Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X
IMPACT FACTOR: 7.056

IJCSMC, Vol. 14, Issue. 10, October 2025, pg.121 – 134

A Novel Deep CNN Model For Breast Cancer Classification and Localisation Using Parallel Hyper-Parameter Optimisation Techniques

Karthik M*[1]; Thangavel K[2]; Sasirekha K[3]

[1*]Department of Computer Science, Periyar University, Salem, India
[2]Department of Computer Science, Periyar University, Salem, India
[3]Department of Computer Science, Periyar University, Salem, India
[1*] mrjrkvgm@periyaruniversity.ac.in; [2] drktvelu@periyaruniversity.ac.in; [3] ksasirekha7@periyaruniversity.ac.in

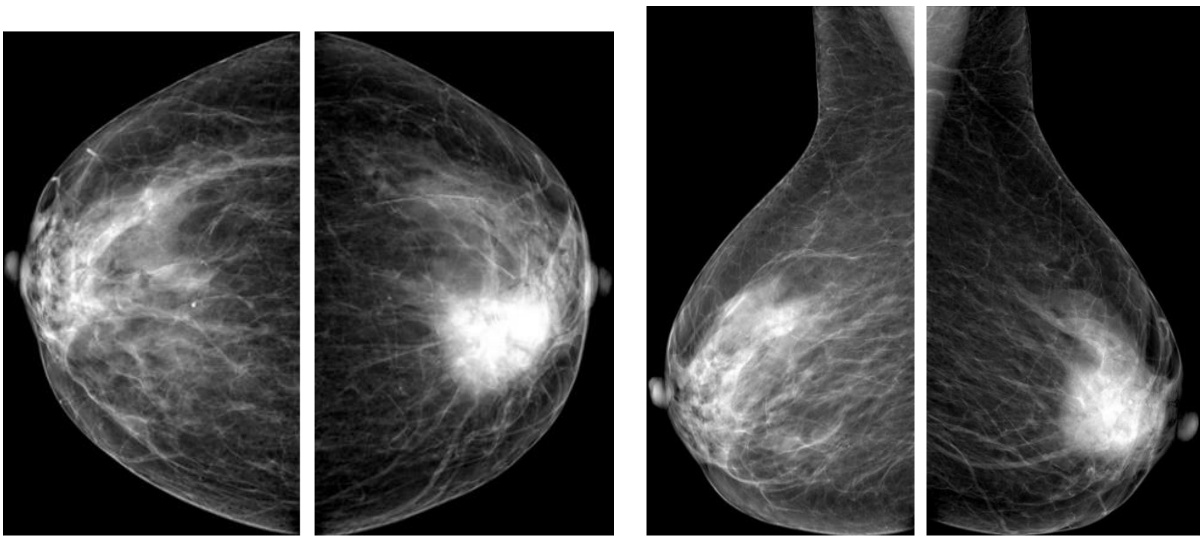
DOI: https://doi.org/10.47760/ijcsmc.2025.v14i10.011

Abstract: The evolution of Deep Learning (DL) enables the solutions for numerous real-time problems, including the finding of malignancies in the medical domain. The Convolutional Neural Network (CNN) delivers attractive and robust results in terms of classifying and detecting useful patterns from images that help radiology experts to detect Breast Cancer (BrC) very accurately and on time. A novel Multi-Output based DeepCNN model with parallel hyper-parameter optimisation using Random Search and Bayesian Search methods is proposed in this study. The proposed model that not only classifies the Mammogram images as normal, benign, and malignant but also determines the location of tumors and malignancies with an optimised approach. The tuning of the hyper-parameters process involves two different stages and parallel tuning of classification and localisation of the proposed model. The first stage involves tuning hyper-parameters in hidden layers, namely the number of filters, learning rate, number of epochs, batch size, and hidden activation. The second stage of hyper-parameter tuning involves the dense and classification layers that include dropout rate, model optimiser, and output activation. Further, the performance of the proposed deepCNN model is compared with six state-of-art transfer learning models, DenseNet169, EfficientNetB4, InceptionV3, ResNet101V2, VGG19, and Xception. Finally, the results of all the models are evaluated using metrics such as precision, recall, F1 score, and IoU that show the proposed model achieves the best test accuracy rate of 96.39% and IoU accuracy of 84.41% using the Bayesian Search approach as compared with Random Search approach and pre-trained models with the same configuration.

Keywords: deep learning, convolutional neural network, breast cancer, random search, bayesian search, mammogram images.

I. INTRODUCTION

In 2020, the World Health Organization (WHO) announced that breast cancer cases exceeded lung cancer in terms of finding new cancer cases compared with 2.21 million lung cancer to 2.26 million breast cancer [1]. Early detection of breast cancer will save millions of lives, as well as the financial stability of their families is endangered. Breast cancer can occur in anyone with or without any underlying medical history or condition and usually has increased risk in women compared to men who inherit mutations to specific genes, such as BRCA1 and BRCA2 [2, 3]. Tumors in Breast cancer are generally classified as benign and malignant [4]. In most cases, benign tumors are caused due to abnormalities of epithelium cells. They cannot grow further, but in some rare cases, they turn out to be cancerous and grow abnormally known as cancer [5]. Among the various screening techniques available to detect breast cancer, the mammogram imaging technique is quite easy to access and reliable worldwide. For each breast, mammography allows for the recording of two views: the mediolateral oblique (MLO) view, which is a side view, and the craniocaudal (CC) view, which is a top-to-bottom view, as shown in Figure 1a to Figure 1d. The modern era of Artificial Intelligence (AI) and the recurrent development of Machine Learning (ML) and deep learning algorithms provide solutions to many real-time problems that were unsolved previously, particularly in the medical domain. Some of the top algorithms like K-Nearest Neighbour (KNN), Support Vector Machine (SVM), Naïve Bayes (NB), Random Forest(RF), and Deep Convolutional Neural Networks (DCNN) produce the best performance on image classification and object detection problems [6]. However, DeepCNN gives the predominant performance compared to classical machine learning algorithms due to its ability to automatic feature extraction and learning of complex features in the huge volume of images. The architecture of the typical DeepCNN model is shown in Figure. 2. It contains the principle components of an input layer, output layer, and two or more hidden layers that are arranged in a three-dimensional shape ($Width \times Height \times Depth$). Subsequently, the hidden layers are made up of convolutional layers, different types of pooling layers, batch normalisation layers, dense layers, and finally end with, dropout layers as optional [7]. Together with this, the evolution of Transfer Learning [8] makes ImageNet types of DeepCNN classifiers more popular than any other machine learning algorithm for image classification tasks in all domains. Despite the success rate of TL and DeepCNN algorithms, Hyper-Parameter Optimisation (HPO) offers felicity to data scientists to adjust the vital learning parameters in various layers in the model, which improves the stability of the model and provides the finest result in almost all the time [9, 10].



(a). Right-side CC **(b).** Left-side CC **(c).** Right-side MLO **(d).** Left-side MLO

Fig. 1. Craniocaudal(CC) and Mediolateral Oblique(MLO) views of Breast Mammogram

In the past few years, many researchers proposed classical machine learning models as well as various deep learning models to classify the different types of images that diverge from natural images to medical images. The requirement of TL and DeepCNN model is crucial for many fields. The authors of [11] presented work on multi-class breast cancer classification using both ML and DL methods, employed a histopathological image dataset for their breast cancer classification task, and three transfer learning models, VGG16, VGG19, and ResNet50, achieved an accuracy rate above with 93% in all three models. The development of TL not only saves the developer's implementation time but also allows for finding the best solutions for any uncovered domains that do not require knowledge in TL. The authors of [12] implemented the VGG16 and GoogleLeNet models for the Crop types mapping process using Malawi and Mozambique crops datasets, which gives a decent amount of accuracy rate of 83% for the Malawi dataset and 90% for the Mozambique dataset. The authors of

[13] performed to find the Return Oriented Programming (ROP) based cyberattacks in computer chips using the CNN-based DeepReturn model and RNN-based DeepVSA models that attain a high ROP detection rate of 95.2% with a minimal error rate of 0.29%. The authors of [14] established to build four distinct DeepCNN models based on the ensemble techniques to classify breast cancer using two TL models such as VGG16 and VGG19. The authors also used a hyper-parameter optimisation process to optimise the learning rate and batch size of the models to improve the performance of the model. Finally, the authors got an average accuracy rate of 93.51% for fully trained VGG16 with VGG19 models and 95.29% for fine-tuned VGG16 with VGG19 models. The flexibility of TL is very high from using a minimal volume of data and adjusting the configuration of the model structure up to every single parameter in the hidden layers. The authors of [15] published drug response prediction in treated and untreated liver cancer cells using ResNet101 based customised model. They retrained only the last three layers of the model and got the accuracy values. Their proposed model achieved 97.5% accuracy when compared to non-retrained models ResNet18 and InsecptionV3. Most of the researchers applied DeepCNN and TL models only for the classification problem, while few of them had performed object localisation using DeepCNN. The author of [16] provided a special learning approach to the customised DeepCNN model that was specially built for moving object tracking and localisation using natural image datasets. The author used the Center Point Matrix (CPM) and Region of Interest (ROI) to fix the object in the input images. But due to the complexity of finding and localising the objects in the input images (frames), finally, the author claims that the proposed model does not perform well as expected and does not provide the result metrics in the numerical format.

