Deep-Ensemble-Learning based Breast-Cancer Classification

By

19BCE229 – Rushir Bhavsar
19BCE260 – Smit Shah

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
Ahmedabad 382481

Deep Ensemble Learning-based Breast Cancer Classification

Report

Submitted in partial fulfillment of the requirements

For the degree of

Bachelor of Technology in Computer Science & Engineering

By

19BCE229 – Rushir Bhavsar
19BCE260 – Smit Shah

Guided By

Prof. Parita Oza
[DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING]

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
Ahmedabad 382481

CERTIFICATE

This is to certify that the Minor Project entitled **Deep Ensemble Learning based Breast Cancer Classification** submitted by Rushir Bhavsar (19BCE229) and Smit Shah (19BCE260), towards the partial fulfillment of the requirements for the degree of Bachelor of Technology in Computer Science and Engineering of Nirma University is the record of work carried out by him under my supervision and guidance. In my opinion, the submitted work has reached a level required for being accepted for examination.

Prof. Parita Oza
Assistant Professor
Computer Science and Engineering Dept.,
Institute of Technology,
Nirma University,
Ahmedabad

Dr. Madhuri Bhavsar,
Professor and HOD,
Computer Science and Engineering Dept.,
Institute of Technology,
Nirma University,
Ahmedabad

ACKNOWLEDGEMENT

First and foremost, we would like to express our special thanks of gratitude to our Project Guide Mrs. Parita Oza who gave us this opportunity to work on this wonderful project of “Deep Ensemble learning based Mammogram Classification”, while also being a source of inspiration and guiding us on the path of learning new things about the problem domain. We are extremely grateful to Nirma University and Centre of Excellence for providing the needed resources utilized in the process of working on this project. Secondly, we would like to thank our classmates and colleagues for giving their wonderful support in the time of working on this project. Lastly, we would like to thank our family, siblings, and friends for their essential help and undivided support.

ABSTRACT

Any sickness that is detected early enough may usually be treated with minimal human effort. Most people miss their sickness before it develops into a chronic condition. It causes the death rate to rise globally. One of the diseases that can be treated is breast cancer if it is discovered early on and before it has spread to every part of the body. Due to misunderstanding, the medical professional could make a wrong diagnosis. The computer-aided diagnostic (CAD) is an automated tool that provides doctors with accurate results when assessing the seriousness of their patients' diseases. A CAD system for automated breast cancer diagnosis is shown in this chapter. Transfer learning significantly contributes to the accurate detection and classification of cancer for limited datasets of medical pictures. In this project we have implemented CAD method for breast cancer classification using a combination of transfer learning models (DenseNet, MobileNet, XceptionNet, Inception-ResNet-V2) and Deep Ensemble Learning Models on the combination of the following 5 datasets (SureMaPP, DDSM, MIAS, InBreast, and KAU). Our implemented method provides the best average validation accuracy for binary classification of benign or malignant cancer cases of 58.7%, 56.66%, and 48.94% for XceptionNet, InceptionResNetV2, and MobileNet, respectively. Lowest Achieved accuracy was for DenseNet models @42.4%. Post subjecting the models over various combination under deep ensemble learning, a ensemble model accuracy of 80% was achieved for Direchlet Ensembling while average accuracy of only 65% for Basic stack ensembling.

CONTENTS

1. Introduction

Increase in daily cases by over 50% in the last 3 decades, early breast cancer diagnosis and treatment have become a necessity for the good health of women around the globe. To curb the issue of early detection and bring efficient resolution to this problem, Computer Vision, and Deep Learning (DL) are one of the best options for solving the task of efficient detection of breast cancer. Independent DL classification models have been efficient in detecting and classifying breast cancer to some extent; however, the combined performance of such models is far more efficacious in practical deployment scenarios. Deep Ensemble Learning or simply ensemble modeling involves combining various models wherein the combined bias and variance of different models help reduce the drawbacks faced for inlying feature extraction in the successful detection and further assists in expanding the image universe i.e., accepting all variations in the medical images.

1.1 Topic Title

Deep Ensemble Learning-based Breast Cancer Classification

1.2 Objective

- Identify and collect required Datasets
- Pre-process the Data for Ensemble Modelling
- Identify the Transfer Learning Models for Single Model performance evaluation
- Based on Accuracy find the 3 Best Transfer learning Models and create Deep Ensemble Learning Model
- Classify the mammograms as Benign, Malignant, and Normal class evaluate the ensemble modelling performance.

