Sustainable Operations and Computers 3 (2022) 296–302

Contents lists available at ScienceDirect

Sustainable Operations and Computers

journal homepage: http://www.keaipublishing.com/en/journals/sustainable-operations-and-computers/

Evaluation of deep learning models for detecting breast cancer using histopathological mammograms Images

Subasish Mohapatra[a], Sarmistha Muduly[a], Subhadarshini Mohanty[a], J V R Ravindra[b], Sachi Nandan Mohanty[c,*]

[a] Department of Computer Science & Engineering, College of Engineering & Technology, Bhubaneswar, Odisha, India
[b] Department of Electronics and Communication Engineering, Vardhaman College of Engineering (Autonomous), Hyderabad, India
[c] Department of Computer Science & Engineering, Vardhaman College of Engineering (Autonomous), Hyderabad, India

ARTICLE INFO

Keywords:
Deep Learning
Deep Convolution Neural Network
Medical Imaging
Mammograms (MGs)

ABSTRACT

Breast cancer detection based on the deep learning approach has gained much interest among other conventional-based CAD systems as the conventional based CAD system's accuracy results seems to be inadequate. The convolution neural network, a deep learning approach, has emerged as the most promising technique for detecting cancer in mammograms. In this paper we delve into some of the CNN classifiers used to detect breast cancer by classifying mammogram images into benign, cancer, or normal class. Our study evaluated the performance of various CNN architectures such as AlexNet, VGG16, and ResNet50 by training some of them from scratch and some using transfer learning with pre-trained weights. The above model classifiers are trained and tested using mammogram images from the mini-DDSM dataset which is publicly available. The medical dataset contains limited samples of data due to low patient volume; this can lead to overfitting issue, so to overcome this limitation data augmentation process is applied. Rotation and zooming techniques are applied to increase the data volume. The validation strategy used here is 90:10 ratio. AlexNet showed an accuracy of 65 percent, whereas VGG16 and ResNet50 showed an accuracy of 65% and 61%, respectively when fine-tuned with pre-trained weights. VGG16 performed significantly worse when trained from scratch, whereas AlexNet outperformed others. VGG16 and ResNet50 performed well when transfer learning was applied.

1. Introduction

Breast Cancer can be viewed as common cancer in women and is the second most prime cause of death worldwide. It is caused by abnormal cell growth in the breast tissue, which results in the formation of a tumor and hence poses a severe risk to women's health and life. A lump in the breast, nipple discharge, and shape change of breast are all signs of breast cancer. A cancer is an uncontrollable growth of the cells which can spread to other parts of the body which carries life risk to the patient, so early detection is necessary to avoid for spreading further. By detecting lumps in their early stages, the mortality rate can be significantly reduced. Various medical imaging methods such as Magnetic Resonance Imaging (MRI), mammography, breast sonography, and magnetic resonance tomography (MRT) are widely used for breast cancer diagnosis [1].

Mammograms are the most preferred imaging methodology for screening early breast cancer. It helps in detecting suspicious lesions like masses and micro calcification. Due to its cost effectiveness and high sensitivity towards the minor lesion, mammograms are widely used for screening. It is regarded as the most reliable cancer detection method, as it has less radiation exposure than other alternatives available [2]. There are two significant findings seen in mammogram images, including masses and calcification. The calcification is distinguished as a coarser, granular, popcorn, or ring shape characteristics with higher density and more dispersed [3].

Mammograms show masses as medium grey or white regions in the breast, and their shapes can be oval, irregular, lobular, and with margin types that can be circumscribed, speculated, ill-defined, or obscured [4]. A mass can either be classified as benign or malignant. Benign tumors are usually distinguished as round or oval, whereas tumors having a partially round shape with an irregular outline are malignant, which is shown in Fig. 1, [5].

Architectural distortion is another indicator of breast cancer, but it is less significant as compared to others. The mammograms can be acquired in two different views (i) craniocaudal view and mediolateral oblique view, [6] as shown in Fig. 2.

During the actual diagnosis process, various factors such as image quality, radiologist expertise, the complexity of the breast structure af-

* Corresponding author.
E-mail address: sachinandan09@gmail.com (S.N. Mohanty).

https://doi.org/10.1016/j.susoc.2022.06.001
Received 15 March 2022; Received in revised form 27 April 2022; Accepted 1 June 2022
Available online 2 June 2022

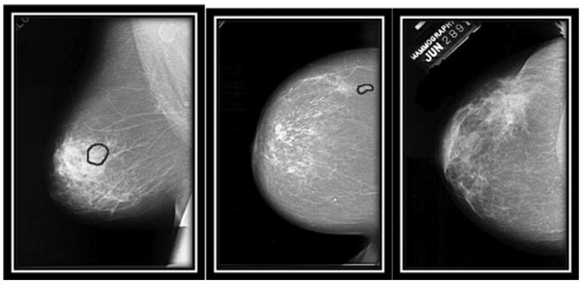

Fig. 1. Samples of the mini-DDSM dataset a) Cancer b) Benign c) Normal.

