

Convolution Neural Network for Breast Cancer Detection and Classification Using Deep Learning

Basem S Abunasser^{1*}, Mohammed Rasheed J AL-Hiealy¹, Ihab S Zaqout², Samy S Abu-Naser²

Abstract

Objective: Early detection and precise diagnosis of breast cancer (BC) plays an essential part in enhancing the diagnosis and improving the breast cancer survival rate of patients from 30 to 50%. Through the advances of technology in healthcare, deep learning takes a significant role in handling and inspecting a great number of X-ray, MRI, CTR images. The aim of this study is to propose a deep learning model (BCCNN) to detect and classify breast cancers into eight classes: benign adenosis (BA), benign fibroadenoma (BF), benign phyllodes tumor (BPT), benign tubular adenoma (BTA), malignant ductal carcinoma (MDC), malignant lobular carcinoma (MLC), malignant mucinous carcinoma (MMC), and malignant papillary carcinoma (MPC). **Methods:** Breast cancer MRI images were classified into BA, BF, BPT, BTA, MDC, MLC, MMC, and MPC using a proposed Deep Learning model with additional 5 fine-tuned Deep learning models consisting of Xception, InceptionV3, VGG16, MobileNet and ResNet50 trained on ImageNet database. The dataset was collected from Kaggle depository for breast cancer detection and classification. That Dataset was boosted using GAN technique. The images in the dataset have 4 magnifications (40X, 100X, 200X, 400X, and Complete Dataset). Thus we evaluated the proposed Deep Learning model and 5 pre-trained models using each dataset individually. That means we carried out a total of 30 experiments. The measurement that was used in the evaluation of all models includes: F1-score, recall, precision, accuracy. **Results:** The classification F1-score accuracies of Xception, InceptionV3, ResNet50, VGG16, MobileNet, and Proposed Model (BCCNN) were 97.54%, 95.33%, 98.14%, 97.67%, 93.98%, and 98.28%, respectively. **Conclusion:** Dataset Boosting, preprocessing and balancing played a good role in enhancing the detection and classification of breast cancer of the proposed model (BCCNN) and the fine-tuned pre-trained models' accuracies greatly. The best accuracies were attained when the 400X magnification of the MRI images due to their high images resolution.

Keywords: Breast cancer- deep learning- convolution neural network

Asian Pac J Cancer Prev, 24 (2), 531-544

Introduction

Breast Cancer is considered one of the most common types of cancers and is taken as the 2nd to cause death between women worldwide. In 2002, it was the 2nd most common cancer globally, by means of exceeding one million new cases. Despite the enhancements in early detection and knowing of the molecular foundations of the biology of breast cancer, nearly 30% of the patients with "early-stage" breast cancer have disease recurrence (Gore et al., 2022; Hashemi et al., 2021; Olayide et al., 2021).

Determining the most real and least toxic therapy, molecular and clinical features of the tumor in certain need of inspection. General treatment of breast cancer comprises hormonal agents, immunotherapy and cytotoxicity. These medications are used in adjuvant,

transitional, and neoadjuvant modes. Overall, systemic agents are vigorous at the start of management in 90% of main breast cancers and 50% of metastases. Though, after a flexible period of time, progress does follow. At the present, opposition to treatment is not only popular but anticipated" (Cahoon et al., 2002; Yasmin et al., 2013; Ahmed et al., 2020; Patil et al., 2020).

Diagnosing BC early is the most important feature of its treatment. Amongst the diverse types of diagnosing technique of BC, imaging is the main diagnostic method that can offer important data on the patients of BC. It was revealed that a few techniques of imaging e.g.: "Magnetic Resonance Imaging (MRI), Single-Photon Emission Computed Tomography (SPECT), mammography, Computerized Tomography (CT), and Positron Emission Tomography (PET)" can be put into use for the identifying

¹University Malaysia of Computer Science & Engineering (UNIMY), Cyberjaya, Malaysia. ²Faculty of Engineering and Information Technology, Al-Azhar University, Gaza, Palestine. *For Correspondence: p05210002@student.unimy.edu.my

and observing of the patients of BC in diverse stages. Moreover, techniques of imaging, and the use of biomarkers of biochemical e.g.: “proteins, mRNAs, DNA, and microRNAs” can be used as an innovative analytic and therapy gears for BC patients (Hafemann et al., 2014; Tajbakhsh et al., 2016; Widiana and Irawan, 2020).

Deep Learning (DL) as a type of Machine Learning method has its working mechanism inspired from the way human brain neurons process information. The most basic element of the DL networks are small nodes known as artificial neurons, which are usually arranged in layers wherein each neuron has connections to every neuron in the subsequent layer via weighted connections. Recently, the rise of DL technique has encouraged different areas of study to solve complex problems or enhance performances of existing study using the new technology. Examples of application of DL include machine translation, speech recognition, sentiment analysis, image recognition, face recognition, signal processing, etc. (Waks and Winer, 2019; Jafari et al., 2018; El-Habil and Abu-Naser, 2022; Barhoom et al., 2022; Arqawi, 2022; Alzamily et al., 2022; Alayoubi et al., 2022; Almasri et al., 2022; Almasri et al., 2023).

The current trend of applying DL technique in medical applications has reaped great success with the potential of DL technology to perform faster analysis with higher accuracy when compared with human practitioners. To give an example, a notable study by Google on the diagnostic classification of diabetic retinopathy has shown remarkable performance that exceeds the capabilities of domain experts (Joshi and Mehta, 2017). In addition, the application of transfer learning techniques can be seen in several studies. As opposed to training a model from scratch, transfer learning method allowed the use of weights trained previously on a specific task to be reused as the starting point for a model on another task (Abunasser et al., 2022; Saleh et al., 2020; Arqawi et al., 2020; Mady et al., 2020; Albatish and Abu-Naser, 2019).

It comes with the benefits of shorter training time and is possible to deliver better results. CNN models such as Alex Net, VGG, ResNet, and Xception are the most dominant models applied in the transfer learning approach. Generally, there are two ways of applying transfer learning technique; First, a pre-trained model with weights trained on ImageNet dataset can be used as feature extractor; Second, fine-tuning of pre-trained model on a new problem. Study utilized transfer learning technology with Xception neural network to speed up the training of models for distinguishing subjects (Kumar et al., 2017; Abu Ghosh et al., 2016; Naser, 2006-2010; Sachdeva et al., 2021).

Literature Review

Many researchers have employed artificial intelligence, expert systems, and neural networks in the diagnosis of BC to increase the screening accuracy. Usually hospitals use x-rays for diagnosis of BC, but lately, hospitals have been using mammography images as a substitute for x-rays, due to their ease in analyzing and studying through intelligent models, which have increased the efficiency and accuracy of BC diagnosis.

A number of models and methods were proposed for increasing the efficacy of the diagnosis of BC. These methods includes: “Linear Regression (LR), Artificial Neural Network (ANN), K-Nearest Neighbors (KNN) search, Softmax Regression, and Support Vector Machine (SVM), and Convolutional Neural Network (CNN).”

Kiyan and Yildirim, (2004) ; Gonzalez-Angulo et al., (2007) collected their datasets from Kaggle depository. They proposed a prediction model to predict whether or not a person has breast cancer and to provide awareness or diagnosis about it. The authors of these studies made comparisons using accuracy of each result of the SVM, random forest (RF), Naive Bayes classifier, and logistic regression on the dataset to deliver a precise model for breast cancer prediction. The outcome of their experiments indicated that the techniques of machine learning models that were applied in their studies predicted breast cancer disease with an accuracy between 52.63% and 98.24%.

In these studies, the authors proposed new KNN models for the need for early diagnosis and precise diagnostic procedures that clinicians can use to classify whether cancer is benign or malignant. The main objective of their study was to compare the results of supervised learning classification algorithms and to combine these algorithms using a classification technique called voting. Voting was a grouping method because they could combine multiple models to achieve higher classification accuracy. The datasets were collected from the University of Wisconsin. »Ponraj et al., (2012) achieved 98.90%, Pavithra et al., (2016) achieved an accuracy of 97.60%, Asri et al., (2016) achieved 97.13%, Gayathri et al., (2013) achieved 99.9%, Ak, (2020) achieved 98.10% Asri et al., (2018) achieved 98.23% and Kumar et al., (2017) achieved 83.45%.

In the following studies like (Agarap, 2018; Pritom et al., 2016; MurtiRawat et al., 2020; Selvathi and Aarthi, 2017; Lg and At, 2013) the authors argued that early detection and prevention can significantly reduce the chances of passing away. «An important fact about breast cancer prognosis was to improve the likelihood of cancer recurrence. Their studies aimed to find the probability of breast cancer recurrence using different machine learning techniques like SVM. The authors presented new approaches in order to improve the accuracy of these models. Cancer patient data were collected from the Wisconsin Dataset of the UCI Machine Learning Repository. The dataset contains a total of 35 attributes in which they applied the Naive Bayes, C4.5 Decision Tree and SVM algorithms and measured their prediction accuracy. The efficient feature selection algorithm helped them improve the accuracy of each model by reducing some lower-order features. Not only are the contributions of these traits much lower, but their addition also misleads the classification algorithms. After careful selection of higher-order attributes, they significantly improved accuracy rate for all algorithms.

