Early Breast Cancer Detection Based on Locally Available Mammography Data Using Convolutional Neural Network

Eze C.E.[1], Azubogu A.C.O.[2], Akpado K.[3], Ifeyinwa Dimson[4], Nnebe S.U.[5], Ezeogu S.[6], Ibrahim N.B.[7]

[1,2,3,4,5] **Department of Electronic and Computer Engineering, Nnamdi Azikiwe University, Awka.**

[6] **uLesson Education Limited**

[7] **Outsource Global Technologies Limited**

ABSTRACT

In this work, convolutional neural network model was developed to read mammography image scans to predict malignancy. Firstly, mammography image data was sourced locally from University of Abuja Teaching hospital, Gwagwalada, Nigeria and then combined with publicly available mammography image dataset from The Mammographic Image Analysis Society (MIAS) database. The data obtained were then preprocessed using Contrast Limited Adaptive Histogram Equalization (CLAHE), Image Denoising, Padding and Formatting. Transfer learning was implemented by re-architecting pre-trained VGG-16, VGG-19, MobileNet-V2 and DenseNet-121 models to have an output layer with two neurons and softmax activation function, each signifying the degree to which calcifications in the mammography image is benign or malignant. The model was then trained on the preprocessed data and evaluated, the following metrics were achieved: VGG-16 and VGG-19 based models achieved an accuracy of 86% each, precision of 87% and 88% respectively and recall of 92% and 93% respectively. MobileNet-V2 has accuracy of 90%, precision and recall of 93% while the DenseNet-121 model achieved an accuracy of 95%, precision of 93% and 100% score in recall. MobileNet-V2 performed best in the computational complexity analysis with 336.34 MFLOPs, followed by DenseNet-121 with 5.69GFLOPs. VGG-16 and VGG-19 have computational complexity of 15.3 GFLOPs and 19.6 GFLOPs respectively.

INTRODUCTION

On an annual basis, Pathologists diagnose about 14 million new patients with cancer around the world. There were close to 20 million new cases of cancer in the year 2022 [1]. Recent research shows that most of the world's cancer cases are now in developing countries [2]. The incidence of preventable malignancies such as cervical, lung, colorectal cancer, etc. has decreased in developed countries, but remains unchanged in most developing countries. Approximately 70% of deaths from cancer occur in low-and middle-income countries [2]. In Nigeria, it accounts for 22.7% of all new cancer cases among women [3] Since 2007, breast cancer death rates have been steady in women younger than 50, but have continued to decrease in older women. From 2013 to 2018, the death rate decreased by 1% per year. These decreases are believed to be the result of finding breast cancer earlier through mammography screening [4].

Mammography is one of the most efficient techniques for the detection of early stage breast cancer. Mammograms are for the most part investigated by radiologists to distinguish the early stage of breast cancer [5]. The essential indications of anomalies and disease acquired from mammography are Calcifications, Localized increment in thickness, Masses, symmetry variation between the left and right breast images and Architectural distortion [6]. A mass can be malignant or benign. These masses can shape as an after effect of diverse interior techniques that influence the breast in distinctive ways. Malignant breast masses can either be kept to the ducts where they are framed or they can be invasive spreading through the channels to lymph hubs

and other inaccessible locales. Mammography is considered as the best strategy for right on time recognition of breast cancer, and the rate of patients that can be cured at ahead of schedule stages is generally high [7].

Recently, deep learning techniques, especially Convolutional Neural Networks (CNNs), have shown promising results in medical image analysis, including breast cancer detection. CNNs are a type of deep learning algorithm that can automatically learn and extract features from images, which can be used to accurately classify and diagnose breast cancer. CNNs can learn complex features from images and automatically detect patterns that may be difficult for human observers to recognize. CNN-based Computer Aided Detection (CAD) systems have demonstrated improved performance compared to human level accuracies in reading and classifying mammogram image scans.

However, the development of CNN-based CAD systems for breast cancer detection requires a large dataset of mammography images with corresponding ground truth labels. Majority of the studies conducted and models developed uses mammography data obtained from developed countries. This is largely due to non-availability of mammography image scan from developing countries in public domains like Kaggle, etc.

Therefore, this study aims to develop and evaluate a CNN-based model for early breast cancer detection using digital mammography images which was obtained from Nigerian mammography scan centers combined with publicly available datasets from developed countries. The study involved collection of dataset of digital mammography images from patients diagnosed with early-stage breast cancer, as well as healthy individuals with no known history of breast cancer. The dataset will be preprocessed and made available in the public domain for further use in research.

Review of Related Works

[8] Compared three of the most popular ML techniques commonly used for breast cancer detection and diagnosis, namely Support Vector Machine (SVM), Random Forest (RF) and Bayesian Networks (BN). The Wisconsin original breast cancer data set was used as a training set to evaluate and compare the performance of the three ML classifiers in terms of key parameters such as accuracy, recall, precision and area of ROC. The results obtained in this paper provide an overview of the state of art ML techniques for breast cancer detection

[9] Contributed to eliminate histopathological image dataset gap, by introducing a new publicly available image dataset named BreaKHis, which contains histopathological images of breast tumors. The work also presented an alternative approach to classify these challenging images, avoiding any explicit segmentation. Such approach explores hand-crafted textural descriptors and automatic representation, particularly using Convolutional Neural Networks as well as the paradigm Multiple Instance Learning. The obtained experimental results have demonstrated the feasibility of breast cancer detection using histopathological dataset, giving directions for improvement in such model.

