DOI: 10.15849/IJASCA.231130.12

Int. J. Advance Soft Compu. Appl, Vol. 15, No. 3, November 2023
Print ISSN: 2710-1274, Online ISSN: 2074-8523

Breast cancer diagnosis with an ensemble deep neural network

Nourane Laaffat[1], Ahmad Outfarouin[2], Walid Bouarifi[1], and Abdelilah Jraifi[1]

Mathematics Computer Science and Communication Systems Research Laboratory, Cadi Ayyad University, Safi, Morocco
e-mail: nourane.laaffat@gmail.com
Management and Decision Support Laboratory, Ibn Zohr University, Dakhla, Morocco
e-mail: ah.outfarouin@gmail.com

Abstract

Breast cancer (BC) is highly prevalent and lethal among women. Mammographic imaging enables accurate visualization and early detection of abnormalities in breasts. An efficient classification technique that categorizes images based on the stage of pathology is crucial in analysing breast abnormalities. Deep convolutional neural networks (DCNNs) have shown promising results in medical image classification and computer-aided diagnosis (CADx). Recently, deep transfer learning (TL) approaches are most widely used to detect and classify BC. In this article, we propose an approach based on TL and ensemble learning techniques to classify BC using mammographic images from the Mini-MIAS dataset. Our approach adopts the pre-trained models and employs the model averaging of their fine-tuned architectures for classification. It provides a very efficient and robust classifier that achieves an accuracy, a sensitivity and a specificity of 100%.

Keywords: *Breast Cancer, Mammographic Images, Classification, DCNNs, Transfer Learning, Stacking.*

1 Introduction

1.1 Breast Cancer

Breast cancer was initially given attention and importance due to the fact that it is the most prevalent and typical form of cancer in the world and that, if discovered at an extremely early stage, a very high cure rate is attainable [1,2,3,4,5]. We are focusing on addressing this issue to safeguard the lives of numerous individuals affected by this illness. In 2020, a total of 2.3 million women worldwide were diagnosed with breast cancer, resulting in 685,000 deaths, according to the World Health Organization [6]. Starting in puberty, BC may strike women everywhere in the world. Compared to other cancer types, it is the one where women experience a greater loss of life years. The presence of aberrant cells that proliferate uncontrolled is referred to as cancer. BC cells may remain in the breast or travel to other parts of the body via blood or lymphatic channels. The majority of the time [6], BC progresses over several months or even years. BC occurs in 15% of the cells in the glandular tissue lobules of the breast and 85% of the cells that line the breast ducts.

Received 15 September 2023; Accepted 2 November 2023

Initially, a cancerous tumor remains in the original duct or lobule, usually without causing any symptoms and being unlikely to spread (metastasize).

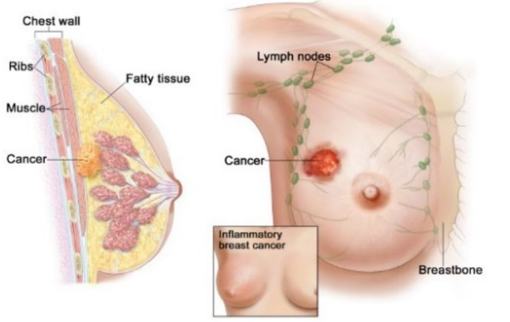

Figure 1. Breast Cancer.

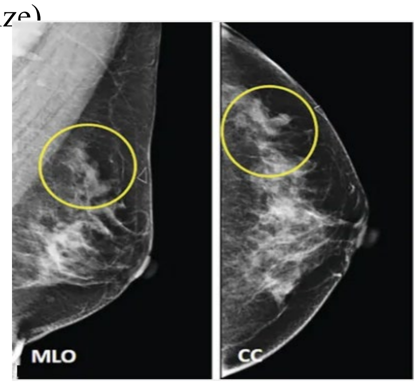

Figure 2. Different incidences taken by mammography of the left breast.

A woman may notice a lump in her breasts, so she thinks about going to the hospital. First, the doctor asks the patient a series of questions to distinguish whether the lump is benign or malignant. First, he asks questions about the development of the lump: When did the patient feel it? Is the duration short or long? and other questions, such as: Is the lump painful or not? animated or not? accompanied by bulging under the armpit or not? Does the patient have fluid discharge from the nipple or not? After the step of asking the patient questions, comes the step of asking the patient to do a medical examination, through mammography for example, because of its great importance in early detection.