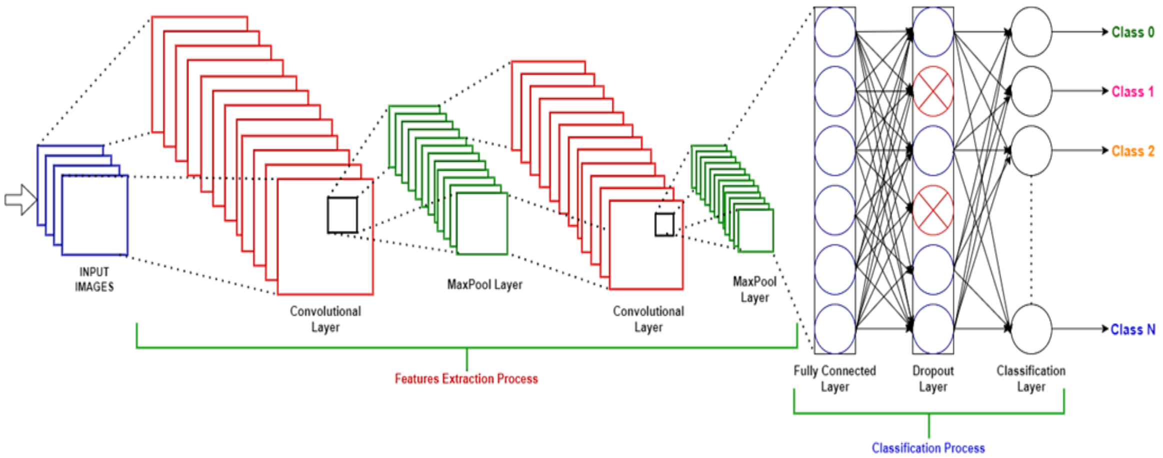


Fig. 2. Typical architecture of DeepCNN Model

Based on the understanding of various authors' works and their implementation nature of TL and DeepCNN models, Few researchers only tried the three class modes normal, benign, and malignant; and the majority of them only used those strategies to categorise breast cancer in binary form. Additionally, we note from earlier studies that only a small number of researchers employ the DeepCNN model alone in conjunction with the localization feature for object categorization. However, due to the complexity of medical images, particularly mammogram images, none of them are capable of handling breast cancer or any other type of medical image data. The following contributions are made in this study to attempt to close these gaps in the field:

i.) To develop a novel Multi-Output DeepCNN model (MODCNN) for tumor localization and breast cancer classification as Normal, Benign, and Malignant classes.

ii.) To use Bayesian search and Random search to implement parallel hyper-parameter tweaking on the suggested model.

iii.) Using the six cutting-edge Transfer Learning models DenseNet169, EfficientNetB4, InceptionV3, ResNet101V2, and Xception, apply the same suggested methodology.

II. METHODS AND MATERIALS

A. *Dataset and Data Augmentation*

Two datasets, the well-known standard INBreast [17] and the recently developed Categorised Digital Database for Low Energy and Subtracted Contrast Enhanced Spectral Mammography (CDD-CESM) [18] that implements Contrast-enhanced spectral mammography imaging modality are used in this study to identify early breast cancer and localise the tumour (benign, malignant) portion in the mammogram image. These two datasets are well organised and follow the breast cancer imaging standards such as Breast Imaging-Reporting and Data System (BI-RADS) [19] and the American College of Radiology (ACR) [20]. Mammogram images are generally classified based on the BI-RADS standard that the assessment carried out by the radiology expert. Here, the INBreast dataset is utilised for model testing, and CDD-CESM is used for the model training process. The second dataset CDD-CESM contains a total of 2006 images, which includes 1003 images in the Low Energy modality and 1003 images in Subtracted Energy modality from 326 women patients. Among them, the 1003 Low Energy images are only considered for this study due to the data commonality that matches with the INBreast dataset. The implementation of Image classification or localisation using DL may require a huge volume of data and may be crucial to attaining a high accuracy rate, but the implementation of various data augmentation techniques, such as Shearing, Zooming, and random erasing, may lead to poor performance of the model due to the nature of medical images, especially the mammogram images [21, 22]. Also, the problem of localisation makes it a very challenging task too to maintain the pixel coordinates in the images. Image flipping (Horizontal) and contrast enhancement may be employed and best suited for both scenarios to increase the volume of the image and fix the object location within the specified pixel coordinates. The data augmentation techniques implemented in this study are shown in Table 1. The total number of images available in the INBreast dataset is 410, from 115 patients. Initially, all the images in the INBreast dataset are de-categorised as per their BI-RADS classes from class 0 to class 6 as Normal, Benign, and Malignant. After de-categorised all 410 images, 272 images are included for the testing process from class 1(Normal – 65 images), class 2(Benign – 224 images), class 4c(suspicion for malignancy – 23 images), class 5 (probably Malignant – 46 images), and 6 (Biopsy proven Malignant – 8 images). In CDD-CESM, all 1003 images are involved in the augmentation process.

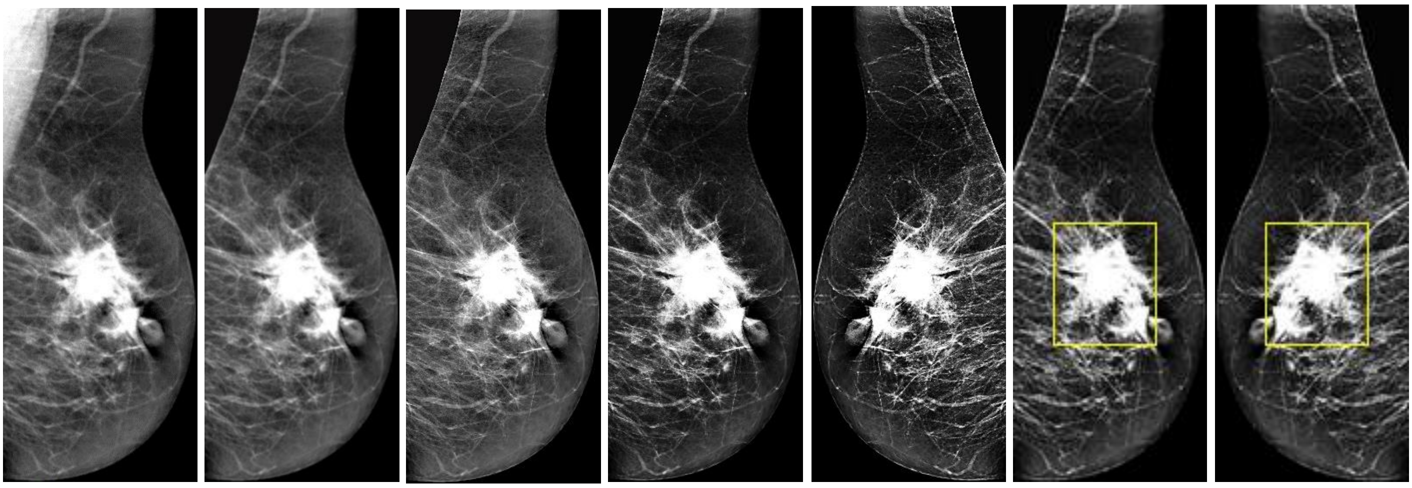
TABLE I
DATA AUGMENTATION ON INBREAST AND CDD-CESM DATASETS

	Dataset Name	**CDD-CESM**	**INBreast**
	Class	**No. of Images**	**No. of Images**
Original Image Distribution	Normal (N)	341	65
	Benign (B)	331	130 (224 – 94)
	Malignant (M)	331	77 (23 + 46 + 8)
Total Images Before Augmentation	**N + B + M**	**1,003**	**272**
Horizontal Flipping	N	682	130
	B	662	130
	M	662	130
Contrast Enhancement using CLAHE	N	1,170	Not Implemented
	B	1,170	Not Implemented
	M	1,170	Not Implemented
Total Images After Augmentation	**N + B + M**	**3,510**	**390**