[10] Proposed Multi-View Feature Fusion (MVFF) based CADx system using feature fusion technique of four views for classification of mammogram. The complete CADx tool contains three stages, the first stage have the ability to classify mammogram into abnormal or normal, second stage is about classification of mass or calcification and in the final stage classification of malignant or benign classification is performed. Convolutional Neural Network (CNN) based feature extraction models operate on each view separately. These extracted features were fused into one final layer for ultimate prediction. Our proposed system is trained on four views of mammograms, after data augmentation. We performed our experiments on publicly available datasets such as CBIS-DDSM (Curated Breast Imaging Subset of DDSM) and mini-MIAS database of mammograms. In comparison with literature the MVFF based system is performed better than a single view-based system for mammogram classification. We have achieved area under ROC curve (AUC) of 0.932 for mass and calcification and 0.84 for malignant and benign, which is higher than all single-view based systems. The value of AUC for normal and abnormal classification is 0.93.

In the work done by [11] transfer learning was implemented from pre-trained deep neural networks ResNet18, Inception-V3Net, and ShuffleNet in terms of binary classification and multiclass classification for breast cancer from histopathological images. They used transfer learning with the fine-tuned network which results in

much faster and less complicated training than a training network with randomly initialized weights from scratch. Their approach is applied to image-based breast cancer classification using histopathological images from public dataset BreakHis. The highest average accuracy achieved for binary classification of benign or malignant cases was 97.11% for ResNet 18, followed by 96.78% for ShuffleNet and 95.65% for Inception-V3Net. In terms of the multiclass classification of eight cancer classes, the average accuracies for pre-trained networks are as follows. ResNet18 achieved 94.17%, Inception-V3Net 92.76% and ShuffleNet 92.27% [12] Combined an ensemble of EficientNet-based classifiers with YOLOv5 and abnormality detection model, to identify abnormalities in mammography scans. The inclusion of YOLOv5 detection was to provide explanations for classifier predictions and improving sensitivity, particularly when the classifier fails to detect abnormalities. Abnormality detection model was incorporated to further enhance the screening process. They achieved an F1-score of 0.87 and a sensitivity of 0.82 but with the addition of suspicious mass detection, sensitivity increased to 0.89, albeit at the expense of a slightly lower F1-score of 0.79.

Summarily, several works have been done in the area of using machine learning techniques in analysis of medical images, and is still on going, especially in the area of breast cancer detection due to the need for a fast and reliable breast cancer detection techniques to circumvent the deaths that occurs due to breast cancer. However, very few of these works have considered using datasets from Nigeria. It is important that models are developed from a wide range of datasets so that the model can better generalize. Hence the need to build breast cancer models using data from developing countries such as Nigeria, having established that these developing countries have most breast cancer mortality rate.

Several breast cancer models have been developed using transfer learning performed on some popular models such as Google Inception-V3Net, ShuffleNet, ResNet18, etc. In this work, novel machine learning models was developed using transfer learning performed on VGG-16, VGG-19, MobileNetV2 and DenseNet121 architecture.

METHODS

In this section, the method adopted in the development of the model for classification of mammography images on breast cancer was shown in fig 1. The dataset is sourced from Mammography Image Analysis Society (MIAS) database and University of Abuja Teaching hospital, Gwagwalada, Abuja, Nigeria. The rest of the development cycle will take the following steps:
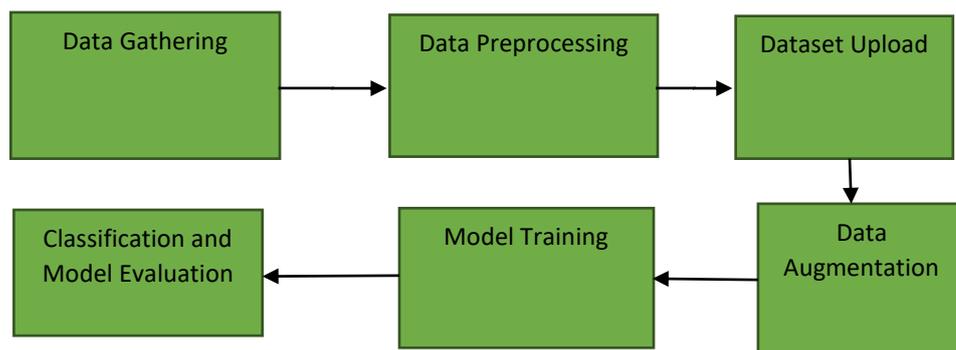
Fig 1: The model development steps taken in this thesis

As shown in Fig 1, the data obtained is then processed by improving the contrast of the images in a process known as contrast limited adaptive histogram equalization (CLAHE), then followed by reformatting, padding and denoising. The dataset is then uploaded to Kaggle under MIT license. Other mammography datasets are sourced from MIAS database which are then combined with the locally sourced dataset. Data augmentation is then carried out by random rotation between angles 2 degree to 10 degree, random flip, and random zoom. The dataset is then trained and evaluated by comparing with existing results.

Data Gathering

Dataset was obtained from The mammographic Image Analysis Society (MIAS) database of mammograms

[13]. It contains 322 images in Portable Gray Map (PGM) format. It was originally digitized at 50 micron resolutions, but now reduced to 200 micron per pixel edge and clipped/padded so that every image is 1024 by 1024 pixels. The MIAS database license allows the public to use the datasets for research purposes only.