1.2 Mammography Images

Mammography [7,8,9] is a medical imaging procedure based on X-rays, particularly intended for the examination of the women's breasts. Its goal is to find anomalies earlier, before they start to manifest as clinical symptoms. It offers details on the anatomy, diseases, and breast morphology. It is particularly useful for both the evaluation of breast mass lesions and the identification and diagnosis of BC. Early identification of BC is a crucial step that enhances the effectiveness of the disease's treatment, saving the lives of affected individuals and enabling longer, normal lives [10]. The mammographic image is created as a result of the attenuation of an X-ray beam passing through the various breast tissues [11,3], which mostly depends on the composition of the tissues passed through. The breast consists mainly of fat, milk-producing lobules and ducts that bring the milk to the nipple. As the fat has a very light physical density, it is regarded as a radiolucent region. This causes it to seem quite black on a mammogram. On the other hand, calcium, which is a crucial component of breast lesions, is present in the radio-opaque patches, which seem clear and correlate to fibro glandular tissue [11]. Based on anatomy and radiolucency information, it can be established that a mammogram generally appears dark while areas with microcalcifications or masses (composed of calcium) are lighter. Considering the intricacy of breast anatomy [11,3], mammography is a procedure that is applied in different directions or incidences. A good impact is made to stretch the breast tissue out as much as possible on the x-ray plate in order to reveal as much of the breast tissue as feasible. Different incidences are adopted, depending on the area of the breast being examined. The most often utilized incidences are the frontal incidence, also known as Cranio Caudale (CC), the external oblique incidence, also known as Medio Lateral Oblique (MLO), and the profile incidence (Figure 2).

The advantage of such an examination is that all of the breast tissue may be examined with just one or two instances.

1.3 Importance and Necessity of Computer Aided Diagnosis (CADx) Systems

It necessitates several interventions by radiologists, a lot of effort and a lot of time to evaluate the massive number of mammographic images that need to be done each year. This is especially true given that the analysis of these images is crucial to a woman's survival or lack thereof. In order to achieve this, a number of research projects have been focused on the construction of systems for the automating the reading of mammographic images. Such systems were first developed to give radiologists a second interpretation to aid in the early detection and diagnosis of suspicious lesions, independent of their type. They are called Computer Aided Detection/Diagnosis (CAD) systems. A CADx system is used to describe a comprehensive mammographic image processing system, from pre-processing to categorization. The success of such systems is due to their speed, consistency, and ability to provide reliable solutions to assist the breast lesion detection stage or alternatively the identification stage.

In our work, we are interested in (CADx) to better diagnose breast cancer. The (CADx) is dedicated mainly to the analysis of mammographic images, which is a sequence of phases that must be performed one after the other, from image acquisition to decision making. Some of these phases are often closely related and inseparable. The stages involved in processing a mammogram can be summed up as follows:

- A pre-processing step that enhances the image’s quality before any manipulations.
- A segmentation phase that enables the identification of the lesion to be studied.
- A description phase that aims to describe mathematically the lesions.
- A classification and decision-making step using a suitable classifier.

A (CADx) often seeks to describe and categorize the discovered anomaly as benign or malignant, abnormal or normal. There are many algorithms for classification and prediction of breast cancer outcomes. In DL, classification entails directly extracting image characteristics from raw images through the adjustment of the convolution and pooling layer parameters. Due to the appearance of TL technique and the rise in image production, which has improved the performance rates of classification models, DL models have achieved significant advancements in classification. Although it is challenging to obtain huge datasets, the lack of medical images has been one issue that TL approaches have been able to solve.

DCNNs are subsequently unable to more effectively train small datasets and become involved in the overfitting phenomena.

The pre-trained DCNNs are used in various ways by TL. Either to maintain the important knowledge after feature extraction and to transfer it to a classifier for use. Either to modify it specifically to get better outcomes.

2 Related Work

Particularly with TL approaches and the growth in data that increases the performance of classification models, the classification of medical images has seen considerable development in the prevention of diseases like BC, skin disorders, pneumonia, etc. In the context of the classification of mammographic images, several studies have been conducted in the literature.

In [12], authors provide a breast lesion detector and classifier. The YOLO detector is used for the detection phase, which yields F1 scores for the DDSM and INbreast datasets of 99.2% and 98.02%, respectively. Regular feedforward CNN, ResNet-50, and InceptionResNet-V2 classifiers are used during the classification phase. The accuracy of CNN classification models for the DDSM and Inbreast datasets, respectively, is 94.5% and 88.7%, ResNet-50 is 95.8% and 92.5%, and InceptionResNetV2 is 97.5% and 95.3%. The work [13] entails comparing the performance of VGG16 and ResNet50 on the IRMA dataset to find the top breast cancer classifier. Tt is VGG16 that has an accuracy rate of 94% compared to ResNet50's 91.7%.