B. *Data pre-processing*

In any image classification, object detection, or object localisation task, pre-processing plays a vital role that enables the desired model to become a more robust and reliable one [23]. Here, augmenting the data volume from 1,275 images to 3,900 images is a reasonably good incrementation; even though the total number of images is arguably low but the model structure and complexity are based on the data availability, and that should maintain the robustness as a top priority. There are three pre-processing techniques applied to the training set images, which are: image resizing, noise removal, and removal of pectoral breast muscle, which is a very important pre-process in breast mammogram image-based tasks [24, 25]. Once the necessary number of images have been obtained by using image augmentation methods, then the removal of breast pectoral muscle operation is performed by Fusion-based Graywolf and Particle Swarm-based U-Net process, and that should be done both on CDD-CESM and INBreast datasets. After the removal of the pectoral breast muscle, the removal of Poisson and Speckle noises was performed on all the images. Particularly mammogram images are more prone to these kinds of noise compared to Gaussian and Salt and pepper types of noises due to the image acquisition process and imperfections of the imaging system design techniques [26, 27]. The genetic algorithm (GA) [28] based denoise autoencoder technique was employed to remove the poison and speckle noises with high PSNR values of 39.43 against the Poisson noise and 37.91 against the Speckle noise. Before the conversion from RGB to Grayscale mode, resize process should be performed on all the images. The DeepCNN model that performs only Classification may have more freedom interms of image resizing as compared to the model that aims to build image classification with an image object localisation feature. Object localisation is regression kind of task that defines the objects in an image using bounding-box concept [29].

In this study, we limit the number of objects in an image to one for the image localisation problem. Define a single object (tumor) in the image, known as single-object localization [30], because of the availability of tumor regions in train-test image datasets that naturally prevent the redundancy of an object in the same image and same position, particularly for the medical image object localisation [31]. Each image in the CDD-CESM dataset has an average image size of 2355 × 1315 in height and width, but the INBreast dataset that is varying image sizes from one image to another. All the images that have benign and malignant tumors in those regions through bounding box coordinates are defined in x-min, y-min, x-max, and y-max format. In addition, the images that fall under the normal class have simply defined these four coordinates as zeros to indicate the images do not have any objects in particular. Initially, the tumor coordinates are defined in all benign and malignant class images to maintain their original size. Once all images are defined with their tumor localisation bounding box along with the class label, the original images are augmented with images using the imgaug and resized to 1024 x 1024. All the images are converted to RGB to Grayscale mode, converting the class labels to one-hot encode method, and all bounding boxes are rescaled from 0 to 1, which is more flexible to handle the data for regression problems. Finally, the datasets are going to be split in the train-test-validation form and maintain the ratio of 70% for training, 20% for validation, and 10% for testing. INBreast may separately maintain for testing, and CDD-CESM makes train-validation samples using 5-fold cross-validation. Figure 3a to 3g shows the results of the images with all the pre-processing techniques mentioned above.



(a) **(b)** **(c)** **(d)** **(e)** **(f)** **(g)**

Fig. 3. Implemented pre-processing techniques on Image samples

Figure 3a in the above-mentioned image sample displays an actual MLO image from the CEE-CESM dataset. The pectoral region removal is shown in Figure 3b, noise eliminated image is shown in Figure 3c, flipping and contrast enhancements have been applied images are shown in Figures 3d and 3e, and bounding boxes have been fixed for the tumor region images are shown in Figures 3f and 3g.

C. *Multi Output-based DeepCNN Model*

The proposed architecture of the Multi-Output DeepCNN model (MODCNN) builds to perform the dual task for breast cancer classification based on the image label and tumor localisation based on regression bounding box prediction using Keras Application Programming Interface (API) shown in Figure 4. Keras is a top-level API on TensorFlow that is used to build various types of Deep Learning Neural Network (DNN) models, allowing the model's developer to do so quickly and with greater flexibility [32]. Among various available API types in Keras, the Sequential API is implemented in a wide range of domain problems, and it consists of layers that have been arranged in a stack-like structure; the inputs and outputs are passed between the layers in sequential order [33]. The sequential model is so easy to build, but the model provides less flexibility to the developer in any structural adjustments [34]. The Functional API model, in comparison, offers greater flexibility in creating and deploying more reliable models. The proposed Multi-Output DeepCNN model has been implemented as a functional model. The model starts the initial sequential type of layers for feature extractions and maximisation and two parallel sets of layer branches from the dense layers level, one for classification and another for localisation. Figure 4 shows the architectural design of the proposed model that implements using the Bayesian Search [35] hyper-parameter optimisation method, which outperforms the other implemented hyper-parameter tuning method called Random Search optimisation [36]. The proposed model receives its inputs in 1024 × 1024 × 1 dimensions, followed by seven consecutive convolutional 2D layers and Maxpooling 2D layers for feature extraction and feature maximisation. Each 2D convolutional layer is fixed by its hyper-parameters separately with a selection of the different number of filters that varies from 8 to 512 and with fixed kernel size, which is 3 × 3. Every convolutional layer in the model has Rectified Linear Unit (ReLU) familybased activation functions that are used for hidden layer activations which are chosen by the hyper-parameter tuning algorithms and pick the best one on the tuning process. Apart from the feature extraction part, the model is designed with two parallel dense layer branches, one for Classification and another one for object localisation at the bottom of the model. Both dense layers have different levels of layer depth. For dense classification, the branch consists of two fully connected layers with 128 and 64 calculation units and has one dropout unit followed by another fully connected layer. As in the dense localisation branch, it adds one more fully connected layer in the processing of the prediction of bounding box coordinates in mammogram images.