The MIAS dataset have folders containing the images in PGM format and a text file containing the labeling of the images. Fig 2 shows the content of the text file already converted to a pandas dataframe in python.

refnum	bg	ab_class	severity	x	y	radius
mdb001	G	CIRC	B	535.0	425.0	197.0
mdb002	G	CIRC	B	522.0	280.0	69.0
mdb003	D	NORM	A	NaN	NaN	NaN
mdb004	D	NORM	A	NaN	NaN	NaN
mdb005	F	CIRC	B	477.0	133.0	30.0

Fig 2: First five rows of the MIAS dataset labeling as pandas dataframe

Column one is the MIAS database reference number which corresponds to the file name of the images in the dataset folder. Column two describes the character of the background tissue, it is either F, G or D representing Fatty, Fatty-glandular and Dense-glandular respectively. Third column represents the class of abnormality present, it could be either of CALC - Calcification, CIRC - Well defined/circumscribed masses, SPIC - Spiculated masses, MISC - Other ill-defined masses, ARCH - Architectural distortion, ASYM - Asymmetry or NORM - normal. Column 4 is the severity of the abnormality which serves as the label for this work. The severity is either B - benign or M - Malignant. A is used to denote the normal mammogram images. Finally, columns 5, 6 are the x and y coordinates of the center of abnormality respectively while column 7 is the approximate radius of a circle enclosing the abnormality.
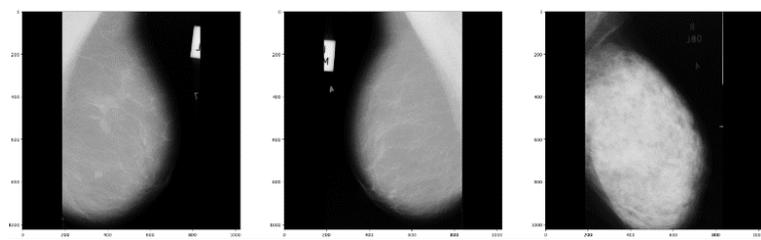
Fig 3: Random samples of MIAS breast scans plotted on Jupyter Lab software (plotted in pixels)

A total of 74 samples of image scans were obtained from University of Abuja Teaching hospital, Gwagwalada, Nigeria. This includes samples of 38 benign, 32 malignant and 4 normal mammogram images. These images were preprocessed and added to the images gotten from MIAS database.
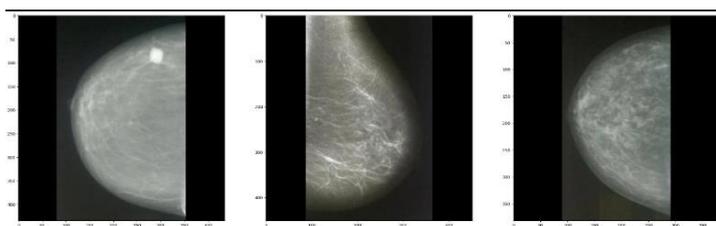
Fig 4: Random sample of mammogram images obtained from University of Abuja mammogram center (plotted in pixels)

Data Preprocessing

The obtained image data is in physical format; it is converted to digital image by way of scanning it using a

scanner. The scanning is preferred to using camera capturing since scanning would produce a cleaner image devoid of noise from reflections of light rays. The image is further denoised since there are several sources of disturbances that can affect mammography images as its been taken which mainly include beam hardening, patient movements, scanner malfunction, low resolution, intrinsic low dose radiation, and metal implants. The denoising of the image is achieved in this thesis using OpenCV image denoising library – cv2.fastNlMeansDenoisingColored().

Image Formatting and Padding

Since the format of the data obtained from the MIAS database is in Portable Gray Map (PGM), the scanned mammogram images in Portable Network Graphic (PNG) format had to go through some preprocessing to maintain the format and the naming convention used to store the MIAS images. First, the image had to go through padding to produce a squared image. The padding process attached the black colored background around the right and left sides of the images since they are already in portrait. The padded images are then resized to 1024 by 1024 pixels and then converted from the PNG format to PGM thereby maintaining the same shape and format as the MIAS image data. These were achieved by a python script.

The denoised image would then go through contrast-limited adaptive histogram equalization to give a detailed image of the mammography scan.

Histogram Equalization

Contrast-limited adaptive histogram equalization was achieved by first of all getting the histogram equalization of the images. Histogram equalization ensures that the images are sharpened so that the hidden details are immediately obvious. This it achieved by spreading the image frequency histogram across the possible gray scale levels, typically between the ranges of 0 to 255, from its initial clustered nature to a better even distribution across the range.

Mathematically, Histogram equalization is achieved thus:

Let m be a given image represented as a r x c matrix of integer pixel intensities.

The pixel intensities ranges from 0 to L-1, where L is the possible number of intensity values, in this case 256.

Let p represent the normalized histogram of image m, then;

$$P_n = \frac{number\ of\ pixel\ with\ intensity\ n}{total\ number\ of\ pixels} \qquad n = 0,1,2,...,L\text{-}1 \quad (1)$$

The histogram equalized image g will be given then as

$$g_{i,j} = floor((L\text{-}1)\sum_{n=0}^{i,j} P_n) \qquad (2)$$

$$T(k) = floor((L\text{-}1)\sum_{n=0}^{k} P_n)1 \qquad (3)$$

The motivation for this transformation comes from thinking of the intensities of m and g as continuous random variables X, Y on [0, L − 1] with Y defined by

$$Y = T(X) = (L-1)\int_0^y Px(X)dx \qquad (4)$$

$P_x(X)$ is the probability density function of m

$\int_0^y Px(X)dx$ is the cumulative distributive function (cdf)

$$Y = T(X) \qquad (5)$$

then, $Py(Y) = Px(X) |\frac{dX}{dY}|$ (6)

given that $Y = T(X) = (L - 1) \int_0^y Px(X)dx$ (7)

$\frac{dX}{dY} = (L-1) Px(X)$ (8)

$P_y(Y) = Px(X)|\frac{1}{(L-1) Px(X)}|$ (8) in (6) (9)

$P_y(Y) = \frac{1}{(L-1)}$ (10)

$T(k) = (\sum_{n=0}^{k} Pn * M)$ (11)

$g(k) = floor(\frac{cdf(k) - cdfmin}{M - cdfmin})x (L-1))$

where cdf = cumulative distribution function

Implementing Histogram Equalization using the openCV Library in Python.