The authors in [14] propose a breast cancer classification system based on CNN. They apply AlexNet and GoogleNet on four mammography databases CBIS-DDSM, INbreast, MIAS and images from the Egyptian National Cancer Institute (NCI) to find the best classifier, based on the evaluation metrics of performance. AlexNet achieves accuracies of 100%, 100%, 97.89% with an AUC of 98.32% and 98.53% with an AUC of 98.95%, respectively on the CBIS-DDSM, INbreast, NCI and MIAS datasets. While GoogleNet achieved accuracies of 98.46%, 92.5%, 91.58% with an AUC of 96.5% and 88.24% with an AUC of 94.65%, respectively on the CBIS-DDSM datasets, INbreast, NCI and MIAS. AlexNet emerges as the most effective classifier.

In [15], the authors propose a fresh approach to the segmentation and categorization of breast cancer. Using MIAS, DDSM, and CBIS-DDSM datasets, they performed the classification using the InceptionV3, DenseNet121, ResNet50, VGG16, and MobileNetV2 models. Using the DDSM dataset, the suggested model with the Inception v3 classifier reaches the best results; an accuracy of 98.87%, AUC of 98.88%, sensitivity of 98.98%, precision of 98.79%, F1-score of 97.99%, and calculation time of 1.2134 s. Nourane et al. [4] utilize the pre-trained DCNN models DenseNet121, InceptionResNetV2, and MobileNetv2 on Mini-Mias dataset to contribute to a breast cancer classifier. The models are employed in two ways: The first is as a feature extractor, then the phase of classification is carried out by Machine learning classifiers (SVM and Random Forest), and the second is for fine-tuning. The hyperparameters employed during training in the second approach are learning rate (0,0005), epochs (10), and batch size (32). The fine-tuned MobileNetV2 succeeds in achieving the greatest accuracy of 97%. Nourane et al [5] conduct a comparative performance study between two approaches based on the VGG-19 model using Mini-Mias dataset, in order to find the most performant breast cancer classifier. The classification in the first approach involves fine-tuning VGG-19 with hyperparameters adjustment, while in the second approach it involves using the machine learning classifier Random Forest after carrying out the feature extraction by VGG-19 model. The VGG-19 performs as the top classifier in the first way; its accuracy is 97 percent, compared to 91% in the second.

S. j. malebary and a. hashmi [16] have devised a new breast mass classification system called BMC. It is based on the development of an architecture that combines k-mean clustering, CNN, random forest, LSTM-RNN, and boosting techniques. They performed the classification into benign, malignant, and normal on MIAS and DDSM datasets. The sensitivity, specificity, F-measure and accuracy achieved by the BMC system for the DDSM dataset are 0.97%, 0.98%, 0.97%, 0.96% with an AUC rate between 0.94% and 0.97% and for the MIAS dataset 0.97%, 0.97%, 0.98%, 0.95% respectively, with an AUC rate between 0.94% and 0.98%. These values are calculated for each training dataset produced utilizing the various patient counts, such as 20, 30, 40, 50, 60, 70, 80, and 100 from the aforementioned datasets.

Imran et al [17] suggest a novel DCNN strategy based on feature fusion and ensemble learning techniques to enhance the identification and categorization of anomalies in mammographic images. After the steps of pre-processing and ROI extraction, the feature extraction and classification are performed by a proposed CNN architecture. It consists of four blocks. Convolution, max-pooling, and dropout layers are used in the first three blocks to extract features, while the fourth one has a flattening layer coupled to three blocks (sigmoid, SVM, RF).

The connection of the first blocks with each individual block in the fourth one forms a sub-network. A majority voting is then applied to find the final prediction. The proposed model achieves in the Mias dataset a sensitivity, specificity, accuracy of 99,5%, 99,4%, 99,4% respectively.

3 Proposed methodology

We present our suggested system in five steps and Figure 3 (Fig.3) shows its architecture.

The first step concerns the fixation/choice of the database of mammographic images (scans of the woman's breast) on which our approach will focus, as well as the data preparation. In our case, we used the Mini-Mias dataset which has been freely available for scientific study. Then we prepare our data by submitting our mammographic images to the Data Augmentation technique which concerns the data preprocessing to enhance the size and quality of the dataset.

The second entails dividing the dataset into three parts (training, testing and validation). The third phase entails the adoption of the TL technique of 4 pre-trained models; VGG-16, Xception, MobileNetV2, NasNetLarge. We present these models thereafter. These are models already pre-trained on the large ImageNet dataset, and the use of TL involves using their weights in order

to improve the accuracy rate and achieve efficient decision-making. Here, we adopt their fine-tuned architecture. Throughout the training phase, only the weights of the last fully connected (FC) layer of each model were adjusted and updated. The pretrained models were modified by removing the classification part. To make this adjustment, the models utilized three hidden layers with 1024, 256, and 28 neurons respectively. Since the mammography classification problem is binary (normal/abnormal), the model was designed with two neurons in its output layer. To combat overfitting, a dropout layer was added after the FC layers, and the "Softmax" activation function was employed in the output layer.