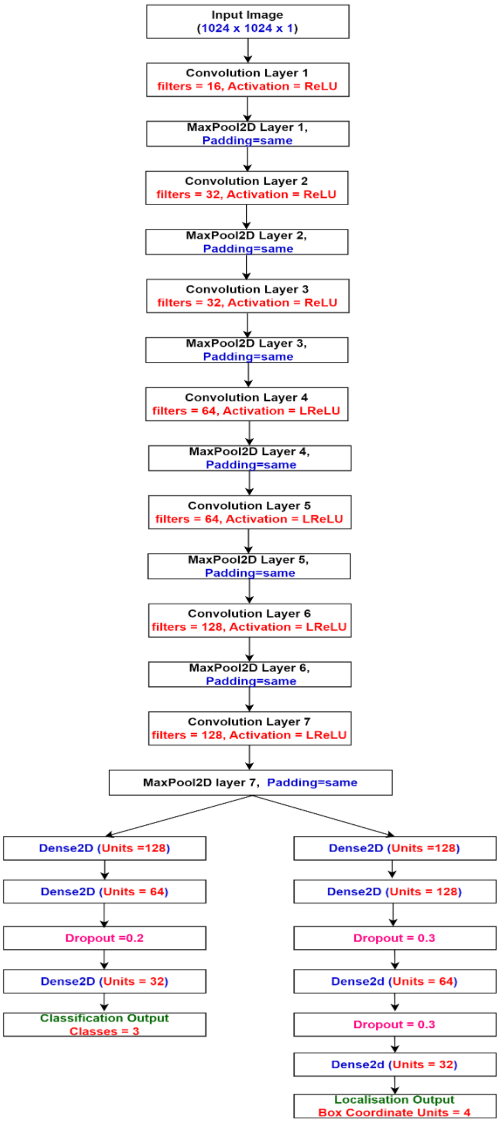


Fig. 4. The Proposed Multi-Output Deep CNN Model

D. *Hyper-parameter optimisation for Multi-Output DeepCNN Model*

The Hyper-parameter is a set of parameters or attributes that should involve in every deep learning model that need to be manually defined and cannot be controlled by the training data itself [37]. Hyper-parameter plays a vital role in the success of the deep learning model for providing optimum results most of the time. Finding an ideal value for the hyper-parameters improves the efficiency and yields a better result for the model because they dictate the model complexity, learnability, and rate of convergence for model parameters. In the proposed Multi-Output DeepCNN model, the tuning of the hyper-parameter is performed in two stages feature extraction level and classification and prediction level in dense layers. Table 2 shows the configuration of hyper-parameters that are involved in the proposed model, and as mentioned in the previous section, the entire model uses Random search and Bayesian search methods for hyper-parameter optimisation. The total number of trials performed to find the best hyper-parameter is 5, each trial was executed and found its best hyper-parameters separately, and the best trial is considered as the proposed one. The number of filters in the proposed model, in each convolution layer, is optimised and fixes the number of filters that varies from 16 to 128. Likewise, with the availability of activation functions in hyper-parameter space, each convolutional layer founds its best-suited activation function also. Most of the hidden layers are settled with ReLU as a hidden activation function, except the third and fifth convolutional layers picked differently. The Leaky ReLU is chosen as the activation function for those two hidden layers. Batch size is another important hyper-parameter that defines the number of samples involved in the training process at the moment of the models' training algorithm. Based on the availability of the image samples and the learning rate selected by the search algorithm, the batch size is fixed. 60 is the selected batch size of the proposed model that randomly selects the samples from the training pool. Apart from the feature extraction part, the model is designed with two parallel arrangments of dense layers, one for classification and another one for object localisation. Both dense layers are executed and find their hyper-parameters with the implementation of two different functions for each and end with classifying the input image based on its class label in the classification layer and fixing the tumor location based on the prediction of bounding box coordinate points in the regression layer. The three hyper-parameters, namely the dropout rate, model optimiser, and output activations, are going to be selected at this level. The range of dropout rates for both classification and localisation is universally set from 0 to 0.5 of the total nodes available at the current dense layer. In the classification branch, there is only one dropout is present, and it fixes at the rate of 0.2, whereas in the localisation part, two dropouts are present and fix both at 0.3 rates. For model optimisation, the Adam optimiser is selected for both classification and localisation. Finally, due to the limited availability of the multi-class output activations, SoftMax was selected as a classifier for breast cancer classification based on the image label, and Sigmoid activation was selected for the prediction of Tumor Bounding Box coordinates. The overall architecture of the proposed Multi-Output DeepCNN is shown in Figure 5.

TABLE II
HYPER-PARAMETERS CONFIGURATION FOR RANDOM SEARCH AND BAYESIAN SEARCH

Level of Optimisation	**Name of the Hyper-parameter**	**Range of the Hyperparameter in Random Search**	**Range of the Hyperparameter in Bayesian Search**
Hidden Layers	Number of Filters	8 to 512	8 to 512
	Learning Rate	0.001, to 0.0000001	0.001, to 0.0000001
	Number of Epochs	10 to 500	10 to 500
	Batch Size	8 to 128	8 to 128
	Hidden Activation	ReLU, LReLU, GReLU, PReLU	ReLU, LReLU, GReLU, PReLU
Dense and Classification layers	Dropout Rate	0 to 0.5	0 to 0.5
	Model Optimiser	SGD, Adam, AdaDelta, AdaGard, RmsProp	SGD, Adam, AdaDelta, AdaGard, RmsProp
	Output Activation	SoftMax, Sigmoid	Linear, Sigmoid

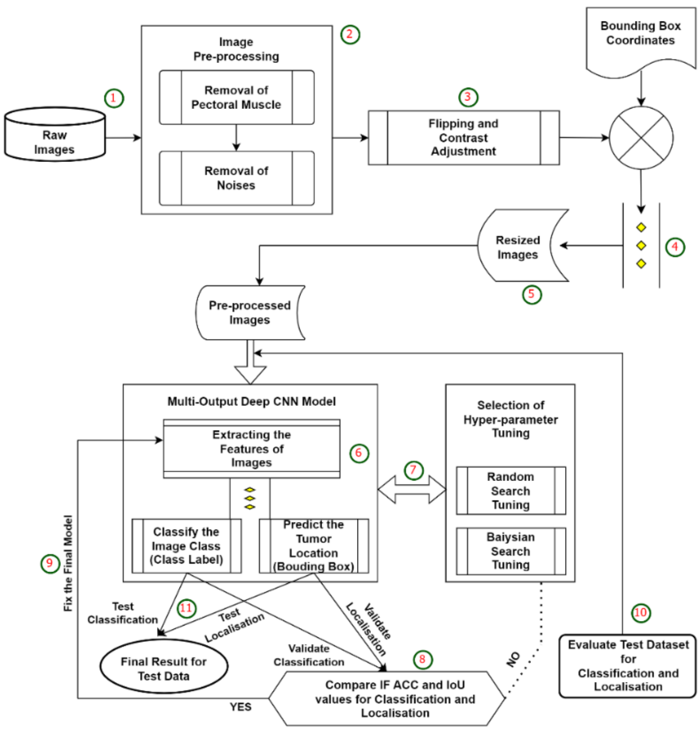


Fig. 5. Architecture of Breast Cancer Classification and Tumor Localisation Model

Both the configuration of Random search and Bayesian search hyper-parameter space mentioned in Table 2 is set equally due to an unbiased chance to find the best hyper-parameters of the proposed model. Additionally, the Early Stopping training monitor is set to only the classification section with three no-improvement iterations. Once the model has been fixed with its best hyper-parameters both in terms of classification and tumor localization, the final Multi-Output DeepCNN model is saved and ready to test the new unseen INBreast dataset samples. Consistently, the pre-trained models, DenseNet169, EfficientNetB4, InceptionV3, ResNet101V2, VGG19, and Xception, are configured with the same hyper-parameter range and methods to compare the performance of the proposed model to find the overall best model among the all.

III. EXPERIMENTAL RESULTS

The proposed Multi-Output DeepCNN model explained in the previous section was implemented in Google Colab with Free GPU selection. All Pre-trained models such as DenseNet169, EfficientNetB4, InceptionV3, ResNet101V2, VGG19, and Xception, with pre-processing techniques, including the removal of the Pectoral breast muscle, noise removal process, and augmentation, were implemented. The Evaluation metrics are methods that are applied to validate or qualify the performance of the DL models [38]. In the evaluation of the proposed model along with pre-trained models, we employed various types of evaluation metrics, mainly for the classification process. The metrics such as accuracy, precision, recall, and F-Score are calculated using the following formulas:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+F} \times 100 \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$F1\ Score = \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

Moreover, the Confusion Matrix has been constructed to validate the model performance on INBreast test data. Further, the evaluation process of the object localisation, Intersection over Union (IoU), and IoU Loss metrics have implemented the prediction of the Bounding box of the tumor portion using Equations 5 and 6 respectively.

$$IoU = \frac{Area\ of\ Overlap}{Area\ of\ Union} \quad (5)$$

$$IoU\ Loss = 1 - IoU \quad (6)$$

IV. Results of the Multi-Output DeepCNN Model

The obtained results from the proposed Multi-Output DeepCNN using Random search and Bayesian search methods are shown in Table 3, and the best hyper-parameters selected using these techniques are presented in Table 4. The results of accuracy, loss for classification, and IoU, IoU loss of localisation are shown in Figure 6 to Figure 9.

TABLE III
PERFORMANCE EVALUATION OF PROPOSED MODEL USING RANDOM SEARCH AND BAYESIAN SEARCH

Performance Measures	MODCNN Random Search	MODCNN Bayesian Search
Validation Accuracy	93.88 %	95.42 %
Test Accuracy	94.69 %	**96.39 %**
Validation Loss	0.39	0.29
Test Loss	0.31	**0.22**
Precision	93.90 %	96.12 %
Recall	93. 85 %	96.25 %
F-Score	93.90 %	96.15 %
Validation IoU	78.66 %	81.90 %
Test IoU	80.32 %	**84.41 %**
Validation IoU Loss	0.39	0.25
Test IoU Loss	0.34	**0.19**