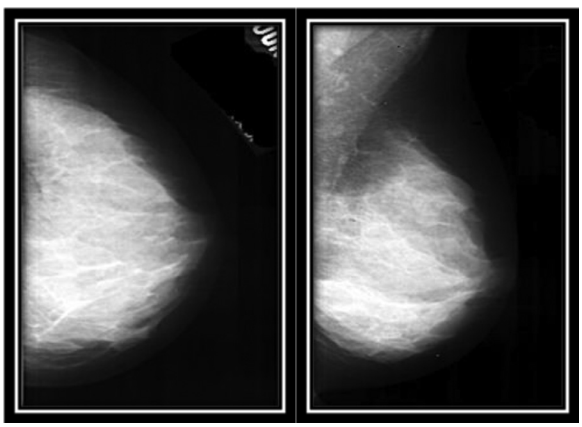

Fig. 2. a) Craniocaudal view of patient right breast b) Mediolateral oblique view of the patient right breast.

fects the cancer detection accuracy. To address this issue Computer-Aided Diagnosis (CAD) system comes into the picture and is very conventional approach [7]

Recently, Artificial Intelligence based CAD systems are used for providing better accuracy and early detection. This has opened many research directions for the researchers and developers to explore in deep learning methods for mammograms. In particular, Convolution Neural Networks (CNNs) have been used for lesion localization, detection, and classification task in mammograms [8]. It aids radiologists for accurate diagnosis and precision analysis of suspicious lesions. In recent years, DCNN has made remarkable advances in medical fields, most notably in image classification tasks; they play an essential role in improving Computer aided diagnosis (CAD) systems' performance for breast cancer diagnosis [9]. This DCNN classifier uses machine learning methods to analyze the mammogram images into different classes such as Benign, Malignant and Normal. Based on the classification done by CAD system the clinicians can predict the breast cancer in early stages which helps the patient for early diagnosis.

In view of the above, we conducted a comparative study of various CNN architectures' results in this paper. The performance of CNN models is compared using two methods: the first is to train the model from scratch, and the second is to train the model with pre-trained weights and from the experiments conducted it was concluded that when the classifiers are applied with transfer learning the accuracy achieved is better than trained from scratch.

The structure of the paper is categorized as follows: Section 2 details about the literature review of the work carried out by the researcher in the field of medical image classification using deep learning methods. Section 3 provides an insight about the motivation of study. Section 4 explains the methodology used to carry out the experiments. Section 5 describes about the simulation environment; Section 6 discusses about the results from the experiments; Section 7 provides the discussion of study and at last Section 8 gives the conclusion of the study and the future work in the respective field.

2. Literature review

In past few years, several researchers have performed classification of malignant and non-malignant classes of breast cancer by using different neural network classifiers.

Ragab DA et al. proposed a computer-aided design model which will classify benign and malignant tumors in mammogram images. It uses the segmentation approach in two ways: the first one physically determines the region of interest (ROI), and the second employs threshold and region-based strategies. A DCNN was applied for feature extraction and to improve the accuracy, the final fully connected layer was replaced with support vector machine. The segmentation procedure was applied to the DDSM dataset; the accuracy achieved for the first segmentation technique was higher than the subsequent one [2,24].

Hua Li et al. proposed a neural network model called Dense Net-II neural network which is improvised version of DenseNet model. It replaces first 3×3 convolution layer with inception layer. It overcomes the overfitting issue caused by deep and wider networks along with huge parameter configuration. The mammogram images are pre-processed, and then a data enhancement technique is applied to avoid the overfitting problem due to insufficient data. Next, the first convolution layers of the DenseNet is replaced by Inception Net. 10-fold cross-validation strategy was used for verifying the results. The pre-processed mammogram images are the input source for the following neural network, i.e., AlexNet, VGGNet, GoogleNet, DenseNet, and DenseNet-II models. [1,26] The results are compared, and it was shown that DenseNet-II neural network performed better than other models. The accuracy of the DenseNet-II network reached 94.55% (sensitivity 95.6%, specificity 95.36%). The dataset used was collected from First Hospital of Shanxi Medical University [12]. A total of 2042 mammogram images are obtained [3].

Dhungel N et al. presented a method that uses a cascade of CNN and a random forest classifier for detecting masses in mammograms. In first step multi-scale deep belief network was used to identify suspicious regions, which was then processed by a series of CNN, and a cascade of random forest classifiers for classification. IN breast and DDSM-BCRP datasets were used, and the sensitivity achieved for both cases is 85%-90%, respectively [4]

Phu T. Nguyen et al. [13] has performed the classification of breast cancer into benign and malignant using Break His dataset. They built a CNN model in which original images are resized and which is used to classify breast cancer classes. There are 7909 breast cancer images in the Break His dataset, categorized as benign or malignant from which 2440 images are in the benign category, and the remaining 5429 images are in the malignant category. It contains four subclasses under benign category and four subclasses under malignant category [10,11].

Shen Li et al. have created a deep learning algorithm based on a convolution neural network method to detect breast cancer using mammograms. Their findings demonstrated improved performance and accuracy on heterogeneous mammography platforms. They performed the experiments on digitized film mammograms from the CBIS-DDSM dataset. The author has performed the experiment by using the combination of VGG16 and ResNet50 and another an ensemble of four best model combination as classifiers [25,27]. The single Model has an AUC of 0.88 for every image four-model has AUC 0.91 with averaging and sensitivity and specificity at 86.1% and 80.1%, respectively. Similarly, for FFDM images of the INbreast database, the single Model has AUC of 0.95 for each image. The four-model averaging has AUC 0.98 with sensitivity: 86.7% and specificity: 96.1% [12,13].