The Study of (Spanhol et al., 2016; Gayathri et al., 2013; Shen et al., 2019; Wadkar et al., 2019; TIWARI et al., 2020; Ragab et al., 2019) «proposed method that uses, CNNs that is a particular type of deep learning, feedforward network, and some preliminary experiments

used the deep learning approach to classify breast cancer mammography images from BreakHis, which is a public dataset. Method based on the extraction of image patches for training the CNN and the combination of these patches for final classification. All their convolutional network techniques for categorizing screening mammograms reached good accuracies. The finest performance on an independent test set of digitized film mammograms from the Digital Database for Screening Mammography was 0.88%, (the sensitivity: 86.2%, the specificity: 80.2%).

The aim of the studies (Westermann et al., 2002; Joo et al., 2004; Salem et al., 2017) was to develop methods for classifying cancers into specific prognostic categories based on gene expression signatures using artificial neural networks. «ANN were trained to use small round blue cell tumors (SRBCTs) as a model. These cancers belong to four distinct diagnostic classes and often present diagnostic dilemmas in clinical practice. ANN properly classified all samples and recognized the genes most related to the classification. The experimental results suggested that the new strategies were able to improve the stability of the selection results as well as the sample classification accuracy. The new algorithms achieved accuracy of classification about 99%. »

Summary of Previous Studies

We made a summary of the studies that we have discussed in the previous section in terms of the following criteria: Machine Learning methods used, programming language used, best result attained, best method, and source of the dataset used in Table 1.

There are numerous studies and suggestions that put on to deal with breast cancer. Each suggestion or study has a different way of detecting or dealing with it. Some of the studies use methods that rely on image processing and the use of systems for this purpose and other studies that address the form of the breast and observe any changes in it. Some studies have enhanced the quality of algorithms that exist for detecting the disease. All these studies concentrate on the detection of whether breast cancer exists or not. None of these studies classified the eight different classes of breast cancer. In the current

study we will concentrate on the classification of the eight classes of breast cancers: “benign adenosis, benign fibroadenoma, benign phyllodes tumor, benign tubular adenoma, malignant ductal carcinoma, malignant papillary carcinoma, malignant mucinous carcinoma, and malignant lobular carcinoma.

Materials and Methods

Dataset

We have collected our dataset from Kaggle depository. The Breast “Cancer Histopathological Image Classification (BreakHis) consists of 7909 microscopic images of breast cancer tissue gathered from 82 patients using various magnifying factors (40X, 100X, 200X, and 400X). It contains 2,480 benign and 5,429 malignant samples (700X460 pixels, 3-channel RGB, 8-bit depth in each channel, PNG image format). This dataset was built in collaboration with the P&D Laboratory – Pathological Anatomy and Cytopathology, Parana, Brazil.

After the dataset was downloaded from Kaggle depository, and stored in Google Drive to be used for the training, validation, and testing of detecting and classifying breast cancer. The images needed to be clearly organized into folders to be ready for usage. Figure 1 shows how the dataset was extracted from the original supplied folder into the new folders.

All the previous studies reviewed in the previous section uses the detection approach of breast cancer. That means that all previous studies looked at 2 classes of breast cancers (benign and malignant cancers); furthermore, they did not take into consideration that there are 4 categories of benign and 4 categories of malignant cancers.

The Dataset is organized so that we can use the 40X magnification as standalone for training and validation and testing with 8 classes of cancers. We will record all the findings for later analysis. We will do the same thing with 100X, 200X, and 400X. After that we will use the complete BreakHis dataset as one dataset comprising 40X, 100X, 200X, 400X together to see which data set has better performance. That means we will use 5 datasets to try and find the best performance of them.

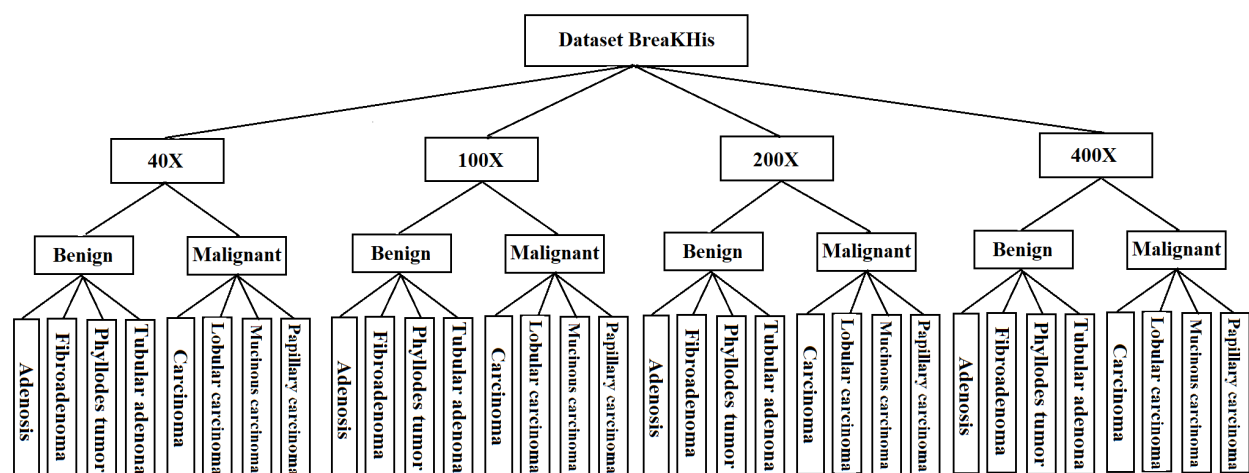


Figure 1. Breast Cancer Histopathological Dataset (BreakHis) Organization

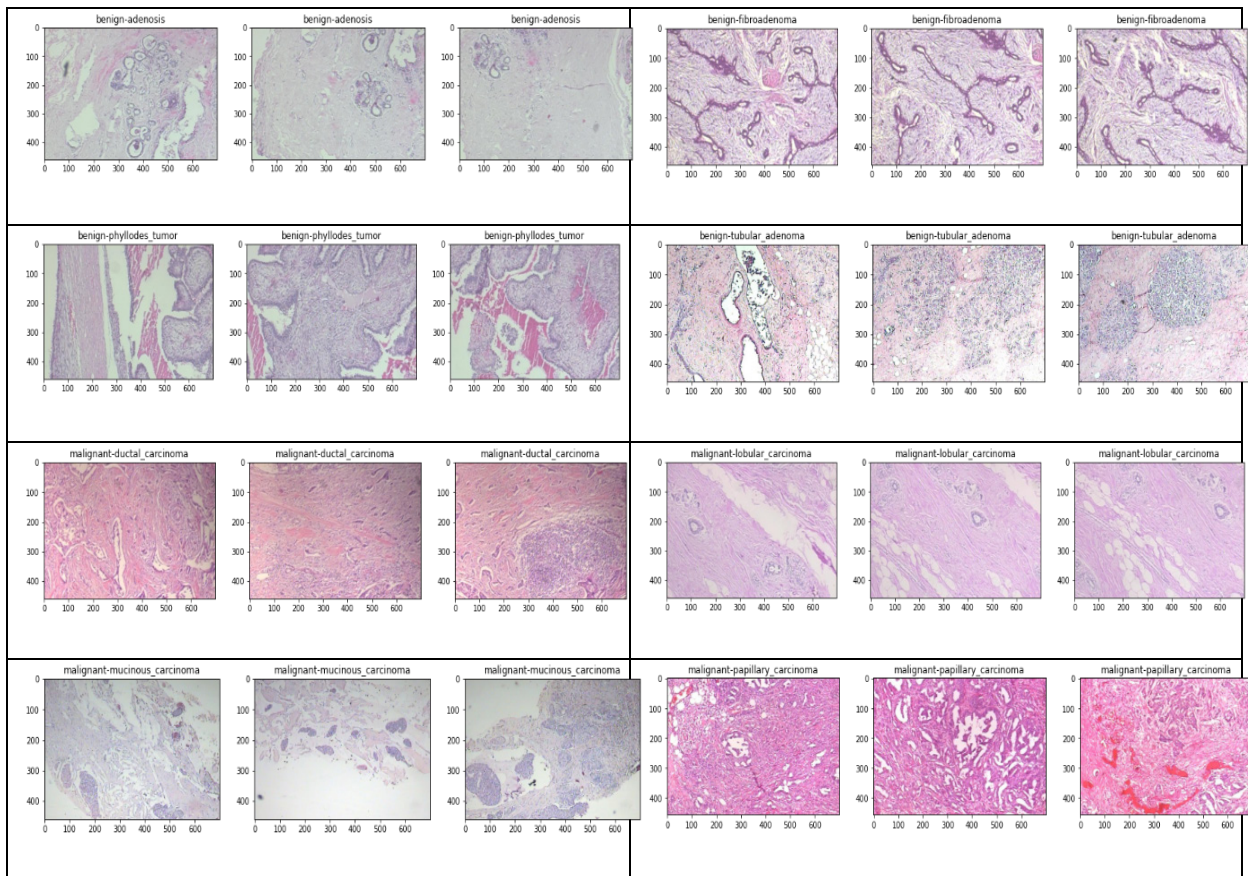


Figure 2. Samples Images from 40X Datasets with Its Eight Classes

Data Generation

The number of images in the original BreakHis dataset has 7909 images. This number of images is considered low and thus we can use new methods to generate more images to boost the dataset.

We can use data augmentation to generate new images from existing images; however, we decided to use “Generative Adversarial Networks” (GANs). GAN is an algorithmic construction that utilizes two NNs, fighting one in contradiction of the other (consequently the “adversarial”) to be able to produce new, fake instances of images that can pass as real images. They are very popular in image generation, voice generation, and video generation (Musleh et al., 2019).