In step 4, we apply the model averaging ensemble technique. It is a reduce form of the stacking technique used in machine learning to combine the predictions of several individual models to obtain a more accurate and robust final prediction, instead of relying only on the predictions of a single model.

The predictions of our pre-trained models are thus grouped together; the average of the predictions is calculated to obtain the final prediction. This reduces bias and variance and improves generalization and the overall performance of the model by combining the strengths of the different adopted models.

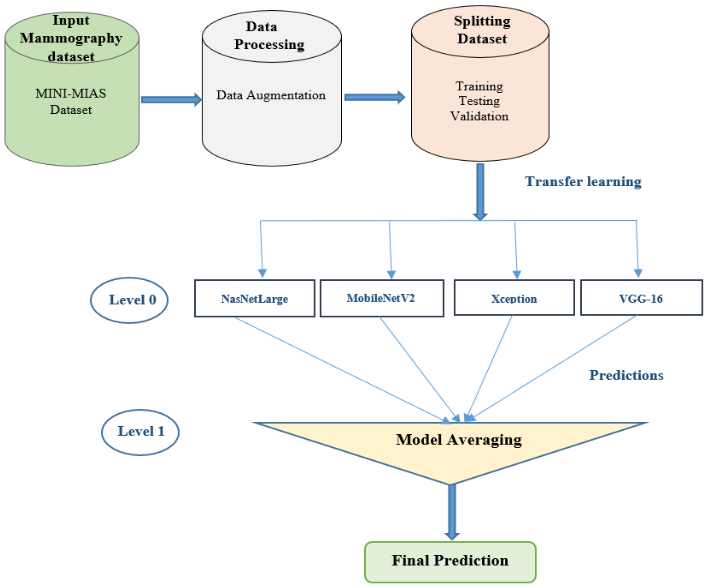

Figure 3. Proposed methodology.

3.1 Dataset and Data Preparation

We make use of the Mini-MIAS dataset [18], which Essex's Pilot European Image Processing Archive University has made freely available for use in scientific research. It was compiled and produced by a team of UK researchers at the Mammographic Image Analysis Society (MIAS), interested in deciphering breast lesions and mammogram interpretation. The dataset comprises of 322 mammographic images with mediolateral oblique views and 1024*1024 pixel resolution, 209 of which are normal and 113 of which have anomalies (abnormal: benign or malignant).

Due of our dataset's modest size, pre-processing is a must in order to reduce overfitting and enhance the suggested model's capacity for learning. We employ the technique of data augmentation for this.

3.2 Data Augmentation

One of the most popular techniques for reducing the overfitting issue and enhancing the performance of DL models that have already been trained on a sizable dataset (ImageNet), and work quite well on it, is data augmentation. It is based on artificially increasing the size of the dataset, especially those of the training set, by creating modified versions of the images which in our case concern mammographic imaging. It consists of geometric image transformation techniques, such as random image rotation, horizontal and vertical random flipping, zooming and shearing.

3.3 CNN Models

CNN models are currently the most powerful models for classifying medical images and diagnosing diseases. the idea of their architecture is inspired by the vertebrate visual cortex's behavior [19]-[20]. It is formed by a stack of three main layers: The convolutional layer, the pooling layer and the fully connected (FC) layer. The task of the convolutions and pooling layers is to learn how to extract the crucial and significant characteristics that will be used in the classification stage carried out in the last step by the fully connected layer, through the application of the kernel (filter) that convolves the image. Dropout and normalization layers which represent the secondary layers play an important role in overcoming overfitting during training and thus improve the performance of CNN models.

We sum up the various types of DCNNs that we have used in our work. These are pre-trained models on ImageNet dataset to recover the knowledge acquired there as their weight to solve a close problem, which is in our case the classification of medical images. And in order to generate the predicted probabilities to classify the mammographic images, we implement the Softmax function at the FC level.

Xception: Xception is a DCNN that was proposed in 2017 by François Chollet [21], as an extreme extended version of the inception model. It consists in introducing new layers of inception created by depth-wise convolutions which are more efficient in terms of computation time instead of classical convolutions. This model defaults to an input image size of 299x299 and has a depth of 126. Among the layers constituting this network, 36 are convolutions responsible for feature extraction and the FC layers are replaced by a global average pooling layer for the reduction of the number of parameters. And to produce predictions, the activation function Softmax is then adopted.