Richa Agarwal et al. have used a patch-based CNN method for detecting masses automatically from full field digital mammograms (FFDM) dataset. [17] The CNN model is first trained with CBIS-DDSM and then tested on the INbreast dataset. Among other CNN architecture, Inception V3 obtains the best performance. The performance achieved due to transfer learning from the CBIS DDSM to the INbreast dataset showed a true positive rate of 0.98 [14].

Levy D et al. used techniques like transfer learning, pre-processing of data, and applying augmentation to train the CNN architecture from start to finish. In their study, three different CNN architectures have been trained, such as shallow CNN, AlexNet, and GoogleNet, among which GoogleNet showed better performance compared to others with an accuracy of 0.92. The Dataset DDSM was used for the experiment [15].

Huynh, Li et al. have done the classification methodology distinguishing between malignant and benign lesions. They have made a comparison between three methods. In first method, they have used pre-trained CNN features with SVM classifier, in the second method segmented tumor-based analytical method with SVM. In the third method, an ensemble classifier averaged between two individual classifier. Among the three methods, the performance of ensemble classifier showed significantly better performance with AUC (0.86) compared to the other two [16].

3. Motivation and objective

Radiologists miss approximately 20% of breast cancer cases due to extremely few non-discernible miniature calcification groups or masses. Due to the complexities of images by low contrast mammograms, it is difficult to diagnose breast cancer in an early stage by clinicians [3,9]. This led to the development of the (CAD)-based system. Convolution Neural Network showed superior performance on many image-classification tasks, and a few research studies have shown that CNNs can perform well on the mammogram classification.

An effective classifier for entire mammograms would offer numerous benefits which incorporates (i) Lot of work that can be saved by annotating mammograms (ii) Reducing the no of patient call-back rate (iii) Reducing false-positive cases and unnecessary follow-up test which overburdens the patients with increased health care costs.

This primary goal is to conduct a comparative analysis of various CNN architectures in the classification of benign and malignant tumors in mammogram images. This method consists of the following steps:

- Collection of mammogram images from a publicly available online database containing cancerous and noncancerous images.
- Developing an automatic detection mechanism to classify breast cancer masses and noncancerous masses from these mammogram images.
- To compare the results among all the CNN models for achieving better accuracy in classification.

4. Methodology

4.1. Dataset collection

The dataset used for experiment is the Mini-DDSM dataset which is publicly available online. It is a condensed version of the popular DDSM (Digital Database for Screening Mammography) data set, which is no longer in use. It includes half-resolution images from the DDSM dataset. The image files in the dataset are of PNG 16-bit format. It includes 9752 mammogram images that are classified as benign, cancerous, or normal. The dataset is divided into training and testing data in 90:10 ratios. The training dataset contains 8679 images with three classes, and the testing dataset contains 1073 images.

4.2. Image preprocessing

Preprocessing is very first and crucial step to enhance the performance of Computer aided diagnosis (CAD) system. Training the convolution neural network on raw images can lead to poor classification problems, so the input raw images need to transform to desired format before being fed to the deep learning classifiers. In our scenario, the medical dataset used contain images of different shape and sizes as compared to the images required by our network classifiers. The images fed should match the input size of the network classifier. To do so, the images need to resize or rescaled to the desired input type. This is an important step to speed up the training process.

4.2.1. Image resizing

The collected input images should be resized to match the required input size of CNN classifiers. When the required size exceeds the size of the input images being fed, less downsizing is required, resulting in less image deformation. On the other hand, a larger image occupies more memory space with a larger neural network. There are two approaches that can be used to accomplish this, one is cropping their border pixels, and the other one is scaling them down. Cropping the images around the border in the first approach may result in missing features and patterns around the border, whereas scaling down may result in deformed features and patterns across the images. Scaling is a reasonable choice compared to others as it is less risky than losing the patterns or features [17]. The images taken from mini-DDSM dataset is of varied sizes $346 \times 216 \times 3$, $365 \times 217 \times 3{,}360 \times 219 \times 3$ etc. which have been rescaled and resized to match the required input size of CNN classifiers. Image Generator method was used to perform the resizing and rescaling of the medical images. The Alex Net model classifiers require input images of size $227 \times 227 \times 3$, whereas the VGG16 and ResNet50 models both require input images of size $224 \times 224 \times 3$.

4.2.2. Data augmentation

To avoid overfitting, DCNN rely heavily on big data, but due to limited volume availability of medical imaging data DCNN suffers from overfitting. Overfitting is how a network model learns perfectly well on training data, but it fails miserably on test data. To avoid overfitting, data augmentation is applied to increase the number of mammogram images from the original dataset, due to limited volume of imaging data. In our approach, we are using augmentation based on geometric transformations which is quite simpler than the other forms. Here each image is augmented by using rotation, flipping, and zooming as they are simple to implement. As compared to flipping the vertical axis, horizontal axis flipping is more commonly used. Rotation augmentations can be performed by rotating the image to the right or left on an axis ranging from 1° to 359° [18]. To implement the augmentation process, we have used ImageGenerator method of the Keras library to generate batches of the images data in real time and this batch size is manually provided as the input parameter to the ImageGenerator method.