OpenCv2 library was used for this purpose. The images is primarily in the Red Blue Green (RBG) format which is not best suited for histogram equalization since it would be required that histogram equalization is performed on the three channel of the image. Instead, the image is first converted to LAB (Lightening, A channel and B channel) format, where the L channel is the lightening channel of the image, A and B saves the color information of the image. Histogram Equalization is then performed on the L channel, after which the equalized L channel is merged back to the A and B channels. The image is then converted back to the RBG format.

Histogram equalization was performed on a sample of mammogram scan, and the result is shown in Fig 5.

The Histogram equalization technique fails when the input image has a large area low-intensity background. Such situation causes severe washing-out of images and thus effectively amplifying the noise in the image. To circumvent this, the technique of Contrast-Limited Adaptive Histogram Equalization is employed which is an improvement of Adaptive Histogram Equalization (AHE).

The image would also go through interpolation, registration, organ windowing followed by normalization, and zero-padding to improve the quality of training for deep learning algorithms. We will employ the use of openCV and Scipy libraries for interpolation, registration and windowing while normalization and zero padding will be achieved by the ScikitLearn libraries. The locally sourced and processed dataset were published to kaggle [14].

(a)   (b)
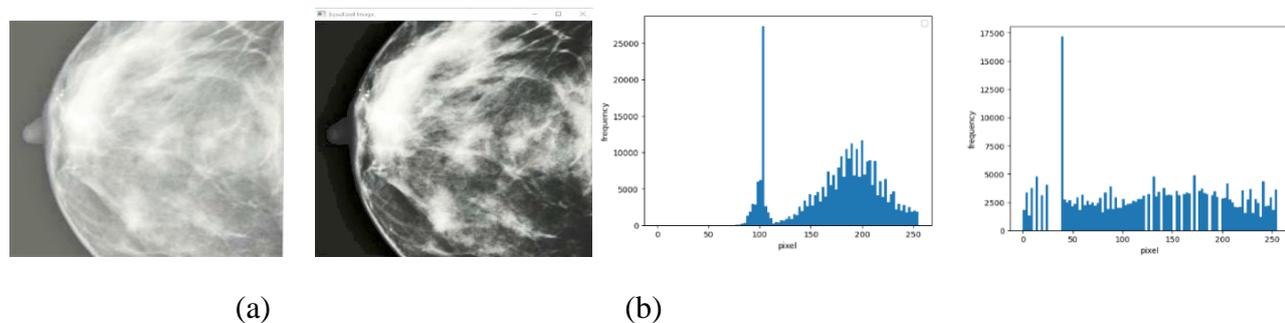
Fig 5: Image and histograms of a histogram equalized mammogram image.

Data Augmentation

Using convolutional neural networks effectively requires a considerable amount of data for learning its

parameters. A standard technique for expanding the training data set is augmentation. Augmentation helps in improving system performance, reduces the chance of overfitting and data imbalance as a result of few datasets. There are many techniques for data augmentation. They include random reflection, rotations and horizontal or vertical translations. For this purpose, Image Data Generator class from the keras library was used to generate augmentation data by using the techniques already stated. The Image Data Generator takes in parameters such as validation_split for splitting the data into training and validation sets; height_shift_range, width_shift_range, rotation_range, rescale for vertical, horizontal, angle rotations and scaling respectively.

Summarily, a total of 322 images were obtained from MIAS database out of which 122 have calcification abnormality. The 122 images with calcification were then extracted from the dataset and added to 74 images obtained locally. The locally sourced images thus accounts for 38 percent of the total number of images used for the training. The data augmentation techniques applied increased the dataset by a factor of 30 thereby resulting in a total of 5,880 mammogram images. The dataset was split to 70 percent for training, 20 percent for validation and 10 percent for testing.

Model Architectures

Deep Neural Networks are one of the most successful techniques used in medical imaging, however, training the model from scratch is usually time consuming and with high computational complexity, requiring strong GPU processors that are usually considerably expensive. Fortunately, we can base our model on the initial layers of known successful neural network models. This is because the initial layers of the neural network focus on feature extraction. Four pre-trained convolutional neural network classifiers were experimented and evaluated on several metrics. They are VGG-16, VGG-19, DenseNet - 121 and MobileNetV2. These convolutional neural network architectures were pre trained with over one million images from ImageNet datasets and could classify one thousand objects in an image.

VGG-16 and VGG-19 are a sixteen and nineteen layers pretrained deep convolutional neural network. MobileNetV2 is a 53 layer convolutional neural network while DenseNet-121 is a dense network architecture with 120 convolutions and 4 average pooling blocks. The pre-trained modelswere imported from tensorflow library. The last output layer was flattened and then connected to a dense layer of 512 neurons and relu activation function. The resultant architecture was then connected to a final dense layer of two neurons and softmax activation function. The final layer indicates the result of the network with each neuron indicating the degree to which the image is benign or malignant.

```python
from tensorflow.keras.applications.vgg16 import VGG16
model16 = VGG16()
model_16 = VGG16(weights = "imagenet", include_top=False,input_shape=(256,256,3))
model_16.trainable = False
flat1 = Flatten()(model_16.layers[-1].output)
class1 = Dense(512,activation="relu")(flat1)
output = Dense(2,activation="softmax")(class1)
model_16 = Model(inputs=model_16.inputs, outputs=output)
```

Fig 6: Code snippet used in importing VGG16 pretrained model in Anaconda environment.