2. Literature Review

In various research, recurrent and convolution-based neural network models were used for training and testing for classification tasks. Further to improve upon baseline DL architecture, various authors used Segmentation masks to add inlay of highlighted features of the Image's ROI. For ensemble modeling, various authors employed simple stacking of various explicitly defined model architectures, whereas various authors simply employed pre-defined ensemble classifiers such as Decision Trees and Random Forests. Highlighting the model architecture details, CNN was majorly used along with an ML classifier, and an advanced Stacked Autoencoder was used. In various research, using ensemble modeling methods, authors achieved more than 95% accuracy across the neural network model variations. Ensemble learning gave better results compared to the independently trained models and had a high change in the precision and recall values. The various Ensemble Models included CNN, SVM, LSTM- CNN combination, and Stacked Auto-encoder, fused with Regression models. The literature surveyed has been tabulated in Table 1 specifying the objective, methodology, and technology used. In the Table 1, various research surveyed employ various advanced deep learning-based vision model which have been able to achieve a superior output via increased computation time for training epochs and use of high-resolution dataset. Hence, due to this we have worked on using basic transfer learning models for training purposes and employed Deep ensemble modelling to achieve better efficiency at mammogram classification.

3. Methodology

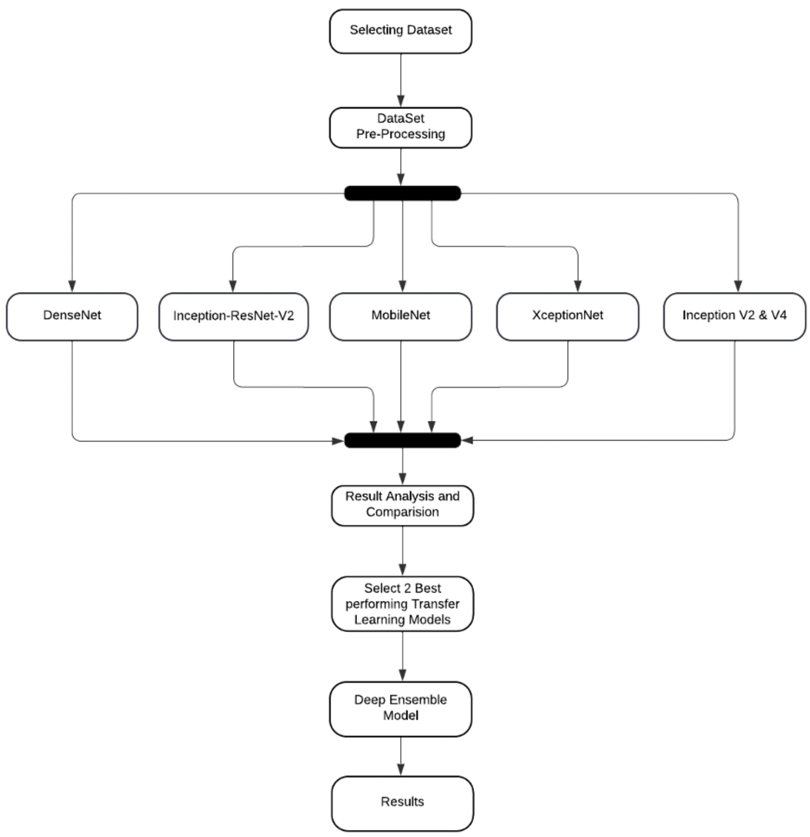


Fig. 1: Overview of the System Model

In this subsection we elaborate the methodology followed in the implementation of deep ensemble modelling for the task of mammogram classification. Firstly, based on the literature survey we identified the need for more mammogram data i.e., increased number of images for better model training. Secondly the collection of images was processed following which the image was subject to individual transfer learning-based model training. Thirdly, after the training – testing of the models, all the models were evaluated for performance and compared based on accuracy metric. After evaluation, best 3 models will be selected and subject to ensemble modelling. Lastly, ensemble modelling results will be evaluated for best classification accuracy.

TABLE I: A comparison of the state-of-the-art methodologies.