The ImageGenerator loads and augment images in batches for image classification tasks.

4.3. Convolution neural networks

Convolution neural network is a deep learning method which is very effective in classifying images, object detection and image recognition. It consists of three types of layers: convolution, pooling and fully connected layers. Convolution layer performs the extraction of the features from the input image by applying relevant filters, whereas the pooling layer reduces the dimensionality of extracted feature generated from the convolution layer and finally the last fully connected layer performs the classification of the features extracted from the series of convolution and pooling layer. AlexNet (Fig. 3), VGGNet (Fig 4), ResNet, are the well-known CNN classifiers used in our approach.

Alex Net architecture was the first one which outperformed a classification and detection task [19]. It was the winning entry in the 2012 ILSVRC. It comprises of eight weighted layers, the first five are convolutional layers, and the last three are fully connected layers. The last layer's output is fed into the SoftMax function, which produces 1000 class label [19].

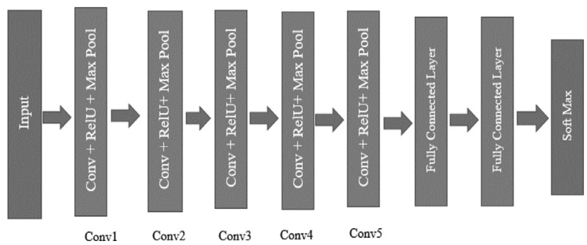

Fig. 3. AlexNet Architecture.

The first convolution layer has 96 kernels of size 11×11 and a stride of 4 pixels to filter the input image (224×224). The output from the first convolution layer's is the input to the second convolution layer, which has filter with 256 kernels of size $5 \times 5 \times 48$. The remaining three layers are linked to each other and in between there is no pooling or normalization layers. The third convolution layer contains 384 kernels of size $3 \times 3 \times 256$. The fourth layer contains 384 kernels of size $3 \times 3 \times 192$, and the fifth convolution layer has 256 kernels of size $3 \times 3 \times 192$. At last two fully connected layers with 4096 neurons each are connected. The number of neurons in the third fully connected layer corresponds to the number of classes that need to be classified [19,20]. The series of convolution, pooling and fully connected layer in Alex Net classifier extracts the features from the mammograms, reduces the dimensionality of extracted feature and finally does the classification of images into Benign, Malignant and Normal classes.

VGG16 is also a CNN architecture which is used for image classification task. It contains stack of convolution layer with Relu as an activation function and filters of size 3×3. The input to the convolution layer is of fixed size 224×224 RGB image, it is then passed to the series of convolution layer where filters is applied to the image for extracting features. The convolution stride is kept constant at one pixel, and the spatial padding of the convolution layer is kept one pixel for 3×3 convolution Layers. Five max-pooling layers are used to perform spatial pooling after few convolutional layers. This max-pooling layer calculates the largest value in each patch of each feature map results to highlight the most present feature in the patch. Next, there is stack of three fully connected layers connected to the previous layers. The first two FCs have 4096 channels each, while the third has 1000 channels for each class. For classification, the final layer includes a soft-max function [19].

ResNet50 network works on the principle of taking a deep Convolution Neural Network and adding a shortcut connection to skip few convolution layers at time. The shortcut connections create residual blocks as shown in Fig. 5. The output of the convolution layer is added to the residual block. The deep residual learning framework is designed to address the degradation issue caused by the deeper network. Th architecture is of 34 layer network in which skip connection or residual blocks are added resulting to a residual network [19].

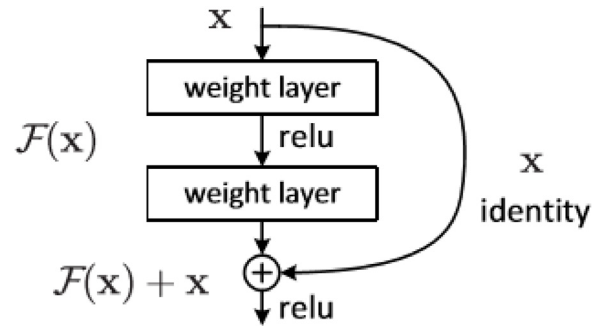

Fig. 5. Residual Learning: a building block [23].

4.4. Transfer Learning

It is a machine learning technique. In this case, the model developed for one task is being reused as the initial starting point for a model on a different task [21]. The basic idea is to apply the Model knowledge from a task with a large amount of available training data to a new task with much less data. Transfer learning is commonly used in fields such as computer vision and natural language processing that require a large amount of computational power [22]. It can be used as a feature extractor, and its parameters are fine-tuned to match the target dataset. Normally to train a deep model it requires days or month to complete, but in this scenario the computing time is reduced to hours when the model is moved to the target application. Currently, the most well-known object classification is based on ImageNet [23]. In this study, pre-trained CNN models have all indicated an initialization on ImageNet without any additional remarks.

5. Simulation environment

The goal of this study is to assess the performance of deep neural networks in classifying mammogram images as cancerous, benign, or normal. The Mini-DDSM dataset is used to evaluate the effectiveness of various CNN models. The experiments are conducted in Google Collaboratory, a free jupyter notebook that runs on the cloud platform. It includes a zero-configuration interface for writing and executing Python code directly from the browser, as well as free GPU access.