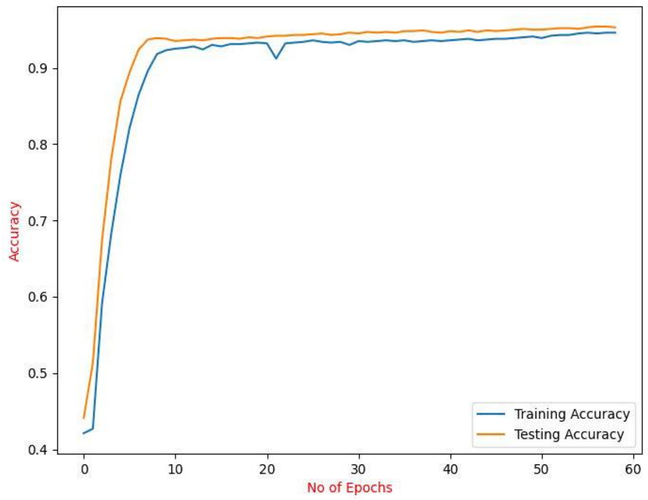


Fig. 6. Accuracy rate of MODCNN using Bayesiyan Method

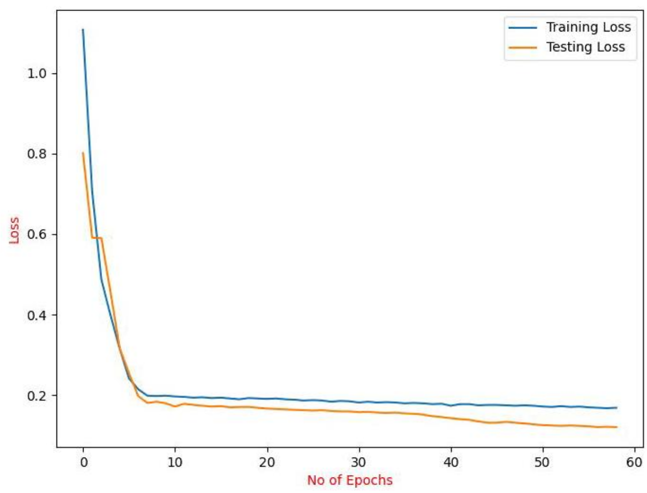


Fig. 7. Loss rate of MODCNN using Bayesiyan Method

TABLE IV
SELECTION OF BEST HYPER-PARAMETERS USING RANDOM SEARCH AND BAYESIAN SEARCH

Hyper-parameters	**Best HPO value using MODCNN Random Search**	**Best HPO value using MODCNN Bayesian Search**
No. of Filters	16, 32, 64 and 128	32, 64, 128
Learning Rate	0.00017	0.00012
No. of Epochs	89	59
Batch Size	38	60
Hidden Activation	GReLU	ReLU
Dropout Rate	0.2 (classification), 0.2, 0.3 (localisation)	0.2 (classification), 0.3, 0.3 (localisation)
Model Optimiser	Adam (classification), AdaDelta (localisation)	Adam (classification), Adam (localisation)
Output Activation	SoftMax(classification), Sigmoid(localisation)	SoftMax(classification), Sigmoid(localisation)

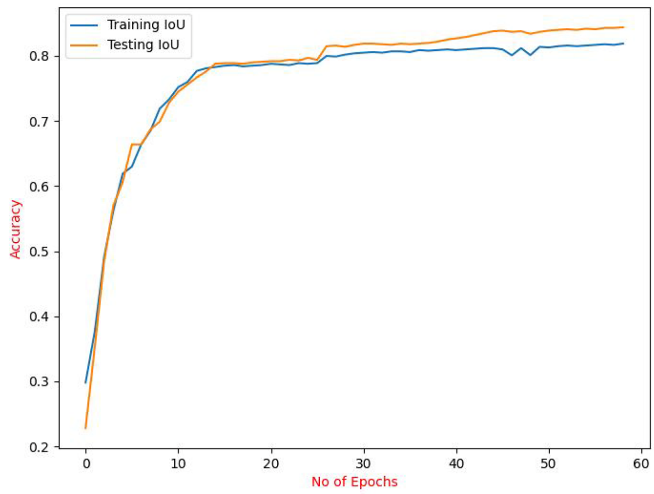


Fig. 8. IoU rate of MODCNN using Bayesiyan Method

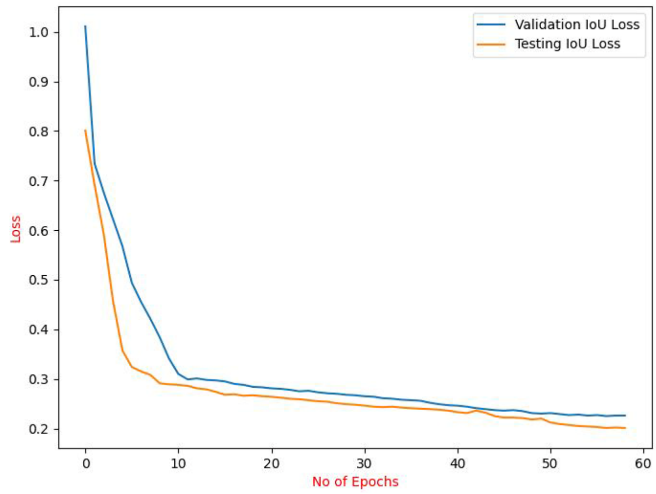


Fig. 9. IoU Loss rate of MODCNN using Bayesiyan Method

From the above Tables 3 and 4, the performance of classification, localisation, and selection of hyper-parameters of the proposed model is displayed. The implementation of both the Random and Bayesian search methods gives the best-fine-tuned hyper-parameters to the proposed model of their own. But, compared to Random search, Bayesian search gives better results in terms of validation accuracy and test accuracy. The model produces 93.88% validation accuracy using the random search method, but this accuracy is significantly increased up to 95.42% when applying the bayesiyan method. Also, the performance of test data shows the same result pattern of 94.69% when applied a random search method and 96.39 % for the bayesian method produces superior performance for classification, as shown in Figure 10, with the Confusion Matrix format. Subsequently, the model utilises the Bayesian hyper-parameter method, producing a better IoU and IoU loss of 84.41% and 0.20 for the tumor localisation task as compared to random search produces only 80.30% and 0.34 IoU and IoU loss shown in Figure 11 to 13. But overall, both techniques perform well in the classification task as compared to the localisation task due to the nature of the distribution features of mammogram images.

A. Results of Pre-trained Models

Apart from the proposed Multi-Output DeepCNN model being built for breast cancer classification and tumor localisation, six state-of-the-art Pre-trained models are also implemented in the same datasets with the same hyper-parameters configuration. All six pre-trained models have utilised the original architectures of their own and frozen their layers in the lower level up to the final pooling layers and added the customised parallel layers, one for breast cancer classification and another for tumor localisation in the model's dense and classification layers level. In the second stage of the hyper-parameter tuning process, all six pre-trained models are tuned only to the Dropout level, Model Optimiser, and Output Activations due to the frozen bottom-level feature extraction layers. Tables 5 to 8 show the implementation results of the CDD-CESM and INBreast data on all six pre-trained models for the classification and localisation process using both Random and Bayesiyan search methods, and Figures 14 to 17 show the overall comparison of the performance of all the models and methods we implemented.