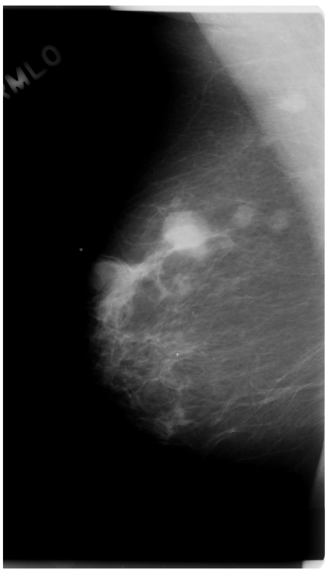
Author	**Year**	**Objective**	**Methodology**	**Technology Used**	**Pros**
sharaf J. Malebary et al.	2021	Create an Automated Breast Mass Classification System with the help of deep ensemble learning algorithm like RNN-LSTM, CNN.	Segmentation using K-mean Clustering. The Output of this segmentation is the clusters. Extract high and low-level features and then perform the comparative analysis with the RNN-LSTM, and CNN techniques.	RNN-LSTM, CNN techniques with the ensemble learning method	RNN-LSTM and CNN methods are proposed to extract the high and low-level features and to find the patches in less-contrast mammography images. ROI is extracted automatically through segmentation to reduce radiologist assistance costs.
Vinod Jagannath Kadam et al.	2019	To create a Breast Cancer Diagnosis with the Help of feature Ensemble Learning	The authors have used two models and the data set used is WDBC dataset. The author has used 3 combinations of these SAE1 and SAE2 models	Feature ensemble based Stacked autoencoder + soft-max regression-based model (FE-SSAE-SM) and Stacked autoencoder and softmax regression-based model (SSAE-SM model)	To find the opti-mal value of parameters , and for both models, we used Grid search. Feature ensemble learning based on stacked sparse autoencoders and softmax regression to classify widely adopted WDBC UCI data sets.
Rong Sun et al.	2021	To Predict the cancer-molecular subtypes using DCE-MRI which is based on CNN combined with ensemble Learning	3 CNN models with the same architecture were consequently trained, followed by preliminary prediction of probabilities from the testing database, Final predictions were made based on ensemble learning which is based on weighted voting	Convolutional Neural Network (CNN), Ensemble Learning	The retrospectively studied preoperative dynamic contrast enhanced-magnetic resonance imaging and molecular information of 266 breast cancer cases with either luminal subtype (luminal A and luminal B) or non-luminal subtype (human epidermal growth factor receptor 2 and triple-negative) were used as a Dataset
Mhd M. Ghiasi et al.	2020	To Classify Breast Cancer with the Decision tree-based ensemble learning	The RF and ET approaches include four main stages: input identification, determination of the optimal number of trees, voting analysis, and final decision.	RF and EF decision-based ensemble learning methodologies	The models implemented in this paper consider important factors such as uniformity of cell size, bland chromatin, mitoses, and clump thickness as the input parameters. Also, the accuracy is high.
Rong Sun et al.	2021	Use of transfer learning technique based on unsupervised learning and ensemble learning for breast cancer molecular subtype prediction	Data Set was prepared then three baseline networks corresponding to the different MRI post-contrast phases were built the pre-trained networks were fine-tuned on the labeled target domain. Finally, prediction results were integrated using weighted voting-based ensemble learning.	Sequence-Based Ensemble Learning, Multiview ensemble learning	3.0T, dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) using T1-weighted high-resolution isotropic volume examination is used.
Moloud Abdar et al.	2019	Use of CWV-BANN-SVM ensemble learning classifier for the diagnosis of breast cancer	The author first tested the SVM algorithm using various values of the C, and parameters. Then CWV-BANN-SVM ensemble learning classifier was used to improve the accuracy	CWV-BANN-SVM which combines boosting ANNs (BANN) and two SVMs	Most important risk factors of breast cancer were extracted using SVM. Benchmark clinical data set is utilized to check the effectiveness of the model.
Woo Kyung Moon	2020	Diagnosis of breast ultrasound images using ensemble learning from CNN.	CNN architecture is used to classify images in CAD Systems. Ensemble Learning is also used to improve accuracy.	Ensemble Learning from Convolutional Neural Networks	Many CNN models on different data sets, including original tumor image, segmented tumor images, tumor mask, and fused image, to compare the diagnostic results.