With Colab it gives the power of using popular python libraries that helps to analyse and visualize data. It has the capability to import image dataset, train classifiers and evaluates the performance of the classifiers. It executes the code in google cloud servers which helps to leverage the power of GPU and TPU, regardless of the power of user machine.

Some of the CNN models were trained from scratch during the experiment, while some of them were pre-trained using the ImageNet dataset. The SoftMax classification function is included in each network's fully connected layer. The optimization of the network is done by Adam opti-

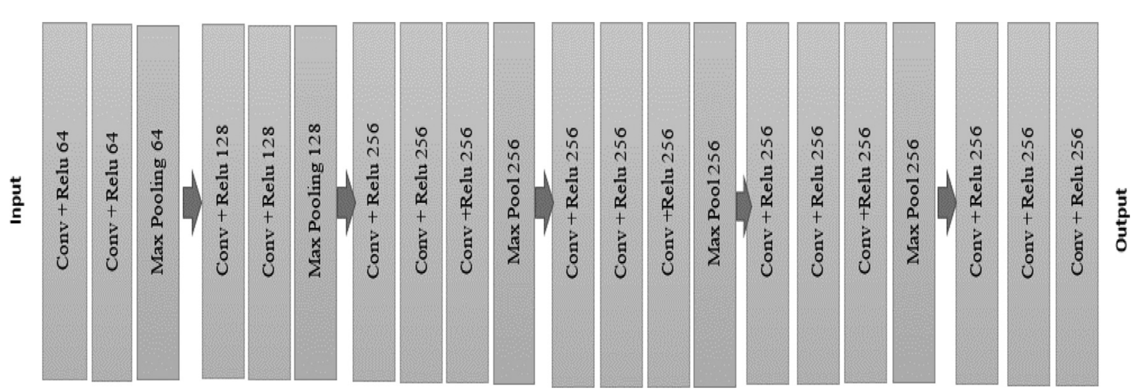

Fig. 4. VGG 16 Architecture.

Table 1
Classification Accuracy of network classifier trained from scratch.

CNN	Epoch	Optimizer	Batch Size	Learning Rate	Accuracy (%)
AlexNet	50	Adam	16	0.001	0.6589
	50	Adam	16	0.06	0.6384
	100	Adam	16	0.001	0.6589
	100	Adam	16	0.06	0.6384
VGG-16	50	Adam	16	0.001	0.312
	50	Adam	16	0.06	0.3756
	100	Adam	16	0.001	0.312
	100	Adam	16	0.06	0.312

mization method which helps to update the network weights efficiently. It is an extension to stochastic gradient which updates weight iteratively with training data. It uses Momentum and Adaptive learning rate to converge faster. Adam optimization is memory efficient and works well with large datasets and parameters. Next the batch size was kept 16 for all models to meet the memory constraints.

6. Experimental results

We have conducted two experiments to evaluate the performance of the classifiers. Experiment I involve training AlexNet and VGG 16 model classifiers from scratch using the mini-DDSM data set, with taking epoch, batch Size, Learning Rate and optimizer as hyperparameters. In Experiment II, the VGG16 and ResNet network classifiers were pretrained using the ImageNet dataset and then fine-tuned using the mini-ddsm dataset. The performance of the model classifiers is evaluated for better accuracy using mammograms in both experiments (Table 3).

Table 1 summarizes the classification accuracy of different network classifiers trained from scratch with different learning rates, which is iterated over 50 epochs and 100 epochs.

6.1. Experiment-I (Training with scratch)

6.2. Experiment II (Pre-trained)

The experiment II was conducted using VGG16 and ResNet architecture which was pretrained using ImageNet dataset and then fine-tuned using mini-ddsm dataset.

7. Discussion

Based on the above experimental results, we can conclude that network models pre-trained with a different dataset outperformed training from scratch, as evidenced by the VGG 16 results in both Table 1 and Table 2. When pre-trained with ImageNet dataset and then with mini-DDSM, the VGG 16 classifier performs significantly better. The accuracy was 0.65 when pretrained and 0.31 when trained from scratch. VGG16 fails miserably when trained from scratch and suffers from underfitting, which could be attributed to the complexity of the VGG 16 architecture [12,13]. However, Alex Net trained from scratch performed better (Fig. 6), with an accuracy of 0.65 (Table 4). In addition, in another experiment, VGG16 and ResNet were pretrained with ImageNet data and fine-tuned with mini-DDSM data. VGG16 performed better than ResNet in this scenario.

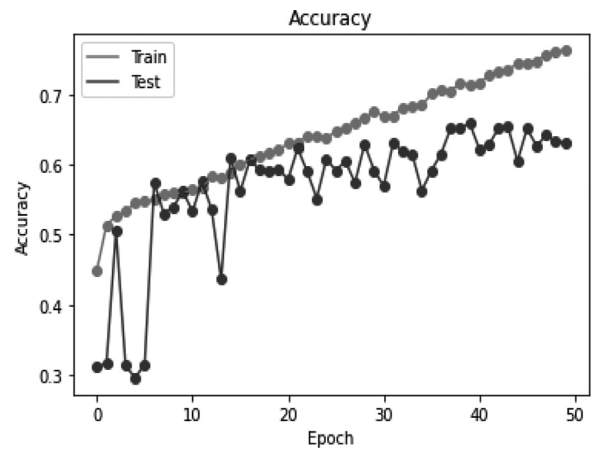

Fig. 6. Accuracy plot of Alex Net Model with lr 0.06.