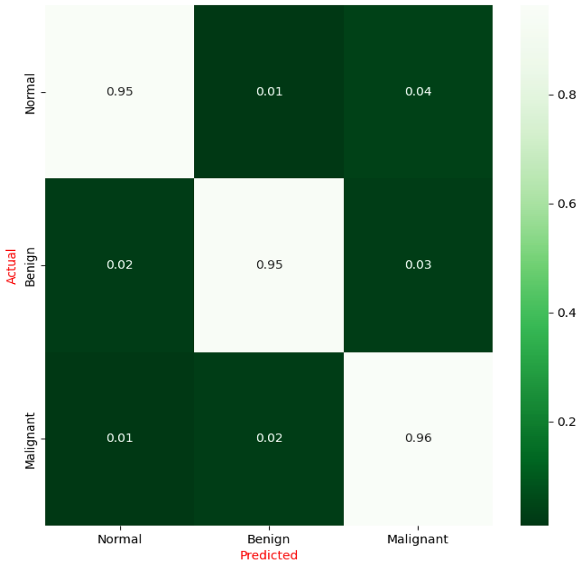


Fig. 10. The Proposed Multi-Output DeepCNN Model

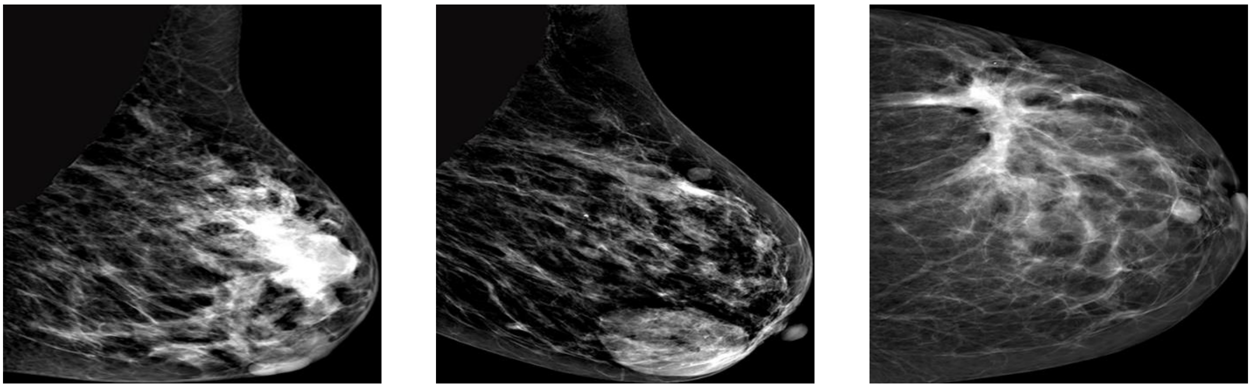


Fig. 11. Original Pre-processed Mammogram Images

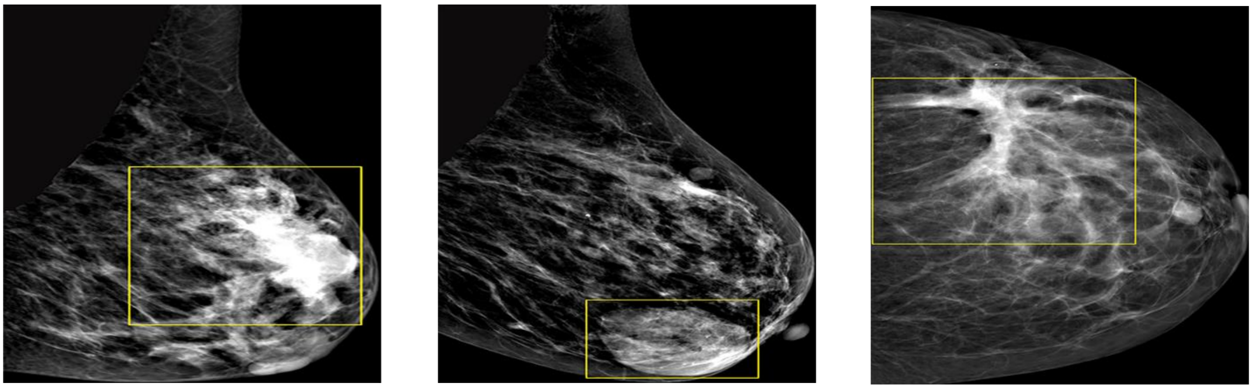


Fig. 12. Original Pre-processed Mammogram Images with Ground truth Bounding Box

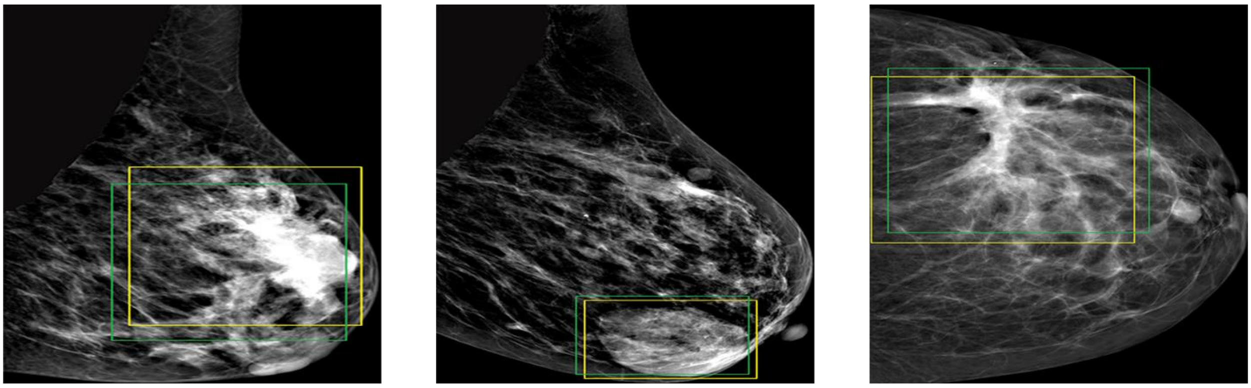


Fig. 13. Bounding Box Precited Images using Multi-Output DeepCNN Model Tumor Classification with Localisation

TABLE V
RESULTS OF PERFORMANCE MEASURES ON PRE-TRAINED MODELS (RANDOM SEARCH METHOD)

Performance Measures	**DenseNet169**	**EfficientNetB4**	**InceptionV3**	**ResNet101V2**	**VGG19**	**Xception**
Validation Accuracy	83.11 %	90.79 %	89.68 %	88.76 %	80.27 %	90.89 %
Test Accuracy	84.91 %	92.38 %	90.65 %	89.87 %	81.89 %	**92.47 %**
Validation Loss	0.46	0.40	0.41	0.45	0.52	0.39
Test Loss	0.40	0.31	0.34	0.36	0.45	**0.32**
Precision	82.80 %	90.71 %	90.00 %	89.10 %	81.60 %	91.97 %
Recall	82.90 %	91.05 %	90.03 %	89.03 %	81.34 %	91.05 %
F-Score	82.83 %	90.90 %	89.98 %	89.91 %	81.54 %	91.78 %
Validation IoU	72.88 %	**79.37 %**	77.67 %	74.39 %	63.80%	79.10 %
Test IoU	73.57 %	80.21 %	78.59 %	75. 11%	65.01 %	79.89 %
Validation IoU Loss	0.57	**0.38**	0.46	0.50	0.66	0.41
Test IoU Loss	0.53	0.32	0.40	0.44	0.61	0.35

TABLE VI
RESULTS OF PERFORMANCE MEASURES ON PRE-TRAINED MODELS (BAYESIYAN SEARCH METHOD)

Performance Measures	**DenseNet169**	**EfficientNetB4**	**InceptionV3**	**ResNet101V2**	**VGG19**	**Xception**
Validation Accuracy	87.68 %	94.21 %	92.83 %	90.36 %	84.89 %	94.56 %
Test Accuracy	89.57 %	94.94 %	93.21 %	91.13 %	86.28 %	**95.97 %**
Validation Loss	0.38	0.27	0.30	0.34	0.43	0.29
Test Loss	0.32	0.23	0.25	0.29	0.37	**0.21**
Precision	87.30 %	94.11 %	92.60 %	90.10 %	83.90 %	94.35 %
Recall	89.30 %	94.76 %	93.05 %	91.03 %	86.14 %	95.67 %
F-Score	89.23 %	94.60 %	93.05 %	90.98 %	86.14 %	95.25 %
Validation IoU	80.67 %	81.80 %	78.69 %	77.30 %	71.96 %	82.50 %
Test IoU	81.54 %	82.62 %	79.88 %	78.38 %	72.75 %	**83.76 %**
Validation IoU Loss	0.48	0.34	0.35	0.37	0.57	0.32
Test IoU Loss	0.41	0.27	0.27	0.30	0.49	**0.26**

TABLE VII

RESULTS OF HYPER-PARAMETERS TUNING ON PRE-TRAINED MODELS (RANDOM SEARCH METHOD)

Hyper-parameters	**DenseNet169**	**EfficientNetB4**	**InceptionV3**	**ResNet101V2**	**VGG19**	**Xception**
Dropout Rate	0.1, 0.2	0.2, 0.2	0.1, 0.3	0.1, 0.1	0.2, 0.2	0.1, 0.2
Model Optimiser	Adam	Adam	AdaGard	Adam	Adam	AdaDelta
Output Activation	SoftMax (C) Sigmoid (L)	SoftMax (C) Sigmoid (L)	SoftMax (C) Sigmoid (L)	SoftMax (C) Sigmoid (L)	SoftMax (C) Sigmoid (L)	SoftMax (C) Sigmoid (L)

TABLE VIII

RESULTS OF HYPER-PARAMETERS TUNING ON PRE-TRAINED MODELS (BAYESIYAN SEARCH METHOD)

Hyper-parameters	**DenseNet169**	**EfficientNetB4**	**InceptionV3**	**ResNet101V2**	**VGG19**	**Xception**
Dropout Rate	0.1, 0.1	0.1, 0.2	0.1, 0.2	0.2, 0.1	0.2, 0.2	0.2, 0.2
Model Optimiser	Adam	AdaDelta	Adam	Adam	AdaDelta	Adam
Output Activation	SoftMax (C) Sigmoid (L)	SoftMax (C) Sigmoid (L)	SoftMax (C) Sigmoid (L)	SoftMax (C) Sigmoid (L)	SoftMax (C) Sigmoid (L)	SoftMax (C) Sigmoid (L)