VGG16: The VGG16 is a DCNN introduced in 2014 by K. Simonyan and A. Zisserman from Visual Geometry group laboratory of Oxford University. This network admits as input images of 224*224 format, it has a depth of 16 [22] layers of weight and applies filters of 3*3 in all convolutional layers to decrease the parameters number. It has proven an excellent performance in the classification of images from the ImageNet dataset as well as in the field of image recognition.

MobileNetV2: The MobileNetV2 is a model whose architecture is proposed by Google and belonging to the family of MobileNet-type models which are designed to operate on devices with high performance constraints such as smartphones and mobile devices in general [23]. The MobileNet architecture adopts depth-wise separable convolutions to reduce the number of trainable parameters, while relying on a technique that reduces the cost of computation and which consists of the separation of the kernel into two smaller kernels; both of which are in charge of convolution, one for depth-wise and the other for point-wise. The input images format of this model is 224*224.

NasNetLarge: NasNet, which stands for Neural Architecture Search, uses reinforcement learning to find the best CNN design. The NasNet architecture is made up of convolutional cells and is based on two main functions: The normal cell and the reduction cell. But since on large datasets, training is costly in terms of resources and intensiveness, the search for an architectural block is conducted first on a tiny dataset, and then transferred to a larger one utilizing the NASNet search space. The NasNetLarge model adopts the ScheduledDropPath regularization technique which greatly enhances generalization of NASNet models.

3.4 Transfer Learning

The input images collection in medical imaging is frequently modest, including mammographic images. It is thus difficult to obtain a large dataset. Therefore, it is strongly recommended not to train the neural network using a random initialization: The risk of overfitting is enormous since the number of parameters to learn is significantly bigger than the number of images. The TL technique avoids all these problems. Its principle is based on the ability to apply knowledge gained by a neural network when addressing a problem in order to solve a more or less similar one [24]. This achieves a transfer of knowledge, hence the name. This requires having a neural network previously trained (pre-trained), ideally on a problem related to the one to tackle. Nowadays, we can easily get one from the Internet, and particularly from DL libraries, such as Keras.

The pretrained DCNN models can be used in two ways: The first is the TL; this involves using the pre-trained models to extract the features; thus, the relevant knowledge is utilized by another model in charge of the classification. The second is Fine-tuning, here it is a more sophisticated technique that gives more efficient results; it consists of making specific modifications to the pre-trained model.

3.5 Ensemble Learning

Ensemble learning [25] is combining models together for the image classification task instead of creating a single model and trying to improve the accuracy of that model. It is always better to create multiple models which can extract different features because each model has its own advantage and disadvantages. So, we use multiple models and try to combine the predictions of those models. For this, there are different techniques to combine the predictions (different ways to stack the models), either we can take a simple average or we can take weighted average. Our approach is based on using a simple prediction technique by taking an average of multiple models. This is the model averaging: A stacked generalization ensemble learning approach. It employs numerous sub-models that each make an equal contribution to the final prediction.

4 Experimental results and Discussion

To judge the performance of our proposed model, we evaluated the performance of all the models on which our approach is based and then compare them with ours, then we conducted a comparison of our model with other existing approaches carried out by other authors.

4.1 Experimental Specification

To perform our experiment, we have used Google Notebook Python Colab. We adopted as optimization algorithm the Adam optimizer, with a learning rate of 0.00005 for the compilation of the models. For their training, we worked with a count of 10 epochs and a batch size of 32. We evaluated our model’s classification performance by calculating certain evaluation metrics such as accuracy, recall, precision, F1-Score, sensitivity, specificity and AUC. The calculation of the evaluation metrics is done on the basis of the calculation of the values of True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN). These values are best presented in the confusion matrix, we clarify their meanings through the table 1.

Table 1 : The signification of TP, TN, FP, FN values

Prediction state	Predictive Class		Actual Class	
	Positive	Negative	Positive	Negative
TP	*		*	
TN		*		*
FP	*			*
FN		*	*	

Accuracy: It is the percentage of predictions that our model properly detected. The accuracy is defined as follows for binary classification: $Accuracy = \frac{\text{TP+TN}}{(\text{TP+TN+FP+FN})}$

Precision: It is the percentage of positive identifications that were actually correct. The precision is defined as follows: $Precision = \frac{\text{TP}}{(\text{TP+FP})}$

Recall: It is additionally known as sensitivity in binary classification. This measure describes the proportion of real positives that were accurately detected. It is defined as follows: $Recall = \frac{\text{TP}}{(\text{TP+FN})}$

F1-score: It is based on a formula of a harmonic mean of precision and recall. It has been designed to work well on imbalanced data. It is defined as follows:

$$F1 - score = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision+Recall}}$$

Specificity: It refers to the percentage of real negatives that were accurately detected. It is defined as follows: $Specificity = \frac{\text{TN}}{(\text{TN+FP})}$

Area Under Curve (AUC): It gauges how well a classifier can discriminate between different classes. The model performs better at differentiating between the positive and negative classes the higher the AUC.