3.1 Dataset Details

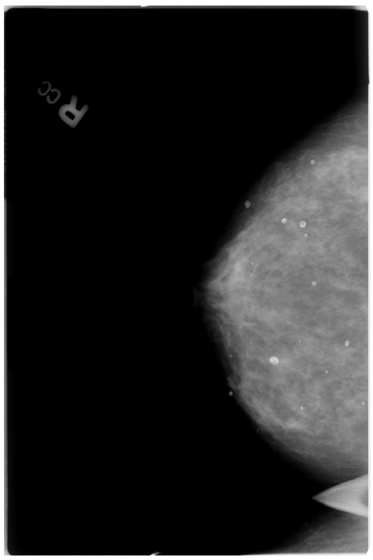
For this project we have collected and employed 5 datasets open to public use. The 5 datasets utilized are listed as follows,

1. DDSM
2. SureMaPP
3. MIAS
4. InBreast
5. KMD

DDSM dataset contains over 2594 Mammograms, InBreast contains 410 Mammograms, KMD dataset contains 2341 Mammograms, MIAS dataset contains 322 Mammograms and SureMaPP dataset contains 339 Mammograms, totalling to 6006 Images as complete PNG set of Mammograms. All the datasets contain Mammograms divided into 3 classes, namely ***Normal, Benign and Malignant.*** Fig. 3 illustrate the 3 classes of Mammograms, wherein the images features are evident via the bright spots being the cancerous and non-cancerous calcifications in the Benign and Malignant class of image. In the normal class image, the bright lines are the veins the in the breast.



(a) Benign



(b) Malignant

(c) Normal

Fig. 2: Mammogram Class Images

3.2 Dataset Pre-Processing

Following the collection and combination of various Datasets, the images from this dataset were converted to .PNG format to retain the original image quality and keep the main features and ROI intact without lost. After such conversion, the unlabeled images were removed and only the 3 classes were retained. The images were then divided into train and test folder which were then downscaled to 228x228 dimensions and converted to NumPy arrays for subjecting them to model training and testing. Also, the images collected were subject to Computer vision techniques for image analysis and image inference. Fig. 3 illustrates the dataset pre-processing steps followed for the employed image dataset.

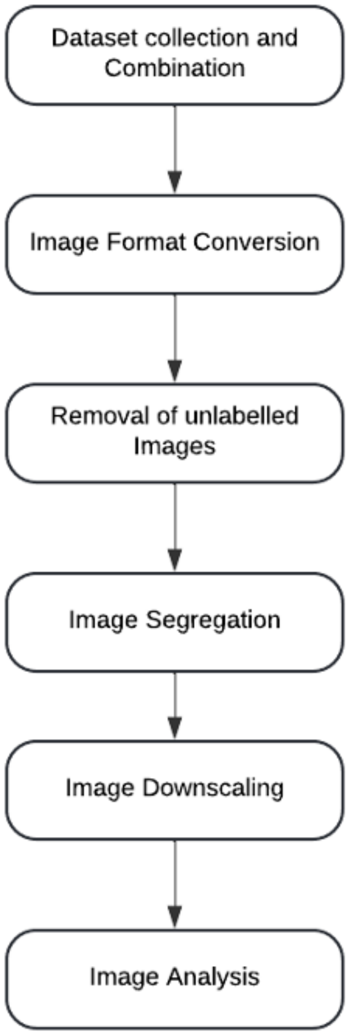


Fig. 3: Image Processing flow chart

The steps followed in the image pre-processing has been elaborated below,

3.2.1 Dataset Collection and Combination

Due to a lack of patient strength across various clinics and hospitals, Mammograms are less compared to the need for computer vision. Hence, datasets need to be independently published by independent doctors and hospitals. Hence, the need for collection and combination.

3.2.2 Image Format Conversion

Various Datasets related to Medical Domain have DICOM image format, hence, to support DL Model processing, the images are needed to be converted to supported.jpg/ .png format. However, to retain the image quality and avoid loss of features all the images are kept in .PNG format only

3.2.3 Removal of unlabeled Images

Unlabeled Images can create training conflict; hence they are removed.