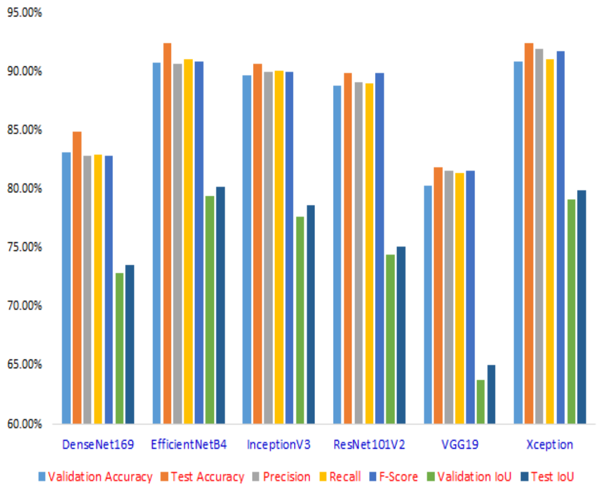


Fig. 14. Accuracy, IoU of Pre-trained Models using Random search optimisation

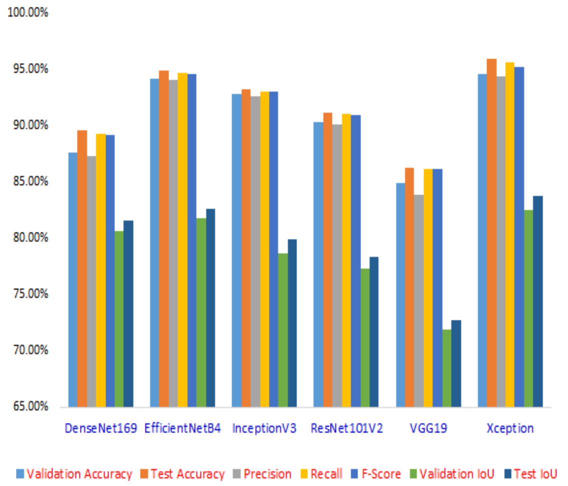


Fig. 15. Accuracy, IoU of Pre-trained Models using Bayesiyan search optimisation

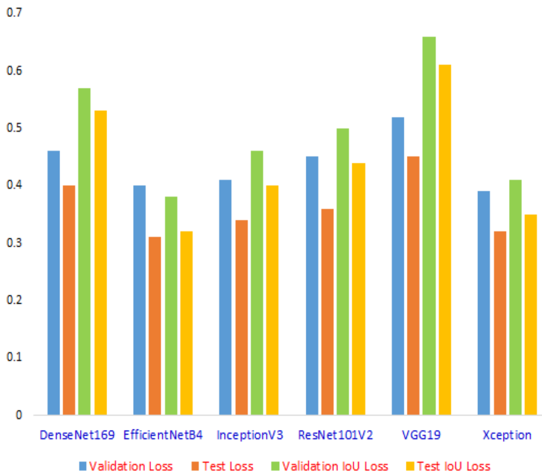


Fig. 16. Acc.Loss, IoU Loss of Pre-trained Models using Random search optimisation

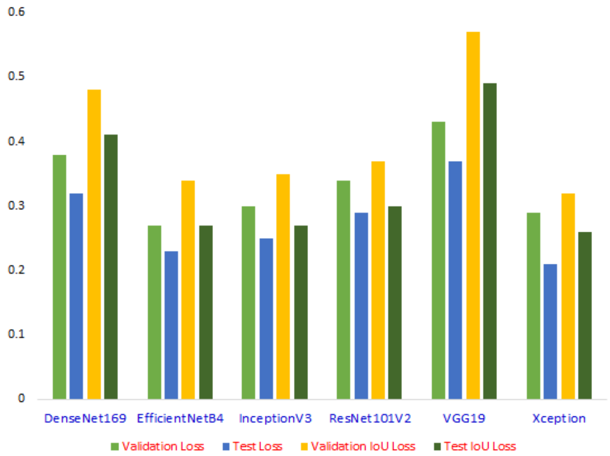


Fig. 17. Acc.Loss, IoU Loss of Pre-trained Models using Bayesiyan search optimisation

The above figures show the comparison of all six pre-trained models' performance against each other's test accuracy, IoU Accuracy, test losses, and IoU losses. From these results, the Xception pre-trained model outperforms all other models with 95.97% of accuracy and 0.21 loss. While EfficientNetB4 also performs well and is very close to the accuracy and loss values of Xception model results with 94.94% and 0.23 loss. Also, in the tumor localisation process, Xception pre-train model shows better IoU accuracy and IoU loss values with

83.76 % and 0.26, followed by EfficientNetB4, which produces 82.62% and 0.27 loss. Notably, in the tumor localisation process, the Random search implemented EfficientNetB4 pre-trained model outperformed compared to Xception model with a very small difference in values, 79.37% against 79.10% and 0.38 against 0.41 in terms of loss. Furthermore, the proposed Multi-Output DeepCNN model with Bayesiyan hyper-parameter optimisation outperforms both the random search implemented proposed model and all six pre-trained models that classify breast cancer that with tumor localisation with a high classification accuracy rate and IoU accuracy rate of 96.39% and 84.41% with a minimal amount of losses 0.22 and 0.19 for model loss and IoU loss respectively.

V. CONCLUSION

The Exponential growth of Computer Vision is an essential need for solving various real-time problems the human community is facing in the current era. Medical domain, particularly in radiology, deep learning-based CNN models are leading to providing reliable solutions for finding various types of cancers, including breast cancer. In this study, a Multi-Output based Deep Convolutional Neural Network Model was presented that not only classifies the breast cancer from the mammogram images but also localises the tumor portion in the mammogram images too. The problem of solving classification with localisation using the CNN model is always complex to achieve high performance, especially using medical images. However, the proposed model with the bayesiyan hyper-parameter optimisation method exhibits better performance as compared to random search optimisation and implementation of various pre-trained models such as DenseNet169, EfficientNetB4, InceptionV3, ResNet101V2, VGG19, and Xception with 96. 39% of classification test accuracy and 84.41 % localisation IoU accuracy.

ACKNOWLEDGEMENT

The First author acknowledges Periyar University for partial financial support to this research work under the University Research Fellowship (URF) Scheme. *Ref.No.* PU/AD-3/URF Order/10191/23F77954/2023, DATED: 08/06/2023.

The Second and Third authors acknowledge the UGC Special Assistant Program (SAP) New Delhi, for partial financial support under the Scheme of UGC SAP PHASE-II, *Project Number:* F. 5-6/2018/DRS-II (SAP-II).

REFERENCES

[1]. https://www.who.int/news-room/factsheets/detail/cancer#:~:text=The%20problem,deaths%20in%2020 20%20(1).

[2]. Nusrat Mohi ud din, Rayees Ahmad Dar, Muzafar Rasool, Assif Assad, Breast cancer detection using deep learning: Datasets, methods, and challenges ahead, Computers in Biology and Medicine, Volume 149, 2022, ISSN 0010-4825, https://doi.org/10.1016/j.compbiomed.2022.106073.

[3]. Petrucelli N, Daly MB, Pal T. BRCA1- and BRCA2-Associated Hereditary Breast and Ovarian Cancer. 1998 Sep 4 [Updated 2022 May 26]. In: Adam MP, Mirzaa GM, Pagon RA, et al., editors. GeneReviews® [Internet]. Seattle (WA): University of Washington, Seattle; 1993-2023. Available from: https://www.ncbi.nlm.nih.gov/books/NBK1247/

[4]. Ataollahi MR, Sharifi J, Paknahad MR, Paknahad A. Breast cancer and associated factors: a review. J Med Life. 2015;8(Spec Iss 4):6-11. PMID: 28316699; PMCID: PMC5319297.

[5]. R. Sathesh Raaj, Breast cancer detection and diagnosis using hybrid deep learning architecture, Biomedical Signal Processing and Control, Volume 82, 2023, ISSN 1746-8094, https://doi.org/10.1016/j.bspc.2022.104558.

[6]. Kim, E.-K. et al. Applying Data-driven Imaging Biomarker in Mammography for Breast Cancer Screening: Preliminary Study. Sci. Reports 8, 2762 (2018).

[7]. Agarwal, R., Diaz, O., Lladó, X., Yap, M. H. & Martí, R. Automatic mass detection in mammograms using deep convolutional neural networks. J. Med. Imaging 6, 031409 (2019)

[8]. Aa Dhungel, N., Carneiro, G. & Bradley, A. P. The Automated Learning of Deep Features for Breast Mass Classification from Mammograms. In Ourselin, S., Joskowicz, L., Sabuncu, M. R., Unal, G. & Wells, W. (eds) Medical Image Computing and Computer- Assisted Intervention – MICCAI 2016: 19th International Conference, Athens, Greece, October 17–21, 2016, Proceedings, Part II, 106–114 (Springer International Publishing, Cham, 2016).

[9]. Dipali Sinha & K. Thangavel (2022) Automatic epileptic signal classification using deep convolutional neural network, Journal of Discrete Mathematical Sciences and Cryptography, 25:4, 963-973, DOI: 10.1080/09720529.2022.2072419.