So, we used the GAN algorithm to boost the BreakHis dataset from 7909 images to 40,000 images. Table 2 illustrates the summary of images after using GAN to generate new images. Samples of the dataset are shown in Figure 2.

Dataset Splitting

To guarantee unbiased model classification performance, each dataset was ensured to have balanced classes. The dataset was divided into 3 different sets which are (1) training dataset, (2) validation dataset, and (3) testing dataset with a split ratio of 60:20:20 respectively. The first 4 datasets (40X, 100X, 200X, 400X) each consisted of 10,000 images; there were 6,000 training images, 2,000 validation images, and 2,000 testing images. For the last dataset (complete Dataset), it

consisted of 40,000 images; there were 24,000 training images, 8,000 validation images, and 8,000 testing images.

Deep Learning Models used in all of the Experiments

We have 6 deep learning models: The proposed model (BCCNN) and five pre-trained models (ResNet50, VGG16, Xception, InceptionV3, and MobileNet). These pre-trained models are the most popular Deep Learning models.

The proposed model (BCCNN)

The proposed model is Convolutional Neural Network and it is called Breast Cancer Convolutional Neural Network (BCCNN) with 18 layers. The input dimension of the images can be between 48x48 to 224 x 224 pixels. BCCNN is a convolution neural network architecture for classifying Breast cancers. The exclusive thing about BCCNN is that rather than consuming a huge number of hyper-parameters they focus on having convolution layers of 3x3 filter with a stride 1 and it uses the same padding and max-pooling layer of 2x2 filter of stride 2. It uses this arrangement of convolution and max pooling layers regularly all the way through the entire architecture as can be seen in Figure 3. At the end it has one fully connected layers followed by a softmax for output.” This model has a big architecture and it has approximately 2.1×10^7 parameters

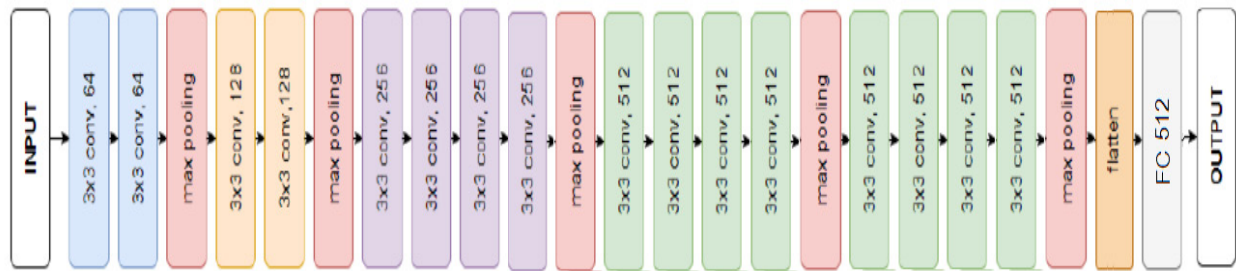


Figure 3. Shows the Structure of the Proposed Model BCCNN

ResNet50

The pre-trained ResNet50 model was used as a feature extractor and swapped a new densely connected classifier for prediction. Residual Networks are the first deeper neural networks that enabled the training of hundreds or even thousands of layers while maintaining compelling performance (He et al., 2016). In 2015, a variant of the ResNet model, which is ResNet50 composed of 50 layers, won the ILSVRC 2015 challenge (as in Figure 4).

VGG16

The pre-trained VGG16 model was used in the form of a feature extractor. In 2014, VGG-16, a very deep CNN model was proposed by Karen Simonyan and Andrew Zisserman from Visual Geometry Group Lab of Oxford

University won the ILSVRC 2014 competition (Simonyan and Zisserman, 2014).

The architecture of VGG16 is appealing due to its uniform architecture. The model requires input in the form of an image with a fixed dimension of 224 x 224 x 3. The overall architecture consists of several blocks of convolutional layer and max-pooling layer followed by a dense classifier that outputs 1000 class scores.

VGG16 comes with 138,357,533 (138 million) trainable parameters attributed to the vast amount of neurons in the fully connected layers, making it one of the largest CNN architectures. Training VGG16 from scratch can be challenging, slow, and computationally costly. However, transfer learning techniques allowed VGG16 with pre-trained weights to be used for feature extraction

Table 1. A Summary of the Previous Studies by Methods, Best Method, Language, Result, Dataset Used

Reference	Methods Used	Best Methods	Language	Best Result	Data Provider
(Westermann et al., 2002)	ANN	ANN	-	96	Private
(Joo et al., 2004)	ANN	ANN	Python	90.3	Kaggle
(Kiyan and Yildirim, 2004)	LR	LR	-	97	Wisconsin
(Spanhol et al., 2016)	CNN	CNN	Python	89.5	BreaKHis
(Salem et al., 2017)	ANN	ANN	-	99	Wisconsin
(Kumar et al., 2017)	KNN	KNN	-	83.45	Wisconsin
(Ponraj et al., 2012)	KNN	KNN	-	98.9	FMC
(Gayathri et al., 2013)	CNN	CNN	Python	90	BreaKHis
(Shen et al., 2019)	CNN	CNN	Python	88	DDSM
(Agarap, 2018)	SVM,KNN	SVM	-	99	Wisconsin
(Pavithra et al., 2016)	SVM,KNN,ANN	SVM	-	97.6	Wisconsin
(Asri et al., 2016)	SVM,KNN	KNN	-	97.13	Wisconsin
(Gonzalez-Angulo et al., 2007)	SVM,LR	LR	Python	98.24	Kaggle
(Pritom et al., 2016)	SVM	SVM	-	92.3	Wisconsin
(Joshi and Mehta, 2017)	SVM,KNN	KNN	-	99.9	Wisconsin
(David et al., 2019)	SVM,ANN	SVM	-	98.82	Wisconsin
(Wadkar et al., 2019)	SVM,KNN,CNN	CNN	-	97	Wisconsin
(TIWARI et al., 2020)	SVM,KNN,CNN	CNN	Python	99.3	Wisconsin
(Ragab et al., 2019)	SVM,ANN	ANN	Python	94	DDSM
(Ak, 2020)	SVM,KNN	SVM	Minitab	98.1	Wisconsin
(Asri et al., 2017)	SVM,KNN	SVM	WEKA	97.13	Wisconsin
(MurtiRawat et al., 2020)	SVM,KNN,LR,ANN	ANN	-	99.3	Wisconsin
(Selvathi and Aarth, 2017)	SVM,KNN	SVM	SAE	98.9	Wisconsin
(Lg and At, 2013)	SVM,ANN	ANN	-	95.7	Private
(Naji et al., 2021)	KNN,LR,SVM,CNN	CNN	Python	97.2	Wisconsin

Table 2. Images in Each Dataset and Each Class after Using GAN Algorithm

Class	Benign				Malignant			
	Adenosis	Fibroadenoma	Phyllodes Tumor	Tubular Adenoma	Carcinoma	Lobular Carcinoma	Mucinous Carcinoma	Papillary Carcinoma
100X	1250	1250	1250	1250	1250	1250	1250	1250
200X	1250	1250	1250	1250	1250	1250	1250	1250
400X	1250	1250	1250	1250	1250	1250	1250	1250
40X	1250	1250	1250	1250	1250	1250	1250	1250

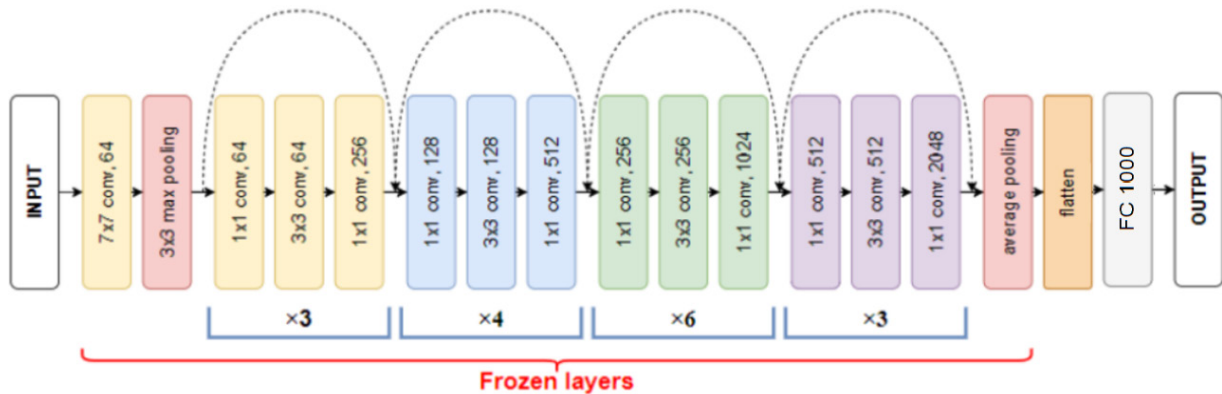


Figure 4. Shows the Structure of the Pre-Trained Model ResNet50

Table 3. Experimental Setup for All Models

Model	BCCNN, ResNet50, VGG16, Xception, InceptionV3, MobileNet
Libraries	Keras, Tensorflow, matplotlib, OS, sklearn
Classes	8
Class Mode	Categorical
Input Layer	75x75x3
Neurons	3*3 Convolutional, 32 filter
Accuracy	F1-score
Pooling	3*3 Convolutional, 64 filter
Max Pooling	Fully
Flatten	Fully
Activation Function	Softmax
Loss	Categorical crossentropy
Learning Rate	0.0001
Batch Size	128
Epochs	120
Optimizer	Adam
Output Layer	8

of images from other domains as shown in Figure 5.