4.2 Comparison of our Proposed Model with each used Pretrained Model

Table 2 presents the evaluation metrics used for our proposed model and for the different architectures on which our approach is based. We find that our proposed model outperforms every pre-trained model that goes into its construction. Our stacked generalized model based on the averaging model manages to reach an accuracy of 100%, and a value of 100% for each of the other metrics; recall, precision, f1-score and AUC.

Generally, fine-tuned DCNN models require less computation time compared to training a DCNN from scratch, thus they perform better and converge faster.

The architectures that we have adopted to constitute our proposed model, in order to classify our mammographic images have achieved good results, but the improvement of the performance metrics still remains a challenge. Thus, given the importance of diagnosing breast cancer early in a more efficient way, efforts can be made to improve the scores of the evaluation metrics that complement the accuracy metric, because it alone is not sufficient. Therefore, we propose a model based on the stacking generalization ensemble learning technique which takes into account the predictions of each model constituting the suggested model and thus considers each model's good results. The results show that our proposed model attains an accuracy of 100% which means that the predicted results are accurate, 100% for specificity, which means that there are no false positive predictions, 100% for sensitivity(recall), which means that there are no false negative predictions which justifies the reliability of the model, and 100% for the AUC, which means that the model has a great capacity to discern between the existing classes.

Table 2 : Comparative results of the suggested model with the fine-tuned architectures that constitute it

Approach	Accuracy %	Precision %	Recall %	F1-score %	Specificity %	AUC
MobileNetV2	97	91	100	95	95	95
VGG-16	97	91	100	95	95	95
Xception	97	91	100	95	95	95
NasNetLarge	94	100	82	90	100	91
Proposed Model	100	100	100	100	100	100

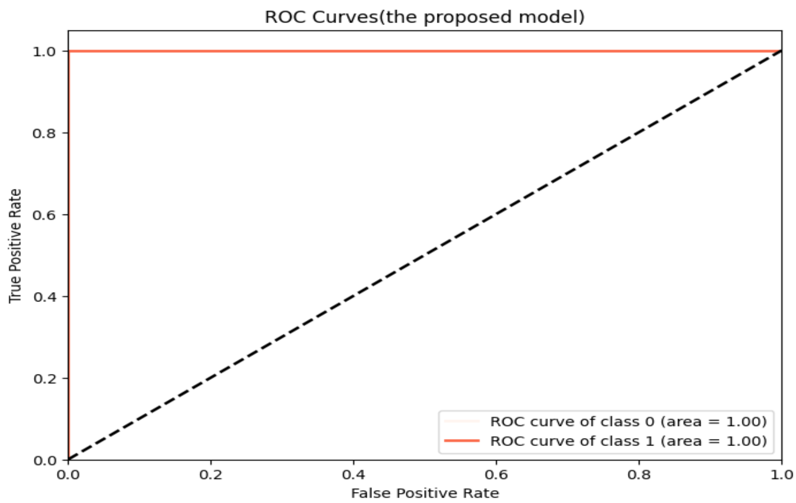

Figure 4. The ROC curves of the proposed model.

Figure 4 illustrates the ROC curve of our classification model. It is a graph that shows the relationship between the sensitivity rate (true positives) and the specificity rate (false positives) for different threshold values of the predictive model. The x-axis represents the specificity values, typically ranging from 0 to 1. The y-axis represents the sensitivity values, also typically ranging from 0 to 1. In our case, the ROC curve is represented by a diagonal line connecting the points (0,0) and (1,1). This diagonal line indicates that the model has perfect performance, where all true positives are identified without any false positives. This results in an AUC=1. Therefore, we can conclude that with an AUC of 1, the model has an excellent ability to discriminate between positive and negative classes. It has successfully separated positive observations from negative observations.

4.3 Comparison of our Proposed Model with other approaches

We examine the results of the suggested model with previous work performed on the Mini-Mias dataset. In Table 3, we sum up the performance evaluators values of the suggested model with those of other approaches. Our contribution achieved an accuracy of 100%, while Malebary et al. [16] achieved an accuracy of 95% using LSTM-RNN, CNN, random forest and boosting techniques. Imran et al. [17] proposed a model consisting of three sub-models based on a CNN architecture where the classification phase is carried out by three ways (sigmoid, SVM, RF), then they apply the technique of the majority vote on the 3 sub-models. The model achieves an accuracy of 99.4%. Hassan et al. [14] studied the performance of two pre-trained DCNN models; AlexNet and GoogleNet, using the TL technique and on several datasets. AlexNet proved to be the best performing model, and on MIAS it achieved an accuracy of 98.53% with an AUC of 98.95%. AlexNet proved to be the best performing model, and on MIAS it achieved an accuracy of 98.53% with an AUC of 98.95%. Thus, it can be said that our contribution provided more accurate findings compared to the other contributions in Table 3.