Model Training

The developed model architectures were compiled with Categorical cross entropy as model loss. For optimization, Adaptive moment estimation, popularly known as adam's optimizer, was used. Early stopping mechanism was applied such that if the model fails to improve after several epochs (specifically 40), the training will be stopped and the model with the best performing weights at the end of its epoch will be restored. For easy resumption of training, model checkpoints were set such that the trained model can be saved after every epoch. Using the tensorflow library, the save_best_only flag was set to true so that the best models are saved. This means that if at the end of an epoch, the produced model does not perform better than the existing best model, then the new model will not be saved. The model was compiled to save the following metrics, Accuracy, Precision, Recall, Area under Curve (AUC), True positives, True Negatives, False

Positives, False Negatives, Sensitivity at specificity and Specificity at Sensitivity. The training algorithm followed the flowchart below.
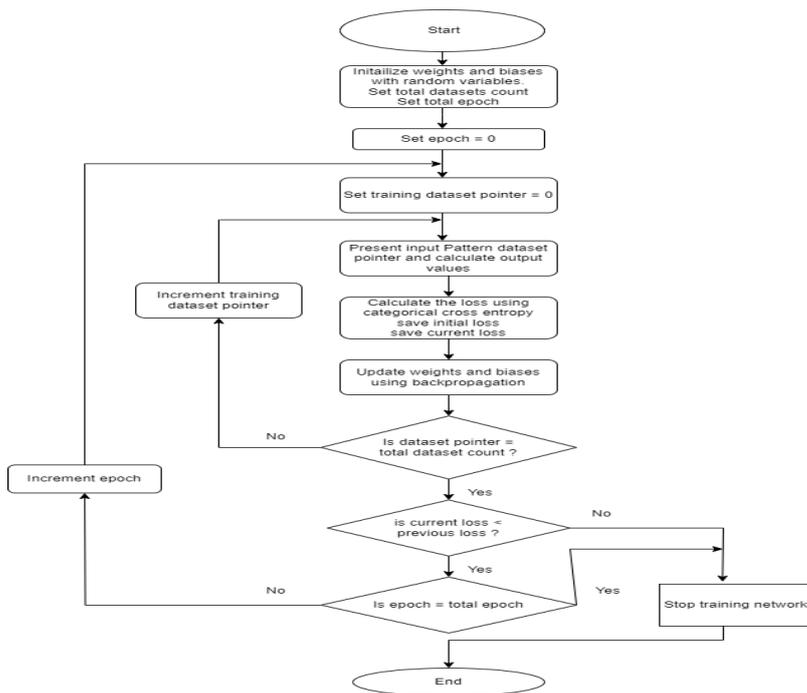
Fig 7: The model training flow chart

RESULTS AND DISCUSION

In this section, the results of the models developed were presented. The following metrics were used to evaluate the performance of the developed models: Accuracy, Precision, Recall, Area under curve, True Positives, True Negatives, False Negatives, False Positives, Sensitivity at specificity and Specificity at sensitivity.

Training Results
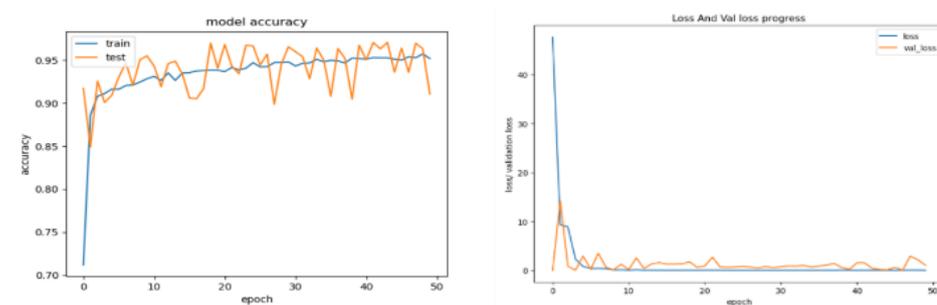
Fig 8: Training and Validation loss and accuracy for VGG-16 based model
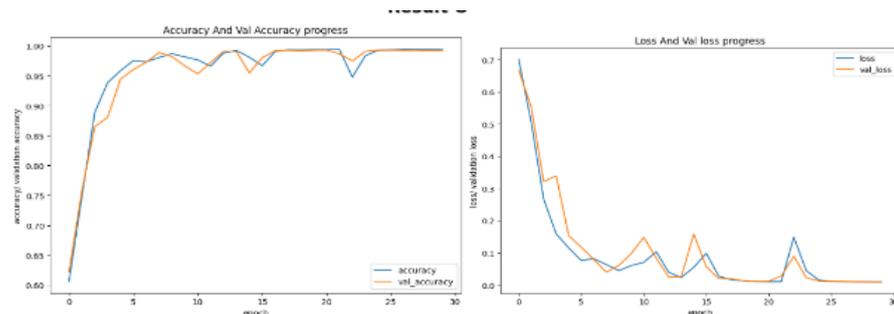
Fig 9: Loss and validation loss and accuracy for VGG-19 based model training
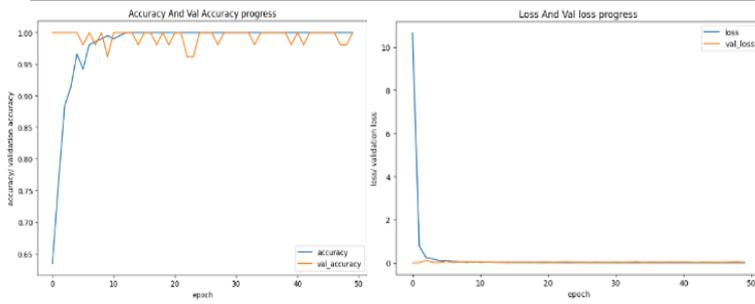
Fig 10: Loss and validation loss and accuracy for MobileNetV2 based model
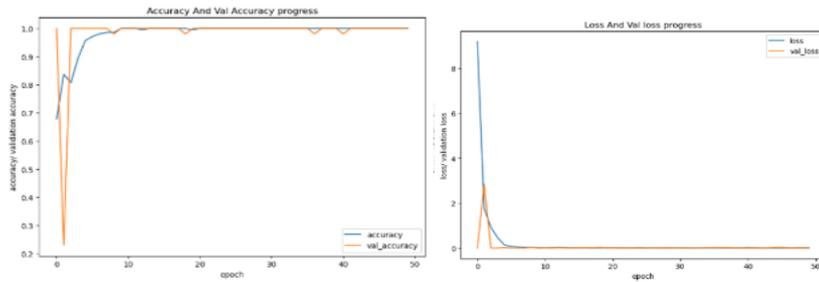
Fig 11: Loss and validation loss and accuracy for DenseNet-121 based model.