Xception

Xception was “proposed by François Chollet, the creator and chief maintainer of the Keras library” (Bovik, 2005). The architecture Xception contains 36 Convnet layers making the feature abstraction base of the network. Except for the first and last modules, “the 36 convolutional layers are divided into 14 modules, all of which contain linear residual connections around them” (Metwally et al., 2018). Xception was recommended by François Chollet, the Keras library’s founder and major maintainer (Miaoa and Miaoa, 2018) (Figure 6).

InceptionV3

InceptionV3 on the ImageNet dataset, is “a widely used image recognition model that has been found to achieve higher than 78.1 percent accuracy” (Alkronz et al., 2019). It has 42 layers and can categorize images to 1000 different object categories: Devices, laptops, pens, and a variety of animals are among the items on display as shown in Figure 7. “Thus, it has built a library of rich

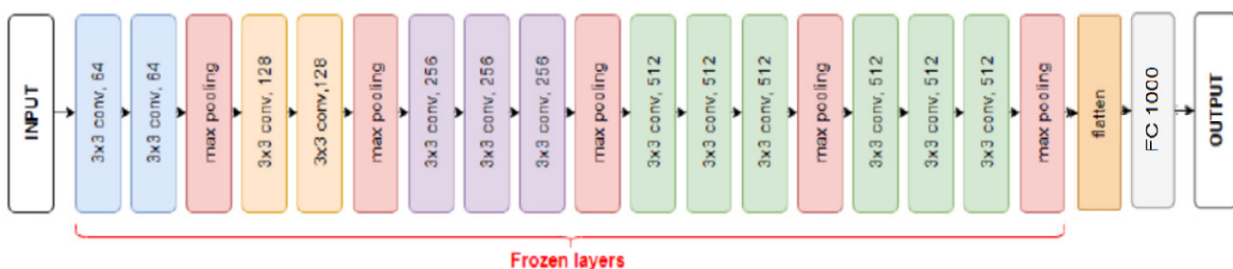


Figure 5. Shows the Structure of the Pre-Trained Model VGG16

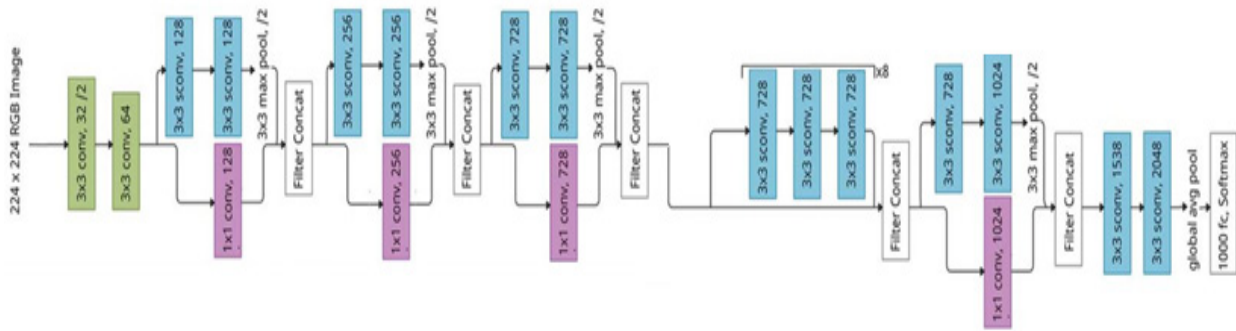


Figure 6. Shows the Structure of the Pre-Trained Model Xception

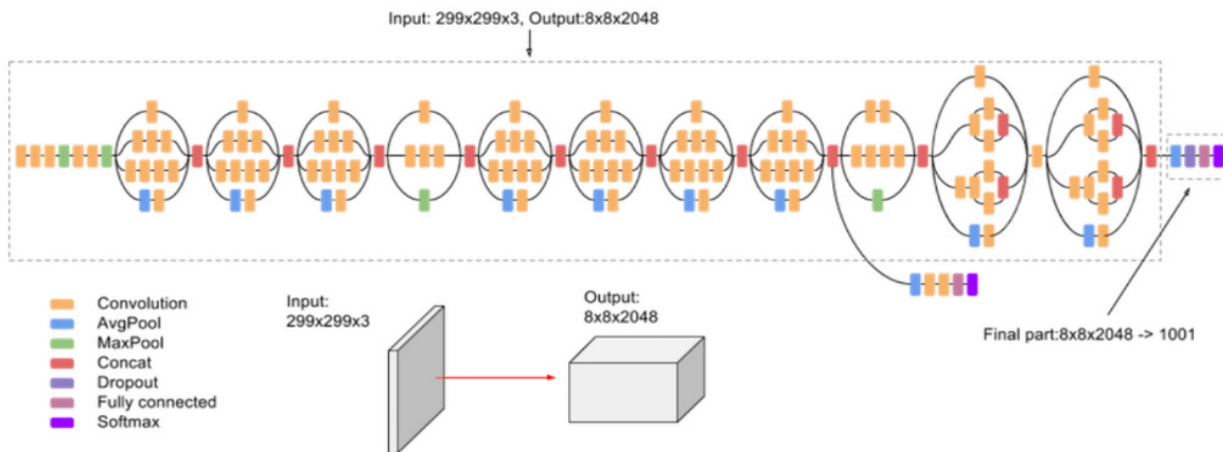


Figure 7. Shows the Structure of the Pre-Trained Model InceptionV3

feature demonstrations for an extensive variety of images. The input image size is 299 by 299 pixels. Convolutions, average pooling, max pooling, concats, dropouts, and fully linked layers are among the symmetric and asymmetric building components in the model. Batch normalization is applied to activation inputs and is used extensively throughout the model. Loss is calculated using Softmax” (Muntasa and Yusuf, 2019).

MobileNet

MobileNet is a plain 53-layer for mobile vision applications, a deep convolutional neural network is used. That is efficient and computationally light. Object detection, fine-grained classifications, facial traits, and localization are just a few of the real-world applications that use MobileNet. Other models exist, but what makes MobileNet unique is that it requires relatively little

computational resources (Sergey and Christian, 2015).

Figure 8 Illustration (A) the overall MobileNet architecture and (B) an in-depth explanation of the Depth-wise Separable (DS) convolutions layer.

Experimental Setup

Table 3 down below shows the experimental setup for each model used. All experiments were running on a Google Colab lab environment (online). Every computer is equipped with Gaming motherboard, 13 GB RAM, and 60 GB Hard Disk space. Using this powerful computer helped to run all the experiments for the whole iterations in an average of 2 hours per testing experiment. Figure 9 shows the structure of the 30 experiments applied in this study.

In Figure 9 the flowchart of each experimental scenario has been presented. In the first 6 experiments, we trained all 6 deep learning models using dataset 40X. The second

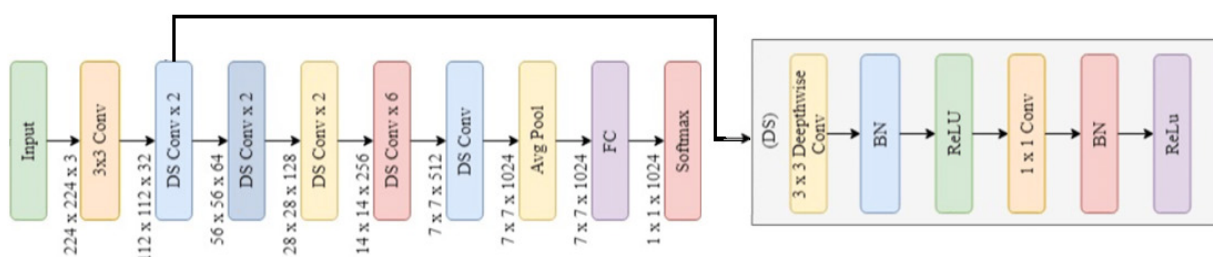


Figure 8. (A) the overall MobileNet architecture and (B) an in-depth explanation of the Depth-wise Separable (DS) convolutions layer