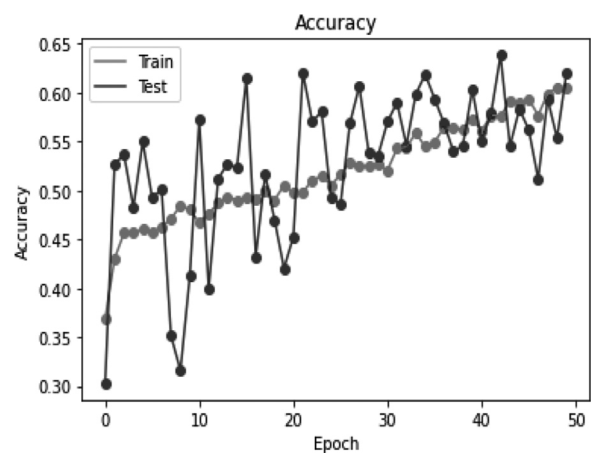

Fig. 7. Accuracy plot of Alex Net Model with lr 0.001.

8. Conclusion

The study concludes that when network classifiers are applied with transfer learning, they perform better than when trained from scratch, as seen in the case of VGG16. Due to the limited volume of imaging data, the network models suffer from low accuracy. Future work in this area will include using different neural network classifiers and combining them with other network algorithms. It would be fascinating to observe and evaluate the performance of the classifiers to improve their accuracy results. Aside from this, other validation split variations can be used to examine the effect on model performance (Figs. 3, 4, 6–13, Tables 3 and 4).

Table 2
Performance comparison of different classifiers.

CNN	Epoch	L.Rate	Specificity	Sensitivity	AUC	Accuracy (%)
AlexNet	50	0.001	0.9194	0.9646	0.8652	0.6589
	50	0.06	0.9133	0.9637	0.8490	0.6384
	100	0.001	0.9194	0.9646	0.8652	0.6589
	100	0.06	0.9133	0.9637	0.8490	0.6384
VGG-16	50	0.001	0.3439	0.3122	0.5000	0.312
	50	0.06	0.3439	0.3756	0.5317	0.3756
	100	0.001	0.3439	0.3122	0.5000	0.312
	100	0.06	0.3439	0.3756	0.5317	0.3756

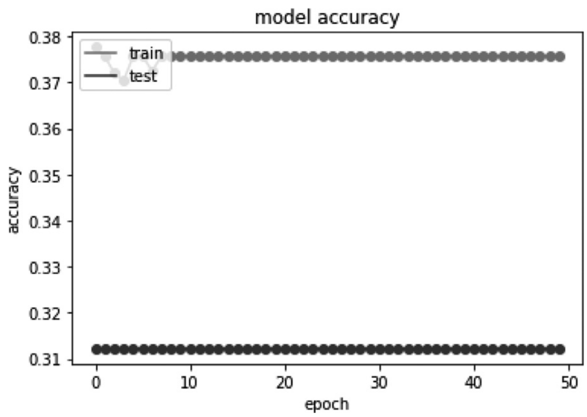

Fig. 8. Accuracy plot of VGG 16 model with lr 0.001.

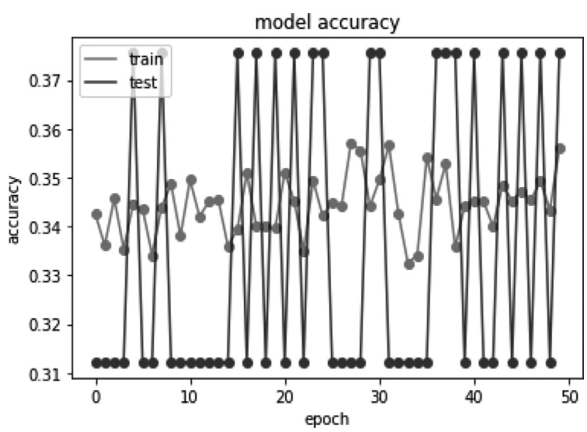

Fig. 9. Accuracy plot of VGG 16 Model with lr 0.06.

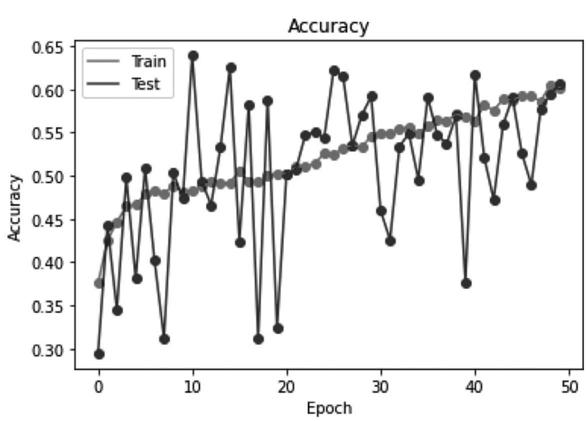

Fig. 10. Accuracy plot of VGG 16 (pre-trained) with lr 0.001.