DenseNet-121 by far performs better than the other three models as evidenced in Fig 8 through to Fig 11. This could be attributed to the high complex network architecture compared to other models used in this report. The accuracy quickly reaches near 100% while the loss also dropped to near zero as we progressed from epoch 1 to 50.

Test Results

The model was tested with 23 random samples of image scans that had not been previously passed to the model before in chapter three, section 3.7 and the following results were obtained for loss value accuracy recall, AUC, sensitivity at specificity and specificity at sensitivity.

The parameters are described as shown below:

1. True Positive (TP): Observation is positive and predicted to be positive.
2. False Negative (FN): Observation is positive but predicted negative.
3. True Negative (TN): Observation is negative and predicted to be negative.
4. False Positive (FP): Observation is negative but predicted positive
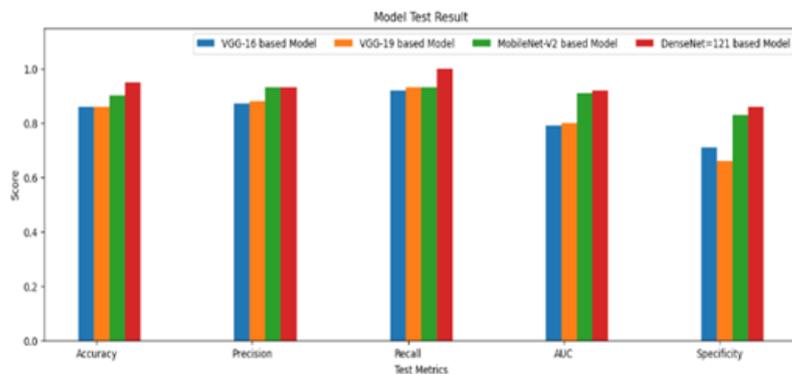
Figure 12: Plot of the model evaluation metric.

As can be seen in the plot of Fig 12, The best performing model is the DenseNet-121 based model. It achieved an accuracy of 0.95, a 5.3% improvement from the closest model - MobileNetV2 based model with an

accuracy of 0.90. While accuracy depicts the percentage of the predictions that are correct, it can, in situations of unbalanced datasets, depict a false sense of high performance. Supposed that this test was done with nineteen benign datasets and one malignant case, a model that predicts benign no matter the input will score an accuracy of 0.95 simply because the dataset is unbalanced.

For a cancer detection application, the rate of true positive is of utmost importance since it is desirable for the model to detect every case of malignant calcifications. The metric that highlights this is the recall as it evaluates the proportion of actual positives that are correctly identified. The DenseNet-121 based model performed highly on this metric as well as other models. Specificity represents the proportion of actual negatives that are correctly identified while Precision is the proportion of true positives out of the total predicted positives. AUC stands for area under curve which generally rates the entire model. Models with AUC of 0.9 and above are considered excellent models, those between the ranges of 0.8 to 0.9 are good models, and fair models are for models with 0.7 to 0.8. A model between 0.6 and 0.7 is considered poor while below 0.6 is a very poor model. Using the AUC, The DenseNet-121 based model and MobileNet-V2 model is considered an excellent model. VGG-19 and VGG-16 are considerably good model as they fall within the range of model considered as such.

Table 1: Confusion Matrix for the models

Confusion Matrix		VGG-16		VGG - 19		MobileNet-V2		DenseNet-121	
TP	FN	13	1	14	2	14	1	14	0
FP	TN	2	5	1	4	1	5	1	6

Computational Complexity Analysis

The time complexity of the computations in a neural network model is established by the number of floating point operations (FLOPs) carried out in its forward pass. To calculate the FLOPs in a model, here are the rules:

Convolutions - FLOPs = 2 x Number of Kernel x Kernel Shape x Output Shape

Fully Connected Layers - FLOPs = 2 x Input Size x Output Size

Pooling Layers - FLOPs = Height x Depth x Width of an image

Pooling Layer with a stride, other than 1 - FLOPs = (Height / Stride) x Depth x (Width / Stride) of the image.