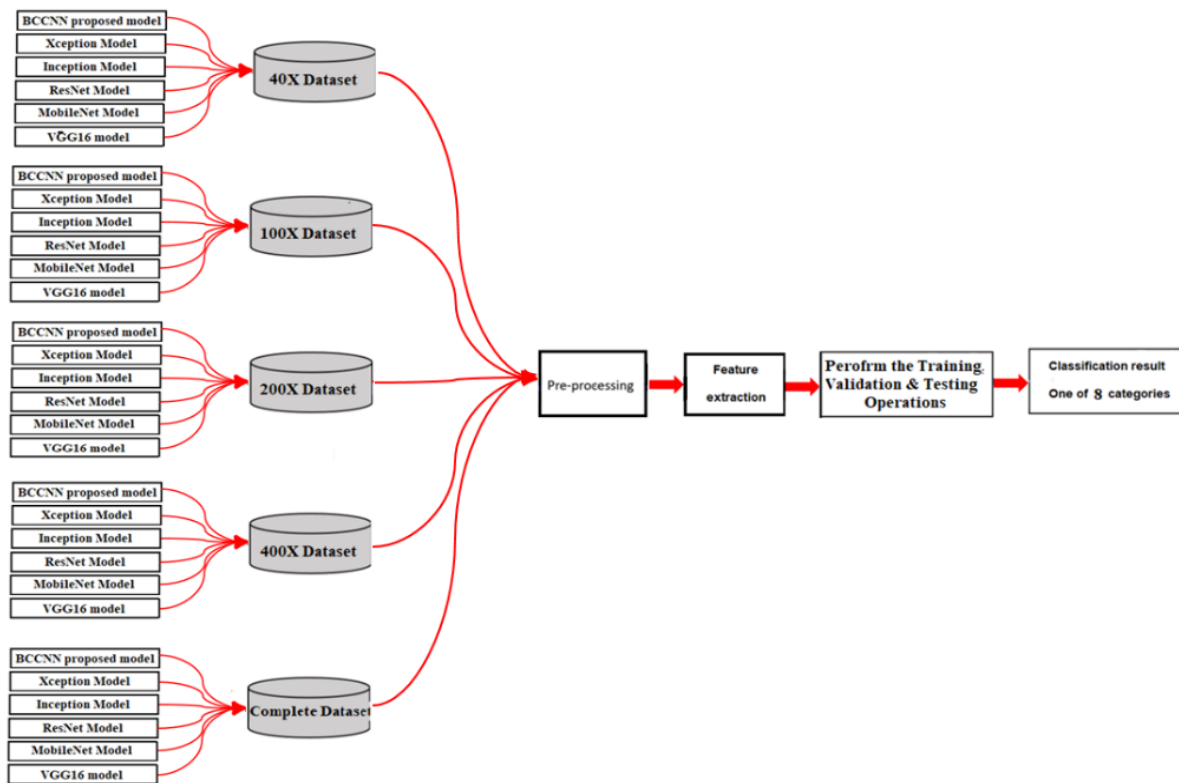


Figure 9. Shows the Structure of the 30 Experiments Applied in This Study

6 experiments, 100X dataset was used, for the third 6 experiments, 200X dataset was used, for the fourth 6 experiments 400X dataset was used and for the last 6 experiments, the Complete Dataset (40X, 100X, 200X, 400X) was used.

Training and Validation of Each Model

We trained the proposed model BCCNN and the five pre-trained model 120 epochs (60 epochs with data augmentation to overcome the problem of overfitting and 60 epochs without data augmentation because the dataset at hand is not that big and we are prone to overfitting problem). Training and validating the proposed model and the five pre-trained models used the same hyper-

parameters to enable us to compare their performance and know which model is the best one for the classification of breast cancers. While training each model, the loss and accuracy metrics are recorded down for comparison analysis.

Results

In each experiment, we used a different dataset. The datasets were organized according to the magnification of the images. There are 4 magnifications of the images: 40X, 100X, 200X, and 400X. Each magnification represents one dataset. Additionally, we grouped all magnifications (40X, 100X, 200X, 400X) into one dataset called All

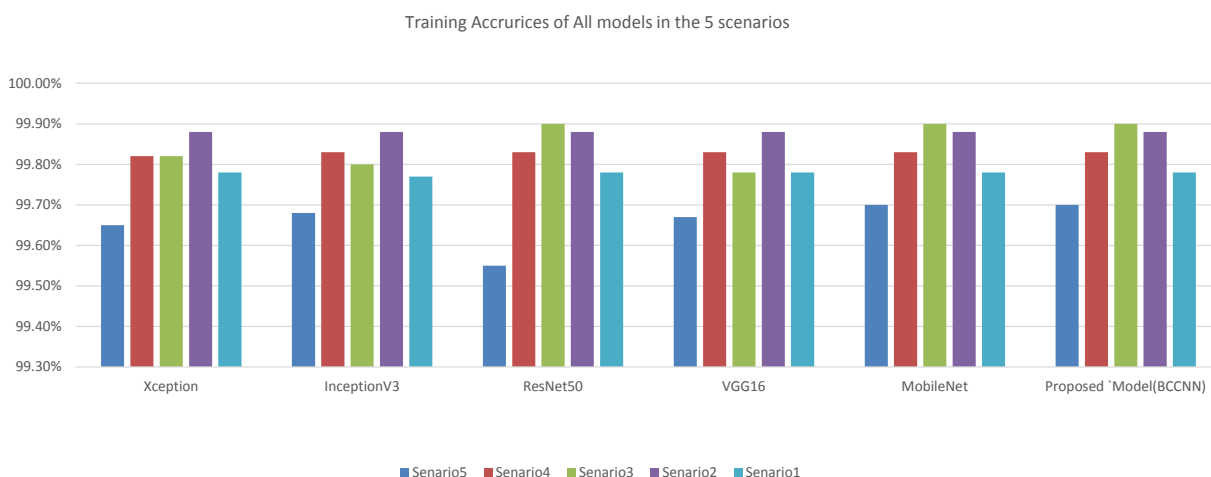


Figure 10. Training Accuracies for the 6 Models in the 5 Scenarios

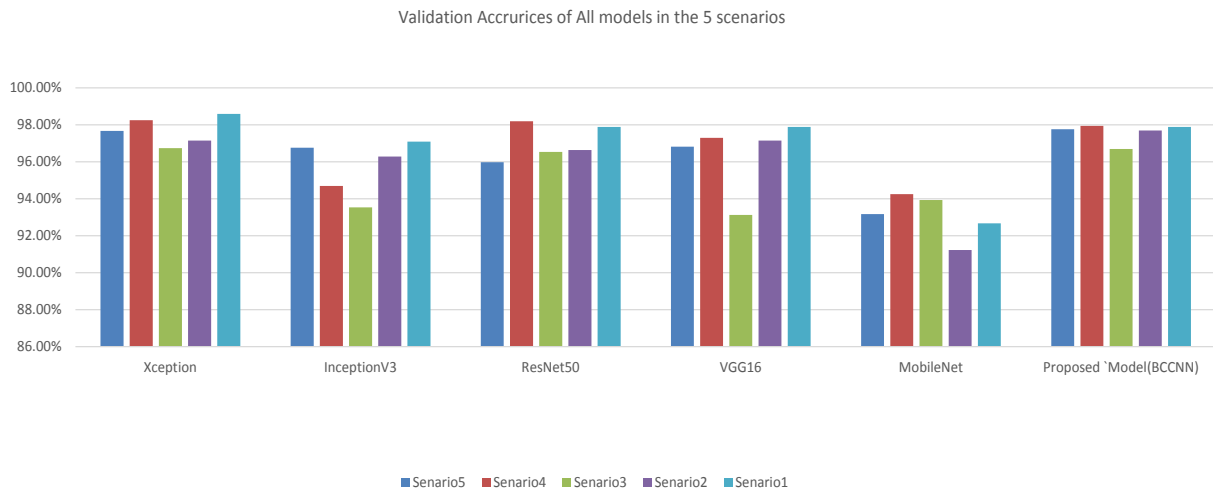


Figure 11. Validation Accuracies for the 6 Models in the 5 Scenarios

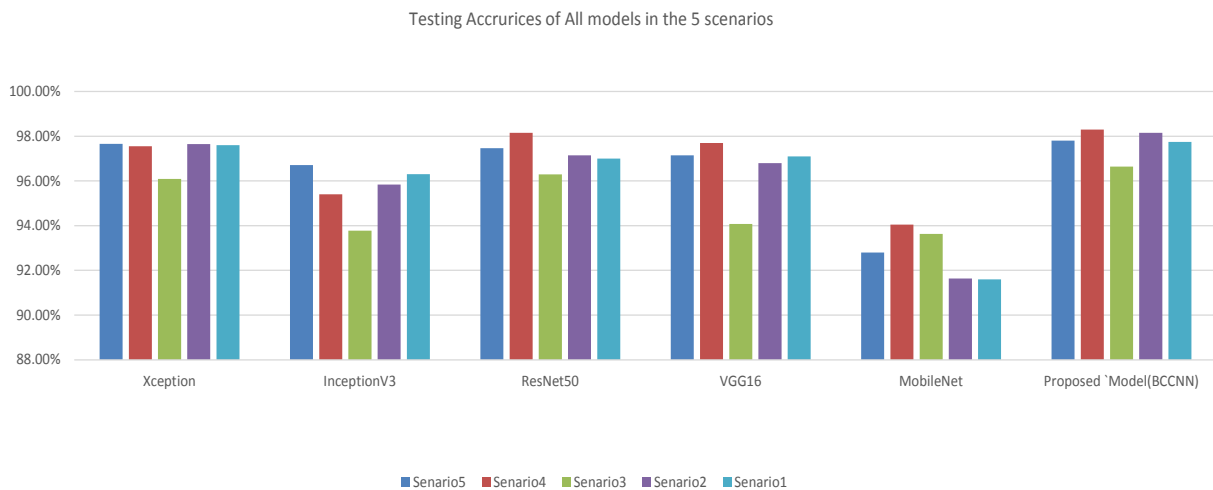


Figure 12. Testing Accuracies for the 6 Models in the 5 Scenarios

Datasets Together.

Thus, we have 5 datasets to experiment with and see which dataset gives better performance in terms of F1-score, Recall, Precision, Accuracy, and Time required for training, validating and testing.

