis of BC. *Phys. A Stat. Mech. Its Appl.* **2020**, *545*, 123592. [[CrossRef](#)]
30. Ting, F.F.; Tan, Y.J.; Sim, K.S. Convolutional neural network improvement for BC classification. *Expert Syst. Appl.* **2019**, *120*, 103–115. [[CrossRef](#)]
31. Nahid, A.-A.; Mehrabi, M.A.; Kong, Y. Histopathological BC image classification by deep neural network techniques guided by local clustering. *BioMed Res. Int.* **2018**, *2018*, 2362108. [[CrossRef](#)] [[PubMed](#)]
32. Ragb, H.; Ali, R.; Jera, E.; Buaossa, N. Convolutional neural network based on TL for BC screening. *arXiv* **2021**, arXiv:2112.11629.
33. Bevilacqua, V.; Brunetti, A.; Triggiani, M.; Magaletti, D.; Telegrafo, M.; Moschetta, M. An Optimized Feed-forward Artificial Neural Network Topology to Support Radiologists in Breast Lesions Classification. In Proceedings of the 2016 Genetic Evolutionary Computation Conference Companion-GECCO ’16 Companion, London, UK, 7–11 July 2016; pp. 1385–1392.
34. Khan, S.; Islam, N.; Jan, Z.; Din, I.U.; Rodrigues, J.J.C. A novel DL based framework for the detection and classification of BC using TL. *Pattern Recognit. Lett.* **2019**, *125*, 1–6. [[CrossRef](#)]
35. Tang, J.; Rangayyan, R.M.; Xu, J.; El Naqa, I.; Yang, Y. Computer-aided detection and diagnosis of BC with mammography: Recent advances. *IEEE Trans. Inf. Technol. Biomed.* **2009**, *13*, 236–251. [[CrossRef](#)] [[PubMed](#)]

36. Agaba, A.J.; Abdullahi, M.; Junaidu, S.B.; Hayatu, H.I.; Chiroma, H. Improved Multi-Classification of BC Histopathological Images using Handcrafted Features and Deep Neural Network (Dense layer). *Intell. Syst. Appl.* **2022**, *6*, 200066.
37. Sharma, S.; Kumar, S. The Xception model: A potential feature extractor in BC histology images classification. *ICT Express* **2021**.
38. Inan, M.S.K.; Alam, F.I.; Hasan, R. Deep integrated pipeline of segmentation guided classification of BC from ultrasound images. *Biomed. Signal Process. Control* **2022**, *75*, 103553. [[CrossRef](#)]
39. Jebarani, P.E.; Umadevi, N.; Dang, H.; Pomplun, M. A Novel Hybrid K-Means and GMM Machine Learning Model for BC Detection. *IEEE Access* **2021**, *9*, 146153–146162. [[CrossRef](#)]
40. Saber, A.; Sakr, M.; Abo-Seida, O.M.; Keshk, A.; Chen, H. A novel deep-learning model for automatic detection and classification of BC using the transfer-learning technique. *IEEE Access* **2021**, *9*, 71194–71209. [[CrossRef](#)]
41. Eroğlu, Y.; Yildirim, M.; Çinar, A. Convolutional Neural Networks based classification of breast ultrasonography images by hybrid method with respect to benign, malignant, and normal using mRMR. *Comput. Biol. Med.* **2021**, *133*, 104407. [[CrossRef](#)] [[PubMed](#)]
42. Chowdhury, D.; Das, A.; Dey, A.; Sarkar, S.; Dwivedi, A.D.; Mukkamala, R.R.; Murmu, L. ABCanDroid: A cloud integrated android app for noninvasive early breast cancer detection using transfer learning. *Sensors* **2022**, *22*, 832. [[CrossRef](#)]
43. Moon, W.K.; Lee, Y.-W.; Ke, H.-H.; Lee, S.H.; Huang, C.-S.; Chang, R.-F. Computer-aided diagnosis of breast ultrasound images using ensemble learning from convolutional neural networks. *Comput. Methods Programs Biomed.* **2020**, *190*, 105361. [[CrossRef](#)] [[PubMed](#)]
44. Xiao, T.; Liu, L.; Li, K.; Qin, W.; Yu, S.; Li, Z. Comparison of transferred deep neural networks in ultrasonic breast masses discrimination. *Biomed Res Int.* **2018**, *2018*, 11. [[CrossRef](#)]
45. Mambou, S.J.; Maresova, P.; Krejcar, O.; Selamat, A.; Kuca, K. Breast cancer detection using infrared thermal imaging and a deep learning model. *Sensors* **2018**, *18*, 2799. [[CrossRef](#)] [[PubMed](#)]
46. Al-Dhabyani, W.; Gomaa, M.; Khaled, H.; Fahmy, A. Dataset of breast ultrasound images. *Data Brief* **2020**, *28*, 104863. [[CrossRef](#)]
47. Paulo, S.R. Breast ultrasound Image. *Mendeley Data* **2017**. [[CrossRef](#)]
48. Mampitiya, L.; Rathnayake, N.; De Silva, S. Efficient and Low-Cost Skin Cancer Detection System Implementation with a Comparative Study Between Traditional and CNN-Based Models. *J. Comput. Cogn. Eng.* **2022**. [[CrossRef](#)]
49. Zheng, M.; Zhi, K.; Zeng, J.; Tian, C.; You, L. A Hybrid CNN for Image Denoising. *J. Artif. Intell. Technol.* **2022**, *2*, 93–99. [[CrossRef](#)]
50. Zhang, X.; Wang, G. Stud pose detection based on photometric stereo and lightweight YOLOv4. *J. Artif. Intell. Technol.* **2022**, *2*, 32–37. [[CrossRef](#)]
51. Masood, F.; Masood, J.; Zahir, H.; Driss, K.; Mehmood, N.; Farooq, H. Novel Approach to Evaluate Classification Algorithms and Feature Selection Filter Algorithms using Medical Data. *J. Comput. Cogn. Eng.* **2022**. [[CrossRef](#)]
52. LeCun, Y.; Bengio, Y.; Hinton, G. Nature. *Deep Learn.* **2015**, *521*, 436–444. [[CrossRef](#)] [[PubMed](#)]
53. Jordan, M.I.; Mitchell, T.M. Machine learning: Trends, perspectives, and prospects. *Science* **2015**, *349*, 255–260. [[CrossRef](#)] [[PubMed](#)]
54. Pan, S.J.; Yang, Q. A survey on TL. *IEEE Trans. Knowl. Data Eng.* **2009**, *22*, 1345–1359. [[CrossRef](#)]
55. Russakovsky, O.; Deng, J.; Su, H.; Krause, J.; Satheesh, S.; Ma, S.; Huang, Z.; Karpathy, A.; Khosla, A.; Bernstein, M.; et al. ImageNet large scale visual recognition challenge. *Int. J. Comput. Vis.* **2015**, *115*, 211–252. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.