Where the output shape of a convolutional operation is = (Input Shape - Kernel Shape) + 1

At the final layers of each network developed, the final 1000 neurons were fully connected to a layer with 512 neurons which were then connected to the output layer with 2 neurons.

Consequently the FLOPs for the last two layers are Penultimate layer: 2 x 1000 x 512 = 1,024,000 FLOPs

Output layer: 2 x 512 x 2 = 2,048 FLOPs

Total: 2,048 + 1,024,000 = 1,026,048 FLOPs

VGG-16 network has 15.3 billion FLOPs, VGG-19 has 19.6 billion FLOPs, MobileNet-V2 has 336.34 million FLOPs while Dense Net has 5.69 billion Flops. The total floating point operations for each of the models developed is obtained by adding 1,026,048 FLOPs to the corresponding FLOPs of the parent network as shown in Table 2

Table 2: Floating Point Operations for each model developed

Model	GFLOPS
VGG-16 Based	15.4
VGG-19 Based	19.7
MobileNet-V2 Based	0.337
DenseNet-121 Based	5.79

Parameters of the neural networks which essentially determines the space complexity of the network and size of the model generated. Table 3 shows details of the network parameters for all the models developed in this work while Table 5 shows the test time for the networks.

Table 3: Number of parameters of the models

Model	Trainable Parameters	Non-trainable Parameters	Total Parameters
VGG-16 based	16,778,754	14,714,688	31,493,442
VGG-19 based	16,778,754	20,024,384	36,803,138
MobileNet-V2 based	41,944,578	2,257,984	44,202,562
DenseNet-121 based	33,555,970	7,037,504	40,593,474
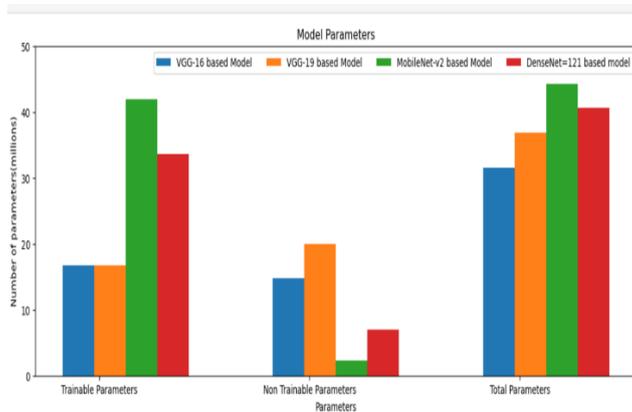
Fig 13: Parameters of the trained model.
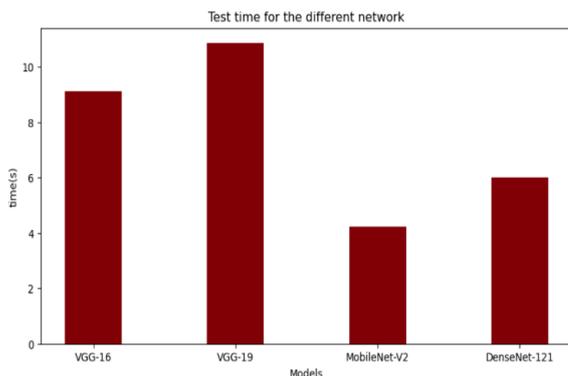
Fig 14: Test time for the different networks built.

As evidenced in the plot of Figure 14, MobileNet-V2 based model has the least test time, followed by DenseNet-121, then VGG-16 and finally VGG-19.

Comparison With Existing Models

For recall, the best performing models – MobileNet-V2 and DenseNet-121 based models - in this thesis were compared with MCDNet developed by Zhao (2022) and the model by Zhang, (2019) which were tested on both INBreast Individual microcalcification image dataset and the results shown in Table 5. The models were also compared with average accuracies obtained by Aloyari, (2020)'s models which were based on ResNet18, Inception-V3 and ShuffleNet to obtain the result in Table 5.

Table 4: Comparison between models based on recall

Model	Recall
Zhao (2022) MDCNet	0.98
Zhang, (2019)	0.89
MobileNet-V2 based model	0.93
DenseNet-121 based model	1

Table 5: Comparison between models based on accuracy

Model	Accuracy
Aloyari, (2020) ResNet18	0.94
Aloyari, (2020) Inception-V3	0.93
Aloyari, (2020) ShuffleNet	0.92
MobileNet-V2 based model	0.90
DenseNet-121 based model	0.95

As shown in Table 4, DenseNet-121 based model achieved better performance both in accuracy and recall among the models compared. In terms of recall, it is followed closely by Zhao, 2022's MDCNet [15]. MobileNet-V2 based model performed better than Zhang, (2019) [16] but achieved less than Zhao, (2022) in terms of recall. In terms of accuracy, Aloyari, (2020)'s models performed better than the MobileNet-V2 based model but since they are based on ResNet18, Inception-V3 and ShuffleNet, the model have higher computational cost than MobileNet-V2 with 1.82GFLOPs, 6GFLOPs and 360MFLOPs respectively.