For each dataset, we fine-tuned 5 pre-trained deep learning models for the detection and classification of breast cancers. Additionally, we proposed a new deep learning model called BCCNN for the detection and classification of breast cancers. We trained, validated, and

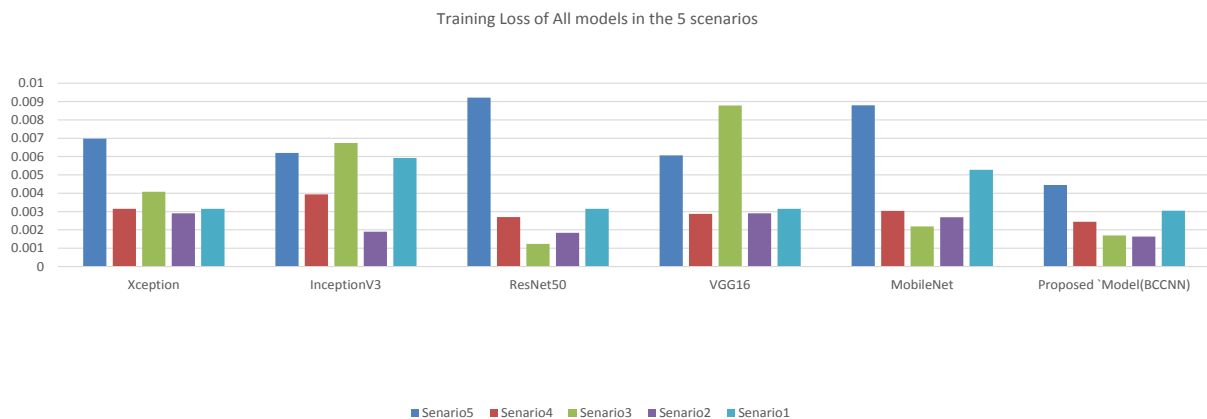


Figure 13. Training Loss for the 6 Models in the 5 Scenarios

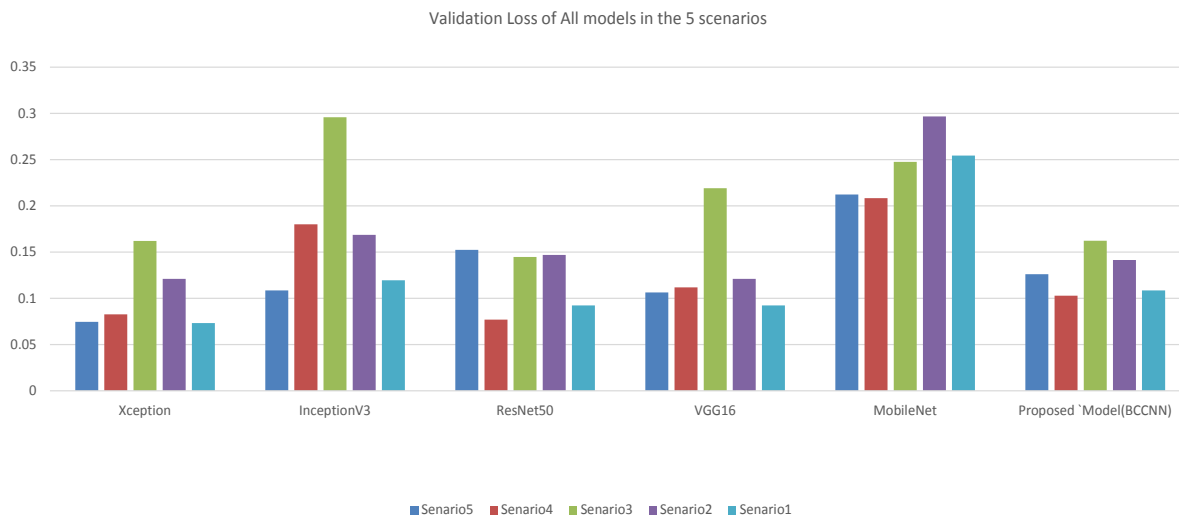


Figure 14. Validation Loss for the 6 Models in the 5 Scenarios

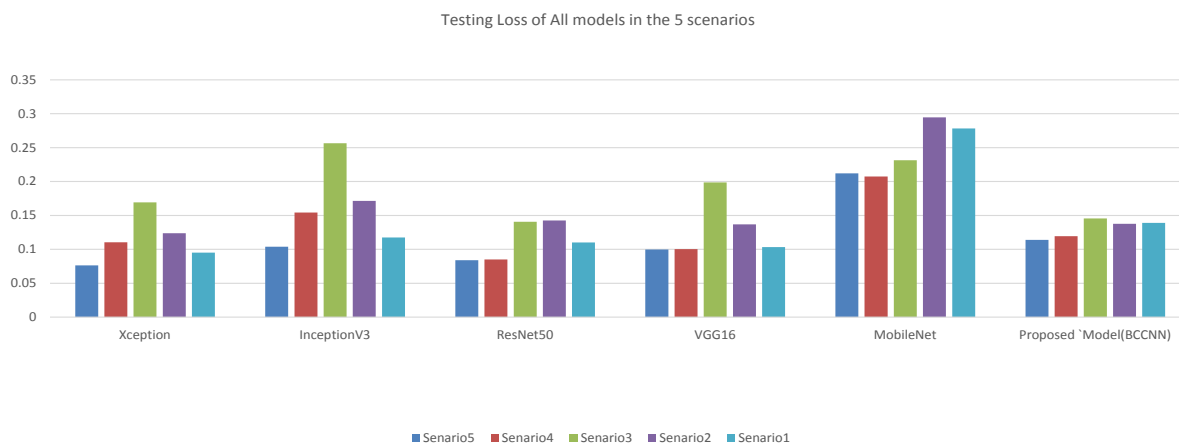


Figure 15. Testing Loss for the 6 Models in the 5 Scenarios

tested all models in each dataset and reported the results.

In this section we will compare the results of each dataset (each scenario) with other scenarios.

Figure 10 shows the training accuracy for the 6 models in the 5 scenarios. The best training accuracy in all scenarios is close to 99%; however the best validation accuracy comes from Xception (98.59%) and proposed

models (97.89%) as seen in Figure 11.

From Figure 12 the best testing accuracy comes from the proposed model (97.80%) followed by the Xception model (97.66%). The best loss of training of all scenarios comes from the proposed model (0.00164) followed by Xception (0.00290) as in Figure 13. The best validation Loss comes from the Xception model (0.07455) followed