3.2.4 Image Segregation

Based on the image class labels, all the images are segregated, further subject to train-test split. The segregation was done via an automated python script.

3.2.5 Image Downscaling

To support DL Architecture design of the transfer learning model, the high-resolution images are downscaled to 224 x 224 for reduced model size and faster processing via OpenCV cv2.resize(function). Post segregation, downscaling was automated via python scripted using OpenCV image processing library which also included conversion of the image into NumPy array.

3.2.6 Image Analysis

For a projected increase in model efficiency, images are analyzed for the empty region to reduce redundant or unimportant image areas. For this task techniques such as image segmentation and contour mapping-based background removal and cropping, were studied

3.3 Transfer Learning

The fundamental idea behind transfer learning is to use a complex and successfully pre-trained model, trained from a large amount of source data, such as ImageNet, and then "transfer" the learned knowledge to a relatively simple task with a small amount of data. To define the concept of transfer learning mathematically, assume the source data δ_S represents ImageNet, and the source label σ_S represents the 1000-category labeling of the ImageNet dataset and f_S denotes the source objective-predictive function,

$$\Gamma = \{\delta_s, \sigma_s, f_s\} \qquad (1)$$

$$\Phi = \{\Delta_T, \Sigma_T, f_T\} \qquad (2)$$

In our case f_S is set of predictive models such as DenseNet, XceptionNet, MobileNet and InceptionResNet. Now for our problem statement, the target triple is such, data Δ_T represents the augmented training set, Σ_T represents the three-class labeling (normal, malignant, and benign breast cancer), and f_T represents the classifier to be established. From Eqn. 1 and 2, resolution for the problem statement for the predictive-classification model training would be,

$$\partial_T = f_T(\Delta_T, \Sigma_T \mid \Gamma) \qquad (3)$$

Where ∂_T denotes the transferred learning model subjected to training and testing over the collection dataset, herein the function would map to the set of predictive models under employment. The basic idea of transfer learning is illustrated in Fig. 4.

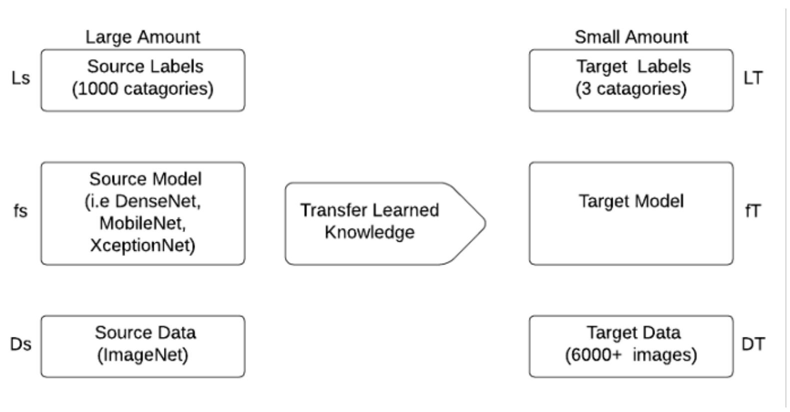


Fig. 4: Basic idea of Transfer Learning

For our problem statement we employed 5 models namely,

1. ***DenseNet121***
2. ***DenseNet169***
3. ***XceptionNetv1***
4. ***MobileNetv2***
5. ***InceptionResNetv2***

3.3.1 XceptionNet

In this study, the Breast Cancer Images were classified using the Xception ImageNet pre-trained model on the ImageNet dataset. Bottleneck features are extracted from the first layer up to the Global Average Pooling layer. The transfer techniques used in this work replaced the pre-trained model's softmax layer with a new randomized weight softmax layer, and then the bottleneck features were used to train the new softmax layer to classify the Breast Cancer Images.

Why do we use Xception??

Because of the strided convolution architecture, the computation time is reduced heavily, hence considered for image classification purposes for large-sized image datasets.