CONCLUSION

In the analysis of the models developed, DenseNet-121 was seen to outperform other models with a recall of 1.00 and accuracy of 0.95. Second best performing model was the MobileNet-V2 with recall and accuracy of 0.93. The best two performing models were then compared with other existing model and it was shown that DenseNet-121 outperforms other models in accuracy and recall. Although MobileNet-V2 could not perform other models in recall and accuracy, it outperforms the rest of the model in computational cost with a very decent accuracy of 0.9 and recall of 0.93. It follows then that the DenseNet-121 based model can be deployed in application with enough computing power such as personal computers and workstations, while the

MobileNet-V2 model can be deployed to mobile phone applications.

REFERENCES

1. F. Bray, M. Laversanne, H. Sung, J. F. ME, R. L. S. I. Soerjomataram and A. Jemal, "Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," CA: A Cancer Journal for Clinicians, vol. 74, no. 3, pp. 229 - 2024, 2024.
2. WHO, "Breast Cancer now most common type of cancer: WHO taking action," 3 February 2024. [Online]. Available: https://www.who.int/news/item/03-02-2021-breast-cancer-now-most-common-form-of-cancer-who-taking-action.
3. Omolara A Fatiregun, "Breast Cancer Research to Support Evidence-Based Medicine in Nigeria: A Review of the Literature," JCO Glob Oncol., pp. 384 - 390, 2021.
4. American Cancer Society, "How Common is breast Cancer," 12 January 2021. [Online]. Available: https://www.cancer.org/cancer/breast-cancer/about/how-common-is-breast-cancer.html#written_by.
5. D. A. Rojas and Nandi, "Toward breast Cancer Diagnosis based on automated segmentation of masses in mammograms.," Pattern Recognition, pp. 1138-1148, 2009.
6. H. Moradmand, S. Setayeshi and H. Targhi, "Comparing Methods for segmentation of Microcalcification Clusters in Digitized Mammograms," International Journal of Computer Science, 2011.
7. A. K. Tucker, Textbook of mammography, Livingstone: Churchill Livingstone, 2001.
8. Dana, B., & Raed, S. (2016). "Comparative Study of Machine Learning Algorithms for Breast Cancer Detection and Diagnosis.," The 2016 IEEE 5th International Conference on Electronic Devices, Systems, and Applications (ICEDSA'2016)., 2016.
9. A. S. Fabio, "Automatic Breast Cancer Classificatin from hispathological images: A hybrid approach," Curitiba : Federal University of Parana, 2018.
10. Khan, H. N., Shahid, A. R., Raza, B., Dar, A. H., & Alquhayz, H. (2019)., "Multi-View Feature Fusion Based Four Views Model for Mammogram Classification Using Convolutional Neural Network.," IEEE Access, pp. 165724-165733, 2019.
11. A. Aloyayri, Breast Cancer Classification from Histopathological Images Using Transfer Learning, Montreal: Department of Computer Science at Concordia University Montreal, Quebec, Canada, 2020.
12. S. R. Kebede, F. Waldamichael, T. Debelee, M. Aleme, W. Bedane, B. Mezgebu and Z. C. Merga, "Dual view deep learning for enhanced breast cancer screening using mammography," Scientific Reports (Sci.Rep), p. 3839, 2024.
13. J. Suckling, J. Parker , D. Dance, S. Astley, I. Hutt, C. Beggis and I. Rickets, "Mammographic Image Analysis Society (MIAS) database v1.21 [Dataset]," 2015. [Online]. Available: https://www.repository.cam.ac.uk/handle/1810/250394.
14. C. E. Eze, "Breast Mammography Scans," Kaggle, 2023.
15. Zhang F, Luo L, Sun X, Inwei S, Zhen Z, Xiuli L, Yizhou Y, Yizhou W, "Cascaded generative and discriminative learning for microcalcification detection in breast mammograms.," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 165724-165733, 2019.
16. Z. H., "Studies on deep learning approach in breast lesions detection and cancer diagnosis in mammograms.," Masters Thesis, Department of computing, Faculty of Techinology, Univerisity of Turku, 2020.
17. S. Maanvi, "Most Of The World's Cancer Cases Are Now In Developing Countries," 15 December 2015. [Online]. Available: https://www.npr.org/sections/goatsandsoda/2015/12/15/459827058/most-of-the-worlds-cancer-cases-are-now-in-developing-countries.
18. A. F. Omolara, T. Oluokun, N. N. Lasebikan, E. Nwachukwu, A. A. Ibraheem and O. Olopade, "Breast Cancer Research to Support Evidence-Based Medicine in Nigeria: A Review of the Literature," An American Society of Clinical Oncology Journal, pp. 384 - 390, 2021.
19. P. W. Sonali B.M, "Research Paper on Basic of Artificial Neural Network," International Journal on Recent and Innovation Trends in Computing and Communication, p. 96 – 100, 2014.
20. Prabhu, "Understanding of Convolutional Neural Network (CNN) – Deep Learning," 4 March 2018. [Online]. Available: https://medium.com/@RaghavPrabhu/understanding-of-convolutional-neural-network-cnn-deep-learning-99760835f148. [Accessed 11 May 2021].

21. Cancer Council, Understanding Breast Cancer, Australia: SOS Print + Media Group., 2020.
22. B. D. Hansa, "Breast Cancer," 8 June 2020. [Online]. Available: https://www.webmd.com/breast-cancer/understanding-breast-cancer-basics.
23. UCSF Health, "Breast Cancer Diagnosis," 25 January 2021. [Online]. Available: https://www.ucsfhealth.org/conditions/breast-cancer/diagnosis. [Accessed 6 29 2021].
24. F. Prinzi1, M. Insalaco, A. Orlando, S. Gaglio and S. Vitabile, "A Yolo Based Model for Breast Cancer Detection in Mammograms," Cognitive Computation , pp. 107-120, 2023.