Table 3 : Comparative results of the proposed model with other approaches

Approach	Accuracy %	Precision %	Recall%	F1-score %	Specificity %	AUC
LSTM-RNN, CNN, random forest and boosting techniques [16]	95	-	97	98	97	94-98
CNN(convolution+max-pooling+dropout layers) +(sigmoid,SVM, RF), Majority vote [17]	99,4	-	99,5	-	99,4	-
AlexNet, transfer learning [14]	98.53%	-	-	-	-	98.95
Proposed model	100%	100%	100%	100%	100%	100%

Our mammographic image classification model based on fine-tuned CNN architectures and model averaging yields excellent results in all performance measures. The use of model averaging technique aids in diminishing the variance of individual models and attaining improved generalization. Through this technique, we mitigate the influence of random errors, and obtain a more stable estimate of the final prediction. As for the optimization parameters used, we adopted the Adam optimizer with a learning rate of 0.00005. The Adam optimizer is known to be effective in training deep neural networks as it automatically adapts the learning rate based on observed gradients. It combines the advantages of two other optimization algorithms, namely AdaGrad to adapt the learning rates to parameter weights, and RMSProp to maintain a moving average of gradients. The use of the Adam algorithm with a learning rate of 0.00005 contributed to the rapid convergence of our model and achieved high performance. We worked with 10 epochs and a batch size of 32. The number of epochs determines how many times the training dataset is traversed by the model, while the batch size determines the number of samples used to update the model weights at each step. These parameters need to be tuned to not be too low to avoid underfitting or too high to avoid overfitting. In our case, these values were well suited to our dataset, leading to perfect results.

5 Conclusion

BC is highly fatal for women and the timely diagnosis of this disease is extremely crucial and significant. Many efforts are being provided to contribute to good diagnosis (performance), such as the construction of computer-aided diagnostic systems for the classification of breast anomalies. In this paper, we present a DCNN model for the classification of BC using mammographic images. Our model uses the fine-tuning technique of pre-trained DCNN models based on the adjusting of hyperparameters, then it adopts the averaging model which is one of the approaches of the stacking generalization ensemble learning technique, to take into account the contribution of each used model (the average of the predictions obtained). We have evaluated our proposed model on the Mini-Mias dataset, we have adopted the data augmentation technique and applied a dropout layer to overcome the overfitting phenomenon. Finally, the evaluation results of our model give a value of 100% for the metrics used such as accuracy and precision and thus prove that data augmentation and stacking generalization are efficient ways to raise classification accuracy rates. For future work, we aim to evaluate our model on other breast cancer datasets and other diseases to be able to prove the performance of our model and its capacity in the diagnosis of breast cancer and other diseases.

References

[1] Kosmia, L., Rafaella, E., & Costas, P. (2023). Computer-aided breast cancer detection and classification in mammography: A comprehensive review. Computers in Biology and Medicine, 153(106554).

[2] Kavitha, T., Mathai, P.P., Karthikeyan, C., Ashok, M., Kohar, R., Avanija, J., & Neelakandan, S. (2022). Deep Learning Based Capsule Neural Network Model for Breast Cancer Diagnosis Using Mammogram Images. Interdisciplinary Sciences: Computational Life Sciences, 14(1), 113–129.

[3] Ichrak, K., & and Najlae, I. (2019,October). Breast cancer image segmentation and classification. In proceeding of the 4th International Conference on Smart City Applications (SCA'19), 2019 (pp. 1–9).

[4] Nourane, L., Ahmad, O., Walid, B., & Abdelali, J. (2022, March). A Deep Learning Model for Breast Cancer Diagnosis Using Mammography Images Classification. In proceeding of the 1st International Conference on Computing, Intelligence and Data Analytics (ICCIDA'22), 2022 (pp. 411-422).

[5] Nourane, L., Ahmad, O., Walid, B., & Abdelali, J. (2022, November). A Deep Convolutional Neural Networks for Detection of Breast Cancer Using Mammography Images. In proceeding of the 4th International Conference Artificial Intelligence and Smart Environment (ICAISE'22), 2022 (pp. 33-41).