Table 3
Classification Accuracy of network classifier using transfer learning and with different hyper parameters.

CNN	Epoch	Optimizer	Batch Size	Learning Rate	Accuracy (%)
VGG 16	50	Adam	16	0.001	0.6570
	50	Adam	16	0.06	0.6542
	100	Adam	16	0.001	0.6412
	100	Adam	16	0.06	0.6514
ResNet50	50	Adam	16	0.001	0.5927
	50	Adam	16	0.06	0.6198
	100	Adam	16	0.001	0.6086
	100	Adam	16	0.06	0.6226

Table 4
Performance comparison of VGG16 and ResNet with Pre-trained weights.

CNN	Epoch	L.Rate	Specificity	Sensitivity	AUC	Accuracy (%)
VGG-16	50	0.001	0.9035	0.9124	0.8295	0.6570
	50	0.06	0.8322	0.8257	0.7414	0.6542
	100	0.001	0.9035	0.9124	0.8295	0.6570
	100	0.06	0.8313	0.7623	0.7437	0.651
ResNet	50	0.001	0.8020	0.6785	0.6966	0.5927
	50	0.06	0.8164	0.7819	0.7173	0.6198
	100	0.001	0.8020	0.6785	0.6966	0.5927
	100	0.06	0.8243	0.8733	0.7240	0.6226

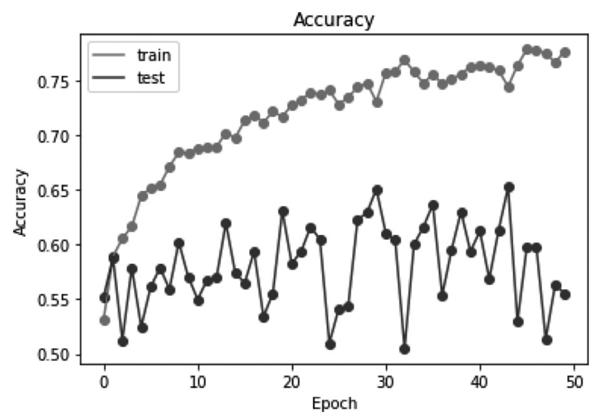

Fig. 11. Accuracy plot of VGG 16 (pre-trained) with lr 0.06.

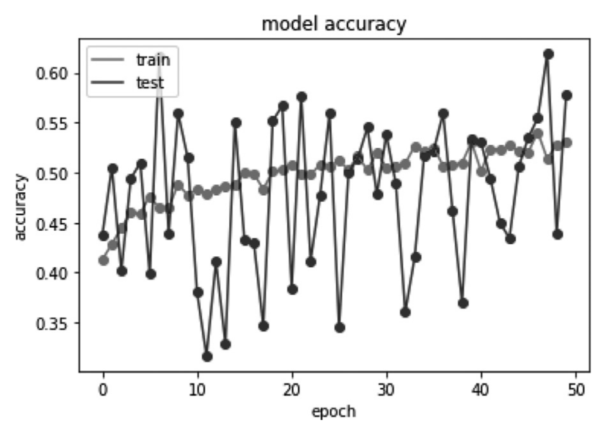

Fig. 12. Accuracy plot of ResNet (pre-trained) Model with lr 0.001.

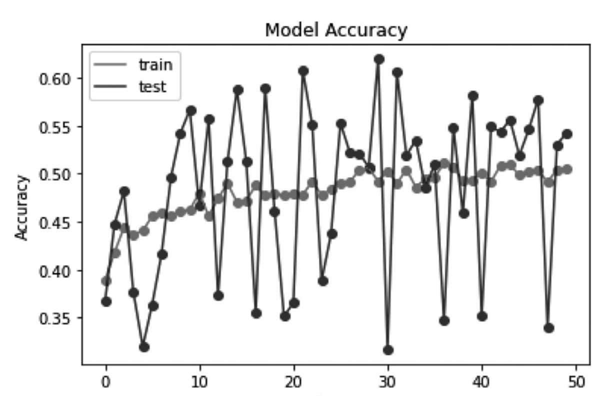

Fig. 13. Accuracy plot of ResNet (pre-trained) Model with lr 0.06.

Declaration of Competing Interest

There is no conflict among the authors in this manuscript.

References

[1] G.H. Aly, M. Marey, S.A. El-Sayed, M.F. Tolba, YOLO based breast masses detection and classification in full-field digital mammograms, Comput. Methods Programs Biomed. (2020) 105823, doi:10.1016/j.cmpb.2020.105823.

[2] DA Ragab, M Sharkas, S Marshall, J. Ren, Breast cancer detection using deep convolutional neural networks and support vector machines. machines, PeerJ. 7 (2019) e6201, doi:10.7717/peerj.6201.

[3] H. Li, S. Zhuang, D. Li, J. Zhao, Y. Ma, Benign and malignant classification of mammogram images based on deep learning, Biomed. Signal Process Control 51 (2019) 347–354, doi:10.1016/j.bspc.2019.02.017.

[4] N. Dhungel, G. Carneiro, A.P. Bradley, Automated mass detection in mammograms using cascaded deep learning and random forests, 2015 International Conference on Digital Image Computing: Techniques and Applications (DICTA), 2015, doi:10.1109/dicta.2015.7371234.