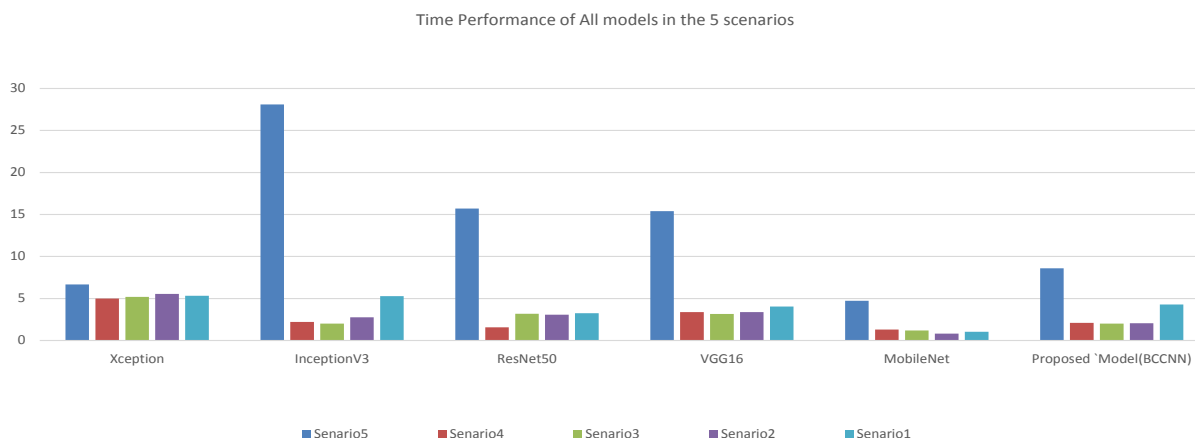


Figure 16. Time Performance for the 6 Models in the 5 Scenarios

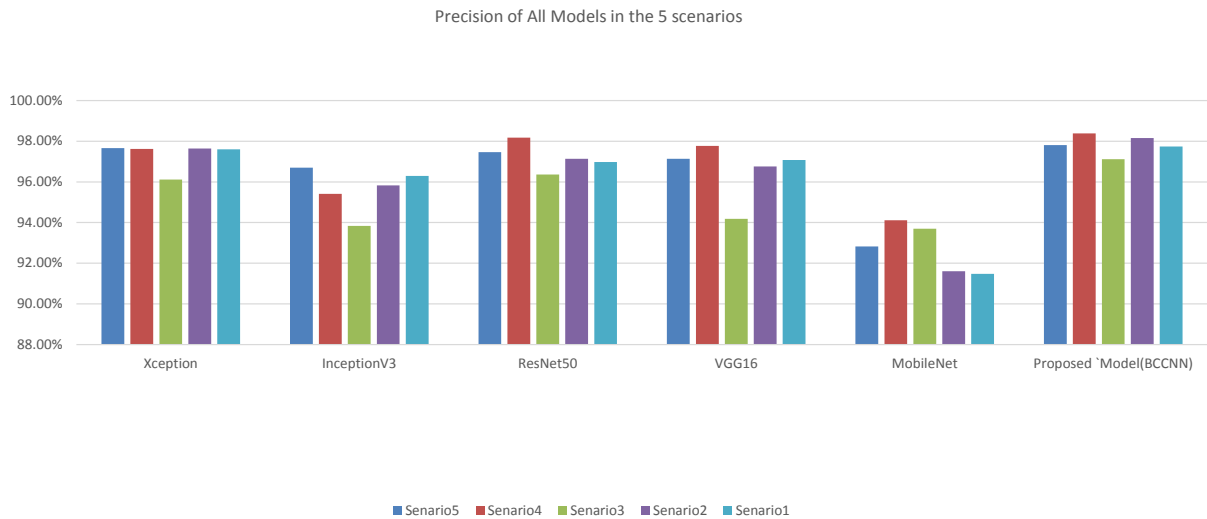


Figure 17. Precision for the 6 Models in the 5 Scenarios

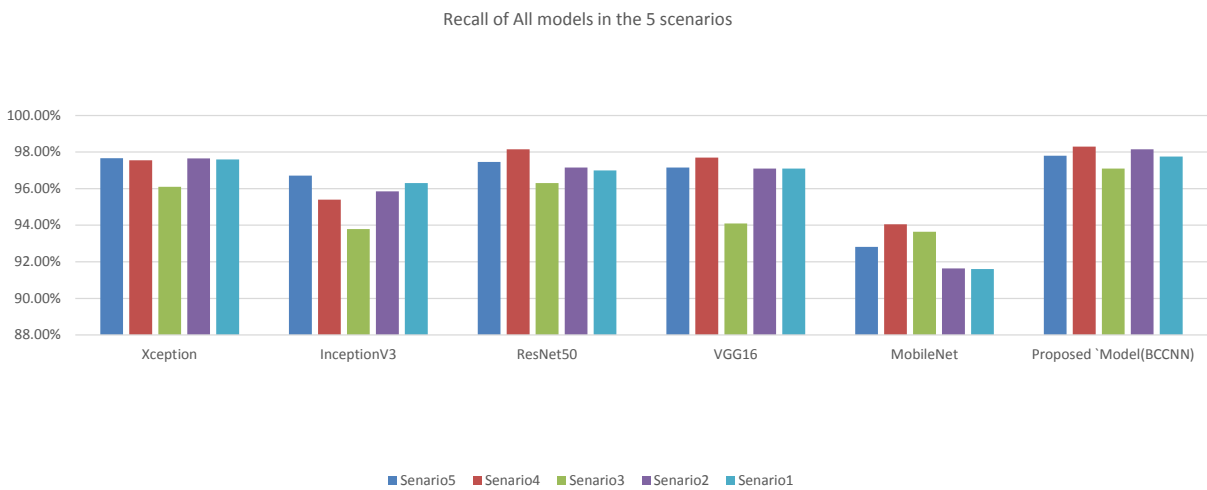


Figure 18. Recall for the 6 Models in the 5 Scenarios

by the proposed model (0.10285) as shown in Figure 14. The best testing loss comes from the Xception model (0.07635) followed by the proposed model (0.11411) as

shown in Figure 15. The best time performance (Time needed for the testing) comes from MobileNet model (1.2) seconds followed by Proposed model (2) seconds

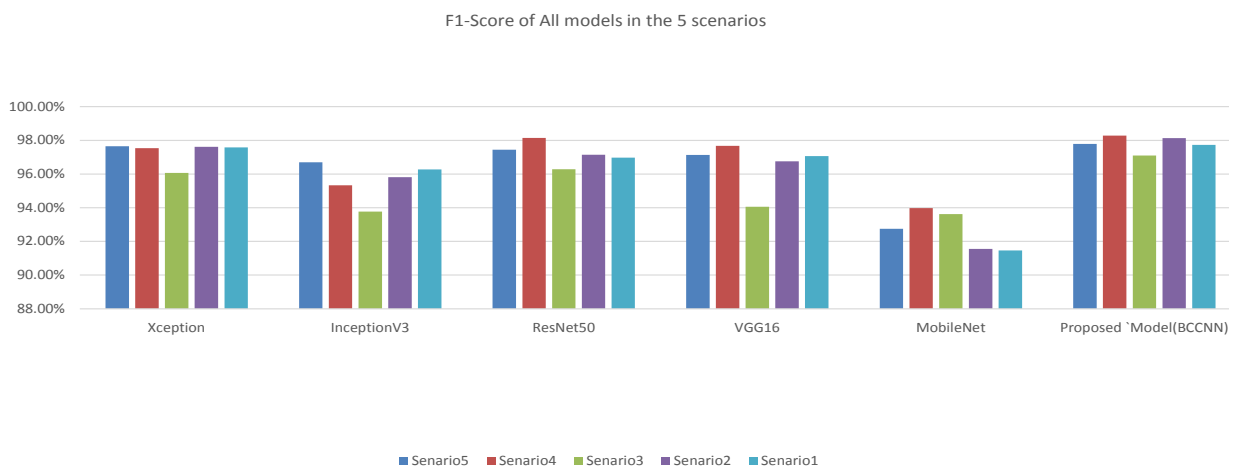


Figure 19. F1-Score for the 6 Models in the 5 Scenarios

as shown in Figure 16. The best Precision comes from the proposed model (98.39%) followed by the Xception model (97.66%) as in Figure 17.

The best Recall comes from the proposed model (98.30%) followed by the Xception model (97.66%) as in Figure 18. The best F1-score comes from the proposed model (98.28%) followed by the Xception model (97.65%) as in Figure 19. From the previous accuracies such as Precision, Recall, F1-score and time performance, it seems that the best model is the proposed model (BCCNN).

Discussion

The focus of this study was on the breast cancer detection and classification based on different datasets (40X, 100X, 200X, 400X, All Datasets together), by balancing the datasets and evaluating the proposed deep learning model and using 5 pre-trained models (Xception, Inceptions, Resnet, VGG16, MobileNet) to compare the result with. The proposed model (BCCNN) achieved the highest accuracy (98.30%), Recall (98.30%), Precision (98.39%), F1-Score (98.28%) and time performance (2.10 seconds). The best scenario results were the fourth scenario with a dataset (400X). These results are promising and can be applied in different applications in various human computer interaction domains.

Author Contribution Statement

The authors confirm contribution to the paper as follows: study conception and design: BA, MR, IZ; data collection: BA, IZ, MR, SA; analysis and interpretation of results: BA. Author; draft manuscript preparation: BA, SA, IZ. All authors reviewed the results and approved the final version of the manuscript.

Acknowledgements

General: The authors would like to thank all staff of the University of Computer Science and Engineering (UNIMY) for their support. The authors would like to thank Staff of Kaggle depository for providing Breast cancer Images for the researchers. Furthermore, the authors would like to thank Google Staff for permitting us to use their computer Labs (Google Colab) for carrying out the Experiments of Breast cancer Free.

Approval

This work was approved by my academic advisors in the University of Computer Science and Engineering (UNIMY), Malaysia.

Ethical Declaration

This study is part of an approved student thesis.

Data Availability

The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

- Abu Ghosh MM, Atallah RR, Abu Naser SS (2016). Secure mobile cloud computing for sensitive data: Teacher services for Palestinian higher education institutions. *Int J Grid Distrib Comput*, **9**, 17–22.
- Abu Naser SS (2009). Evaluating the effectiveness of the CPP-Tutor, an intelligent tutoring system for students learning to program in C++. *J Appl Sci Res*, **5**, 109-114.
- Abu-Naser SS, El-Hissi H, Abu-Rass M, El-khozondar N (2010). An expert system for endocrine diagnosis and treatments using JESS. *J Artif Intell*, **3**, 239-51.
- Abunasser BS, AL-Hiealy MRJ, Barhoom AM, et al (2022). Prediction of Instructor Performance using Machine and Deep Learning Techniques. *Int J Adv Comput Sci Appl*, **13**, 78-83.
- Abunasser BS, AL-Hiealy MRJ, Zaout IZ, et al (2022). Breast Cancer Detection and Classification using Deep Learning Xception Algorithm. *Int J Adv Comput Sci Appl*, **13**, 223-8.
- Agarap AF (2018). On breast cancer detection: an application of machine learning algorithms on the Wisconsin diagnostic dataset. *ArXiv*, abs/1711.07831.
- Ahmed MS, Sayeed A, Mallick T, Syfuddin HM (2020). Knowledge and Practices on Breast Cancer among Bangladeshi Female University Students: A Cross-sectional Study. *Asian Pac J Cancer Care*, **5**, 19-25.
- Ak M (2020). Comparative Analysis of Breast Cancer Detection and Diagnosis Using Data Visualization and Machine Learning Applications. *Healthcare*, **8**, 1-11.
- Alayoubi MM, Arekat ZM, Al Shobaki MJ, Abu-Naser SS (2022). The Impact of Work Stress on Job Performance among Nursing Staff in Al-Awda Hospital. *Found Manag*, **14**, 89-110.
- Albatish IM, Abu-Naser SS (2019). Modeling and controlling smart traffic light system using a rule based system. *Proceedings - 2019 International Conference on Promising Electronic*.
- Almasri A, Obaid T, Abumandil MSS, et al (2023). Mining Educational Data to Improve Teachers' Performance. In: Al-Emran, M., Al-Sharafi, M.A., Shaalan, K. (eds). *International Conference on Information Systems and Intelligent Applications*. ICISIA 2022. Lecture Notes in Networks and Systems, 550. Springer, Cham.
- Almasri AR, Yahaya NA, Abu-Naser SS (2022). Instructor Performance Modeling For Predicting Student Satisfaction Using Machine Learning-Preliminary Results. *J Theor Appl Inf Technol*, **100**, 5481-96.
- Alzamily JY, Ariffin SB, Abu-Naser SS (2022). Classification of Encrypted Images Using Deep Learning –ResNet50. *J Theor Appl Inf Technol*, **100**.
- Arqawi S, Atieh KAFT, Shobaki MJAL, et al (2020). Integration of the dimensions of computerized health information systems and their role in improving administrative performance in Al-Shifa medical complex. *J Theor Appl Inf Technol*, **98**, 1087–119.
- Arqawi SM, Abu Rumman MA, Zitawi EA, et al (2022). Predicting Employee Attrition and Performance Using Deep Learning. *J Theor Appl Inf Technol*, **100**.
- Arqawi SM, Zitawi EA, Rabaya AH, et al (2022). Predicting University Student Retention using Artificial Intelligence. *Int J Adv Comput Sci Appl*, **13**, 315–24.
- Asri H, Mousannif H, Moatassime H, Noël T (2016). Using Machine Learning Algorithms for Breast Cancer Risk