3.3.2 DenseNet

Huang et al. [19] created a CNN model with many links. To achieve a high level of information flow between the levels, each tier in the network is intrinsically linked to every other layer in a feedforward approach. Each layer's extracted features are used as entries, and the specific extracted features are used as inputs for all subsequent layers. DenseNets can be used to solve the problem, which also reduces the number of parameters significantly [23]. The integration of DenseNet-121 and SVM networks was used to generate the DenseNets model with 121 layers, as reported by Huang et al. [23], to classify and detect human postures. ImageNet pre-trained weights are also injected into the approach. Finally, a new FC model was created, this time with the study's unique SoftMax as the top layer. The goal of SVM was to introduce non-linearities into the structural model by using kernel functions, which can improve prediction capacity by allowing the entry of large amounts of data into a small space of feature space [24]. The radial basis kernel function was applied to the DeneSVM model's SVM classifier in the study.

Why do we use DenseNet??

Images utilized are in the range of size 224x224 wherein convolution is a necessity, wherein due to dense convolution block-based architecture, DenseNet is highly suitable for the purpose of Image Classification. Herein DenseNet121 and DenseNet169 are used, wherein the 2 differ with variation in the number of layers in the Convolution Block.

3.3.3 MobileNet

The MobileNet model's weights and features are pre-trained in the source domain ImageNet [28] dataset, then transferred to the target domain for Breast Cancer Classification. Many convolutional layers, pooling layers, and FC-1024 are used in the pre-training MobileNet model. The hidden layer of the FC layer has 1024 neurons and 28 layers (1+21+3+1=28) and is used as the feature extraction layer for welding defects. For improving the accuracy of welding defect classification, the defect classifier includes a Fully Connection layer FC-128 (new layer) and Softmax classifiers. As a result, the TL-MobileNet has 29 layers. The most important component of MobileNet is the Residual Connection Block (RCB). The RCB-1 and RCB-2 structures are used to prevent gradient explosion in the pre-trained MobileNet model. After inputting the first convolution layer Conv3-32, the multiple RCB-1 and RCB-2 blocks are superimposed for welding defect classification. The Conv3-128 indicates that the convolutional layer's filter size is 3x3 and its depth is 128. Conv1-128 denotes that the convolutional layer's filter size is 1x1. In the full connection layer, FC-128 represents 128 neurons. It is worth noting that the Conv1-512 structure has 5 layers.

Why do we use MobileNet??

It is a suitable consideration because of its depth-wise spatial separable convolution blocks. This model architecture has comparatively less runtime.

3.3.4 Inception-ResNet-V2

Figure ___ depicts the InceptionResNetV2 model, which has the same stem layer as the InceptionV4 model, and the rest is made up of (a) 35x35 grid cells. (a) InceptionResNet-A module, (b) 35x35 to 17x17 Reduction-A modules, and (c) 17x17 grid InceptionResNet-B module, (d) 17x17 to 8x8 Reduction-B modules, and the final InceptionResNet-C module (e) 8x8 grid [33]. Following this main architecture, Global Average Pooling was added rather than flatten to prevent overfitting in the convolutional structure due to the lack of parameters to optimize and reinforce the link between feature importance and label category [34]. In addition, when compared to the Flatten method, Global Average Pooling is more parameter efficient. Following that, Szegedy, Ioffe, Vanhoucke, and Alemi [33] added a Dropout layer with a fixed value of 0.8.

Why do we use Inception-ResNet-V2 ?

Selection of the Inception Model along with infused ResNet was considered because of Stride Convolution Blocks in the Model Architecture & the residual connections which incorporate the previously learned features of the image into the high-depth features of the convolution block.

3.4 Deep Ensemble Learning

Individual models have a bottleneck on how the image data is understood and inferred due to differing deep learning convolutional architectures. Hence, to overcome such issue, best provided solution is via deep ensemble learning wherein the combined inference power is utilized to make better and more precise inference of the image subjected as input to the various models. The bias introduced by aggregating predictions from several neural networks balances out the variation of a single trained neural network model. The results are estimates that rely less on the specifics of the training data, the training scheme used, and the luck of a single training run. It can be said that deep ensemble learning is a Meta approach wherein via combinative predictive modelling, learning is established involving power of multiple models with different bias in learning of prediction, hence making the output better than single model. For the problem at hand, we employed and studied 2 ensemble modelling techniques namely,

3.4.4 Dirichlet Ensemble Modelling

It is a Bayesian distribution based weighted average ensemble modelling technique wherein the ensemble is trained for N iterations wherein the model weightage in final inference is computed using the multi-variate beta distribution to increase the prediction accuracy of the ensemble.