[6] WHO Homepage, Breast Cancer, [Online], Available: https://www.who.int/news-room/factsheets/detail/breast-cancer.

[7] NCI (National Cancer Institute), Annual Report to the Nation: Cancer Death Rates Continue to Decline; Increase in Liver Cancer Deaths Cause For Concern, [Online] (2016), Available: https://www.cancer.gov/news-events/pressreleases/2016/annual-report-nation-1975-2012.

[8] Rampun, A., López-Linares, K., Morrow, P. J., Scotney, B. W., Wang, H., Ocaña, I. G., Maclair, G., Zwiggelaar, R., González Ballester, M. A., & Macía, I. (2019). Breast pectoral muscle segmentation in mammograms using a modified holistically-nested edge detection network. Medical Image Analysis, 57, 1-17.

[9] Punitha, S., Amuthan, A., & Joseph, K.S. (2018). Benign and malignant breast cancer segmentation using optimized region growing technique. Future Computing and Informatics Journal, 3(2), 348-358.

[10] American Cancer Society, Breast Cancer, [Online], Available: https://www.cancer.org/cancer/breast- cancer/screening-tests-and-early-detection/american-cancer- society-recommendations-for-the-early-detection-of-breast-cancer.html#:~:text=at%20any%20age.Mammograms,years%20before%20physical%20symptoms%20 develop.

[11] Imane, C. (2012). Description et classification des masses mammaires pour le diagnostic du cancer du sein. (Doctoral dissertation, Évry Val d'Essonne University).

[12] Mugahed, A.A., Seung-Moo, H., & Tae-Seong, K. (2020). Evaluation of deep learning detection and classification towards computer-aided diagnosis of breast lesions in digital X-ray mammograms. Computer Methods and Programs in Biomedicine, 196(105584).

[13] Nur Syahmi, I., & Cheab, S. (2019). Breast Cancer Detection Based on Deep Learning Technique. In Proceeding of International UNIMAS STEM 12th Engineering Conference (EnCon), 2019, pp. 89-92.

[14] Shayma'a, A. H., Mohammed, S. S., Mahmoud, I. A., & Mohsen, A. R. (2020). Breast cancer masses classification using deep convolutional neural networks and transfer learning. Multimedia Tools and Applications, 79, 30735-30768.

[15] Wessam, M. S., & Moustafa, H. A. (2021). Deep learning in mammography images segmentationnand classification: Automated CNN approach. Alexandria Engineering Journal, 60, 4701-4709.

[16] Sharaf, J. M., & Arshad, H. (2021). Automated Breast Mass Classification System Using Deep Learning and Ensemble Learning in Digital Mammogram. In IEEE Access, vol.9, pp. 55312-55328.

[17] Imran, L.H., Haidar, A., Hong Yu, W., Cui, L., & Hazrat, A. (2022). Feature fusion and Ensemble learning-based CNN Model for mammographic image classification. Journal of King Saud University - Computer and Information Sciences, 34, 3310-3318.

[18] University of Essex. The mini-mias database of mammograms. [Online], Available,

http://peipa.essex.ac.uk/info/mias.html.

[19] Majeed, T., Rashid, R., Ali, D., & Asaad, A. (2020). Problems of Deploying CNN Transfer Learning to Detect Covid-19 from Chest X-rays. MedRxiv.

[20] Jonathan, D., Alexandre, D., Matthieu, B., Benoit, E., Harvey, R., Dimitris, K. (2018). Deep Learning for Big Data Applications in CAD and PLM–Research Review, Opportunities and Case Study. Computers in Industry, 100, 227-243.

[21] François, C. (2017). Xception: Deep Learning with Depthwise Separable Convolutions. In proceeding of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017 (pp. 1800-1807).

[22] Dhananjay, T., & Sedamkar, R. (2020). Detecting Affect States Using VGG16, ResNet50 and SEResNet50 Networks. SN Computer Science, 1(79), 1-7.

[23] Mark, S., Andrew, H., Menglong, Z., Andrey, Z., & Liang-Chieh, C. (2018). Mobilenetv2: Inverted residuals and linear bottlenecks. In proceeding of the IEEE Conference on Computer Vision and Pattern Recognition, 2018 (pp. 4510-4520).

[24] Fuzhen, Z., Zhiyuan, Q., Keyu, D., Dongbo, X., Yongchun, Z., Hengshu, Z., Hui, X., & Qing, H. (2021). A Comprehensive Survey on Transfer Learning. In proceedings of the IEEE, 2021 (pp. 43–76).

[25] Ichraq, K., & Najlae, I. (2022). Cervical Cancer Detection and Classification Using MRIs. Jordanian Journal of Computers and Information Technology, 8(2), 141-158.