[10]. Saba Dadsetan, Dooman Arefan, Wendie A. Berg, Margarita L. Zuley, Jules H. Sumkin, Shandong Wu, Deep learning of longitudinal mammogram examinations for breast cancer risk prediction, Pattern Recognition, Volume 132, 2022, ISSN 0031-3203, https://doi.org/10.1016/j.patcog.2022.108919.

[11]. S. Sharma, R. Mehra, Conventional machine learning and deep learn- ing approach for multi-classification of breast cancer histopathology imagesa comparative insight, J. Digit. Imaging 33 (3) (2020) 632–654, doi: 10.1007/S10278-019- 00307- Y/TABLES/10 . https://link.springer.com/article/10.1007/s10278-019-00307-y.

[12]. S.U. Khan, N. Islam, Z. Jan, I. Ud Din, J.J. Rodrigues, A novel deep learning based framework for the detection and classification of breast cancer using transfer learning, Pattern Recognit. Lett. 125 (2019) 1–6, doi: 10.1016/J.PATREC.2019.03.022

[13]. Wang, H., Singhal, A. & Liu, P. Tackling imbalanced data in cybersecurity with transfer learning: a case with ROP payload detection. Cybersecurity 6, 2 (2023). https://doi.org/10.1186/s42400-022-00135-8.

[14]. Z. Hameed, S. Zahia, B. Garcia-Zapirain, J.J. Aguirre, A.M. Vanegas, Breast cancer histopathology image classification using an ensemble of deep learning models, Sensors 20 (16) (2020) 4373, doi: 10.3390/S20164373 . https://www.mdpi.com/1424-8220/20/16/4373/

[15]. Mehdi Hassan1, Safdar Ali, Muhammad Sanaullah, et.al., Drug Response Prediction of Liver Cancer Cell Line Using Deep Learning, Computers, Materials & Continua, 2022, Volume 70, 2, 1546-2226, DOI - 10.32604/cmc.2022.020055.

[16]. Faraz Lotfi, Farnoosh Faraji, Hamid D. Taghirad,Object localization through a single multiple-model switching CNN and a superpixel training approach, Applied Soft Computing, Volume 115, 2022, 108166, ISSN 1568-4946, https://doi.org/10.1016/j.asoc.2021.108166.

[17]. Inês C. Moreira, Igor Amaral, Inês Domingues, António Cardoso, Maria João Cardoso, Jaime S. Cardoso, INbreast: Toward a Full-field Digital Mammographic Database, Academic Radiology, Volume 19, Issue 2, 2012, Pages 236-248, ISSN 1076-6332, https://doi.org/10.1016/j.acra.2011.09.014.

[18]. Khaled, R., Helal, M., Alfarghaly, O. et al. Categorized contrast enhanced mammography dataset for diagnostic and artificial intelligence research. Sci Data 9, 122 (2022). https://doi.org/10.1038/s41597-022-01238-0.

[19]. https://radiopaedia.org/articles/breast-imaging-reporting-and-data-system-bi-rads.

[20]. D.A. Spak, J.S. Plaxco, L. Santiago, M.J. Dryden, B.E. Dogan, BI-RADS® fifth edition: A summary of changes, Diagnostic and Interventional Imaging, Volume 98, Issue 3, 2017, Pages 179-190, ISSN 2211 5684, https://doi.org/10.1016/j.diii.2017.01.001.

[21]. Maitra IK, Nag S, Bandyopadhyay SK. Technique for preprocessing of digital mammogram. Comput Methods Programs Biomed. 2012 Aug;107(2):175-88. doi: 10.1016/j.cmpb.2011.05.007. Epub 2011 Jun 12. PMID: 21669471.

[22]. Zhou K, Li W, Zhao D. Deep learning-based breast region extraction of mammographic images combining pre-processing methods and semantic segmentation supported by Deeplab v3. Technol Health Care. 2022;30(S1):173-190. doi: 10.3233/THC-228017. PMID: 35124595; PMCID: PMC9028646.

[23]. Al-Antari MA, Al-Masni MA, Kim TS. Deep Learning Computer-Aided Diagnosis for Breast Lesion in Digital Mammogram. Adv Exp Med Biol. 2020;1213:59-72. doi: 10.1007/978-3-030-33128-3_4. PMID: 32030663.

[24]. Sushreeta Tripathy, Tripti Swarnkar, Unified Preprocessing and Enhancement Technique for Mammogram Images, Procedia Computer Science, Volume 167, 2020, Pages 285-292, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2020.03.223.

[25]. Abdullah-Al Nahid, Mohamad Ali Mehrabi, and Yinan Kong, Histopathological Breast Cancer Image Classification by Deep Neural Network Techniques Guided by Local Clustering, Hindawi, BioMed Research International, Volume 2018, Article ID 2362108, 20 pages, https://doi.org/10.1155/2018/2362108.

[26]. Hazarika M, Mahanta LB. A New Breast Border Extraction and Contrast Enhancement Technique with Digital Mammogram Images for Improved Detection of Breast Cancer. Asian Pac J Cancer Prev. 2018 Aug 24;19(8):2141-2148. doi: 10.22034/APJCP.2018.19.8.2141. PMID: 30139217; PMCID: PMC6171404.

[27]. Razali NF, Isa IS, Sulaiman SN, Abdul Karim NK, Osman MK, Che Soh ZH. Enhancement Technique Based on the Breast Density Level for Mammogram for Computer-Aided Diagnosis. Bioengineering (Basel). 2023 Jan 23;10(2):153. doi: 10.3390/bioengineering10020153. PMID: 36829647; PMCID: PMC9952042.

[28]. Punitha S., Fadi Al-Turjman, Thompson Stephan, An automated breast cancer diagnosis using feature selection and parameter optimization in ANN, Computers & Electrical Engineering, Volume 90, 2021, 106958, ISSN 0045-7906, https://doi.org/10.1016/j.compeleceng.2020.106958.

[29]. F. Z. Ouadiay, H. Bouftaih, E. H. Bouyakhf and M. M. Himmi, "Simultaneous object detection and localization using convolutional neural networks," 2018 International Conference on Intelligent Systems and Computer Vision (ISCV), Fez, Morocco, 2018, pp. 1-8, doi: 10.1109/ISACV.2018.8354045.

[30]. N. Dhungel, G. Carneiro, A. P. Bradley, Deep learning and structured prediction for the segmentation of mass in mammograms, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 605–612, 2015.

[31]. Elijah Cole, Kimberly Wilber, Grant Van Horn, Xuan Yang, Marco Fornoni, Pietro Perona, Serge Belongie, Andrew Howard, Oisin Mac Aodha, On Label Granularity and Object Localization, Computer Vision and Pattern Recognition, https://doi.org/10.48550/arXiv.2207.10225.

[32]. Jeff Heaton, Applications of Deep Neural Networks with Keras, Heaton Research, Inc., arXiv:2009.05673, ISBN: 9798416344269.

[33]. Yanzhao Wu, Ling Liu, Fellow, IEEE, Calton Pu, Fellow, IEEE, Wenqi Cao, Semih Sahin, Wenqi Wei, Qi Zhang, A Comparative Measurement Study of Deep Learning as a Service Framework, arXiv:1810.12210v2 [cs.PF] 18 Aug 2019.

[34]. Samaya Madhavan, Sidra Ahmed, Vinay Rao, Anto John Developer.IBM, Published March 11, 2021, https://developer.ibm.com/articles/compare-deep-learning-frameworks/

[35]. Tong Yu, Hong Zhu, Hyper-Parameter Optimization: A Review of Algorithms and Applications, arXiv:2003.05689v1 [cs.LG] 12 Mar 2020.

[36]. Lisha Li, Kevin Jamieson, Giulia DeSalvo, Afshin Rostamizadeh, and Ameet Talwalkar. Hyperband: A novel bandit-based approach to hyperparameter optimization. The Journal of Machine Learning Research, 18(1):6765–6816, 2017.

[37]. Pablo Ribalta Lorenzo, Jakub Nalepa, Michal Kawulok, Luciano Sanchez Ramos, and Jos´e Ranilla Pastor. Particle swarm optimization for hyper-parameter selection in deep neural networks. In Proceedings of the Genetic and evolutionary computation conference, pages 481–488, 2017.

[38]. Carvalho, D.V.; Pereira, E.M.; Cardoso, J.S. Machine Learning Interpretability: A Survey on Methods and Metrics. Electronics 2019, 8, 832. https://doi.org/10.3390/electronics8080832