- Prediction and Diagnosis. *Procedia Comput Sci*, **83**, 1064–9
- Barhoom AM, Al-Hiealy MRJ, Abu-Naser SS (2022). Deep Learning-Xception Algorithm for Upper Bone Abnormalities Classification. *J Theor Appl Inf Technol*, **100**.
- Barhoom MA, Al-Hiealy MR, Abu-Naser SS (2022). Bone Abnormalities Detection and Classification Using Deep Learning-Vgg16 Algorithm. *J Theor Appl Inf Technol*, **100**, 6173-84.
- Cahoon TC, Sutton MA, Bezdek JC (2000). Breast cancer detection using image processing techniques, Ninth IEEE International Conference on Fuzzy Systems. *FUZZ- IEEE 2000 (Cat. No.00CH37063)*, **2**, 973-6.
- Communication (ICSPC), 359-363, IEEE 2017.
- El-Habil BY, Abu-Naser SS (2022). Global Climate Prediction Using Deep Learning. *J Theor Appl Inf Technol*, **100**.
- Gayathri B, Sumathi C, Santhanam T (2013). Breast cancer diagnosis using machine learning algorithm a survey. *Int J Distrib Parallel Syst*, **4**, 39-50.
- Gonzalez-Angulo A, Morales-Vasquez F, Hortobagyi GN (2007). Overview of resistance to systemic therapy in patients with breast cancer. *Adv Exp Med Biol*, **608**, 1-22.
- Gore S, Yelake V, Dudhe M (2022). Knowledge and Practice of Breast Self-examination among Women in Yavataml, India, A Cross- Sectional Study. *Asian Pac J Cancer Care*, **7**, 61-4.
- Hafemann L, Oliveira L, Cavalin P (2014). Forest species recognition using deep convolutional neural networks. In *Pattern Recognition (ICPR), 2014 22nd International Conference, IEEE*, pp 1103-7.
- Hashemi EA, Haghight S, Olfatbakhsh A, Jafari M, Yasaei M (2021). Evaluation of Precancerous Breast Lesion Upgrade Rate. *Asian Pac J Cancer Care*, **6**, 461-6.
- Jafari S, Saadatpour Z, Salmaninejad A, et al (2018). Breast cancer diagnosis: Imaging techniques and biochemical markers. *J Cell Physiol*, **233**, 5200-13.
- Joo S, Yang Y, Moon W, Kim H (2004). Computer-Aided Diagnosis of Solid Breast Nodules: Use of an Artificial Neural Network Based on Multiple Sonographic Features. *IEEE Trans Med Imaging*, **23**, 1292-300.
- Joshi A, Mehta D (2017). Comparative Analysis of Various Machine Learning Techniques for Diagnosis of Breast Cancer. *Int J Emerg Technol*, **8**, 522–26.
- Kiyan T, Yildirim T (2016). Breast cancer diagnosis using statistical neural networks. *J Electr Electron Eng*, **4**, 2004.
- Kumar UK, Nikhil MBS, Sumangali K (2017). Prediction of breast cancer using voting classifier technique. 2017 IEEE International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials (ICSTM), pp 108-14.
- LG A, AT E (2017). Using Three Machine Learning Techniques for Predicting Breast Cancer Recurrence. *J Health Med Inform*, **4**.
- Mady SA, Arqawi SM, Al Shobaki MJ, et al (2020). Lean manufacturing dimensions and its relationship in promoting the improvement of production processes in industrial companies. *Int J Emerg Technol*, **11**, 881–96.
- Mousannif H, Asri H, Moatassime HA, Noel T (2017). Using Machine Learning Algorithms for Breast Cancer Risk Prediction and Diagnosis. *Procedia Comput Sci*, **83**, 1064–9.
- MurtiRawat R, Panchal S, Singh VK, Panchal Y (2020). Breast Cancer Detection Using K-Nearest Neighbors, Logistic Regression and Ensemble Learning. International Conference on Electronics and Sustainable Communication Systems (ICESC), pp 534-40.
- Naji MA, Filali SE, Aarika K, et al (2021). Machine Learning Algorithms for Breast Cancer Prediction and Diagnosis. *Procedia Comput Sci*, **191**, 487–92.
- Naser SSA (2006). Intelligent tutoring system for teaching database to sophomore students in Gaza and its effect on their performance. *Inf Technol J*, **5**, 916-22.
- Naser SSA (2008). Developing an intelligent tutoring system for students learning to program in C++. *Inf Technol J*, **7**, 1051-60.
- Naser SSA (2008). Developing visualization tool for teaching AI searching algorithms. *Inf Technol J*, **7**, 350-5.
- Naser SSA (2008). JEE-Tutor: An intelligent tutoring system for java expressions evaluation. *Inf Technol J*, **7**, 528-32.
- Olayide A, Isiaka A, Ganiyu R, et al (2021). Demographic Pattern, Tumor Size and Stage of Breast Cancer in Africa: A Meta-analysis. *Asian Pac J Cancer Care*, **6**, 477-92.
- Omondigbe A, Shanmugam V, Amandeep S (2019). Machine Learning Classification Techniques for Breast Cancer Diagnosis. **5**.
- Patil S, Moafa IH, Mosa Alfaifi M, et al (2020). Reviewing the Role of Artificial Intelligence in Cancer. *Asian Pac J Cancer Biol*, **5**, 189-99.
- Pavithra DR, Preethi S, SriRakshitha AK (2021). Breast Cancer Classification using the Supervised Learning Algorithms. 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS), pp1492-8.
- Ponraj N, Jennifer E, Poongodi P, Manoharan S (2012). Morphological operations for the mammogram image to increase the contrast for the efficient detection of breast cancer. *Eur J Sci Res*, **68**, 494-505.
- Pritom A, Munshi M, Sabab S, Shihab S (2016). Predicting breast cancer recurrence using effective classification and feature selection technique. 2016 19th International Conference on Computer and Information Technology (ICCIT), pp 310-4.
- Ragab DA, Sharkas M, Marshall S, Ren J (2019). Breast cancer detection using deep convolutional neural networks and support vector machines. *Peer J*, **7**, 20-31.
- Sachdeva S, Mangalesh S, Dudani S (2021). Knowledge, Attitude and Practices of Breast Self-Examination amongst Indian Women: A Pan-India Study. *Asian Pac J Cancer Care*, **6**, 141-7.
- Saleh A, Sukaik R, Abu-Naser SS (2020). Brain tumor classification using deep learning. Proceedings - 2020 International Conference on Assistive and Rehabilitation Technologies. iCareTech, 131–6.
- Salem H, Attiya G, El-Fishawy N (2017). Early diagnosis of breast cancer by gene expression profiles. *Pattern Anal Appl*, **20**, 567–78.
- Selvathi D, Poornila A (2017). Performance Analysis of Various Classifiers on Deep Learning Network for Breast Cancer Detection International Conference on Signal Processing and Communication (ICSPC), pp 359-63, IEEE 2017.
- Shen L, Margolies L, Rothstein J (2019). Deep Learning to Improve Breast Cancer Detection on Screening Mammography. *Sci Rep*, **9**, 24-35.
- Spanhol F, Oliveira L, Petitjean C, Heutte L (2016). Breast cancer histopathological image classification using Convolutional Neural Networks, **8**, 20-30.
- Tajbakhsh N, Shin J, Gurudu S, et al (2016). Convolutional neural networks for medical image analysis: Full training or fine tuning?. *IEEE Trans Med Imaging*, **35**, 1299-1312.
- Tiwari M, Bharuka R, Shah P, Lokare R (2020). Breast Cancer Prediction Using Deep Learning and Machine Learning Techniques. Electronic Journal Published.
- Wadkar K, Pathak P, Wagh N (2019). Breast cancer detection using Ann network and performance analysis with SVM. *Int J Comput Eng Technol*, **10**.
- Waks A, Winer E (2019). Breast Cancer Treatment: A Review. *JAMA*, **321**, 288-300.
- Westerman F (2002). Classification and diagnostic prediction of pediatric cancers using gene expression profiling and

artificial neural networks. *Net Med*, **7**, 673-9.

Widiana I, Irawan H (2020). Clinical and Subtypes of Breast Cancer in Indonesia. *Asian Pac J Cancer Care*, **5**, 281-5.

World health organization. <https://www.who.int/>.

Yasmin M, Sharif M, Mohsin S (2013). Survey paper on diagnosis of breast cancer using image processing techniques. *Res J Recent Sci*, **2**, 25-35.



This work is licensed under a Creative Commons Attribution-Non Commercial 4.0 International License.