[5] J. Tang, R.M. Rangayyan, J. Xu, I. El Naqa, Y. Yang, Computer aided detection and diagnosis of breast cancer with mammography: recent advances, IEEE trans. Inf. Technol. Biomed. 13 (2) (2009) 236–251, doi:10.1109/TITB.2008.2009441.

[6] Z. Wang, M. Li, H. Wang, H. Jiang, Y. Yao, H. Zhang, J. Xin, Breast cancer detection using extreme learning machine based on feature fusion with CNN deep features, IEEE Access (2019) 1, doi:10.1109/access.2019.2892795.

[7] H.K. Tripathy, S. Mishra, P.K. MallickA.R. Panda (Ed.), Technical advancements of machine learning in healthcare, Stud. Comput. Intell. (2021), doi:10.1007/978-981-33-4698-7.

[8] SA Hassan, MS Sayed, MI Abdalla, MA Rashwan, Breast cancer masses classification using deep convolutional neural networks and transfer learning, Multimed. Tools Appl. 79 (2020) 30735–30768, doi:10.1007/s11042-020-09518-w.

[9] PT Nguyen, TT Nguyen, NC Nguyen, TT Le, Multiclass breast cancer classification using convolutional neural network, in: 2019 International Symposium on Electrical and Electronics Engineering (ISEE), 2019, pp. 130–134.

[10] Effects of Image augmentation and dual-layer transfer machine learning architecture on tumor classification, in: Proceedings of the 2019 8th International Conference on Computing and Pattern Recognition, 12 Oct 2021 Accessed, doi:10.1145/3373509.3373584.

[11] Shen L, Margolies LR, Rothstein JH, Fluder E, McBride R, Sieh W (2019) Deep learning to improve breast cancer detection on screening mammography. 10.1038/s41598-019-48995-4

[12] ON Oyelade, AE Ezugwu, A deep learning model using data augmentation for detection of architectural distortion in whole and patches of images, Biomed. Signal Process. Control 65 (2021) 102366, doi:10.1016/j.bspc.2020.102366.

[13] R Agarwal, O Diaz, X Lladó, MH Yap, R Martí, Automatic mass detection in mammograms using deep convolutional neural networks, JMI 6:031409 (2019), doi:10.1117/1.JMI.6.3.031409.

[14] Lévy D, Jain A (2016) Breast mass classification from mammograms using deep convolutional neural networks.

[15] BQ Huynh, H Li, ML Giger, Digital mammographic tumor classification using transfer learning from deep convolutional neural networks, JMI 3 (2016) 034501, doi:10.1117/1.JMI.3.3.034501.

[16] L Zou, S Yu, T Meng, Z Zhang, X Liang, Y Xie, A technical review of convolutional neural network-based mammographic breast cancer diagnosis, Comput. Math. Methods Med. (2019) 2019, doi:10.1155/2019/6509357.

[17] C Shorten, TM Khoshgoftaar, A survey on Image data augmentation for deep learning, J. Big Data 6 (2019) 1–48, doi:10.1186/s40537-019-0197-0.

[18] L Tsochatzidis, L Costaridou, I Pratikakis, Deep learning for breast cancer diagnosis from mammograms—A comparative study, J. Imaging 5 (2019) 37, doi:10.3390/jimaging5030037.

[19] SutskeverIlya KrizhevskyAlex, H E, ImageNet classification with deep convolutional neural networks, Commun. ACM (2017), doi:10.1145/3065386.

[20] M Maskey, R Ramachandran, JJ Miller, J Zhang, I Gurung, Earth science deep learning: applications and lessons learned, in: IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium, 2018, pp. 1760–1763.

[21] Transfer learning for deep learningIBM Developer, 12 Oct 2021 https://developer.ibm.com/articles/transfer-learning-for-deep-learning/ Accessed.

[22] SJS Gardezi, A Elazab, B Lei, T Wang, Breast cancer detection and diagnosis using mammographic data: systematic review, J. Med. Internet Res. 21 (2019) e14464, doi:10.2196/14464.

[23] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, doi:10.1109/cvpr.2016.90.

[24] Sksaran Munmin Swain, Jyotrimay Chatterjee, Azwean Abdhullah New Zaman, Sachi Nandan Mohanty, Hybridized machine learning based fractal analysis techniques for breast cancer classification, J. Adv. Comput. Sci. Appl. 11 (4) (2020) 234–245 (2020). ISSN 2158-107X.

[25] S.K. Lakshmanaprab, Sachi Nandan Mohanty, K. Shankar, N Arunkumar, Optimal deep learning model for classification of lung cancer on CT images, Future Gener. Comput. Syst. 92 (1) (2019) 374–382, doi:10.1016/j.future.2018.10.009, ISSN:0167-739X.

[26] Bibhu Prasad Sahu, Sachi Nandan Mohanty, Saroj Kumar Rout, A hybrid approach for breast cancer classification and diagnosis, EAI Endorsed Trans. Scalable Inf. Syst. 6 (20) (2019) 1–8.

[27] Bibhu Prasad Sahu, Sujata Das, Sachi Nandan Mohanty, Saroj Kumar Rout, Ensemble comparative study for diagnosis of breast cancer data set, Int. J. Eng. Technol. 7 (2018) 281–288 Issue (4.15.