3.4.5 Stack Ensemble Modelling

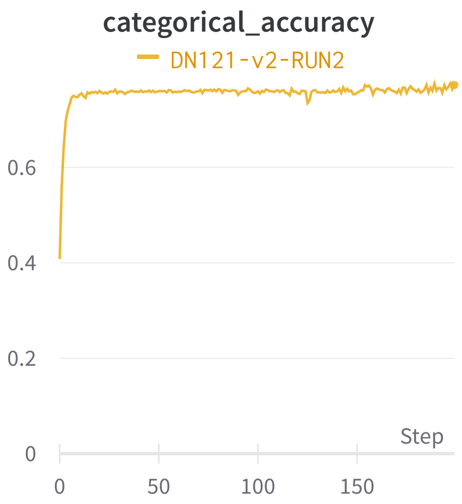
A generalization-based approach, stack ensemble modelling involves equal weightage to all the models stacked in the ensemble. By replacing the linear weighted sum (e.g., linear regression) model used to integrate the predictions of the sub-models with any learning process, this method may be further generalised. Stacked generalization, or stacking for short, is the name of this technique.

Out of the above-mentioned techniques Dirichlet Ensembling is theoretically claimed to give better results, since integration of multi-variate beta distribution helps in unbiased calculation of the model weightage in the final ensemble model inference.

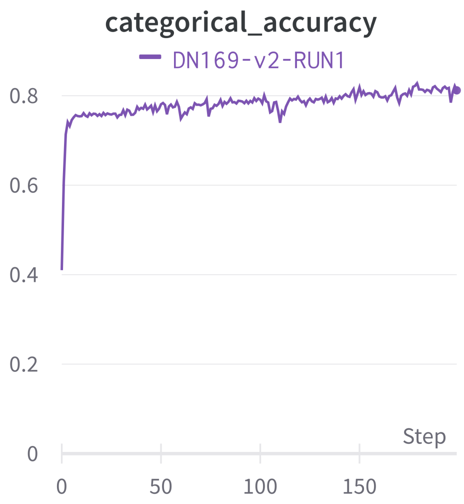
4 Result Analysis and Discussion

Subjecting the image data to training for the 5 transfer learning models for 200 epochs with 0.05 Learning rate and with ADAM optimizer we achieved the average training accuracy of 83% for all the models with average validation categorial accuracy of 53%. Below Mentioned Graphs map the epochs for the categorical accuracy. Using 3 Model and 4 Model combination we achieved Dirichlet Ensemble Model Validation Accuracy of 80% over 4500 Images and Stacking Ensemble Model Accuracy of 64% over 4500 Images.

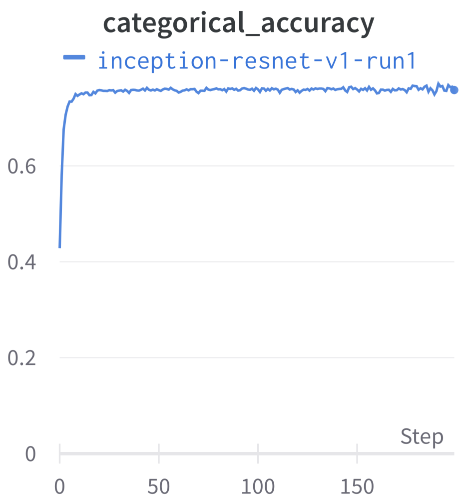
4.1 Categorical Accuracy



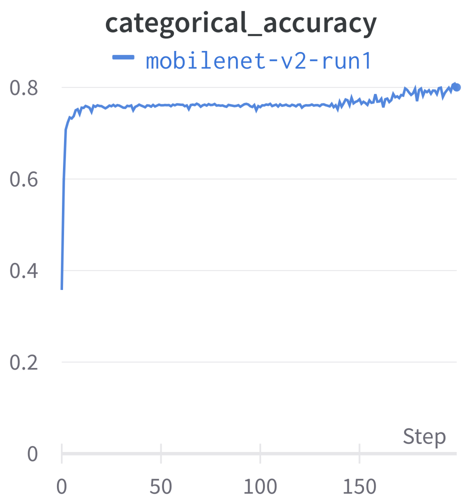
(a) DenseNet121



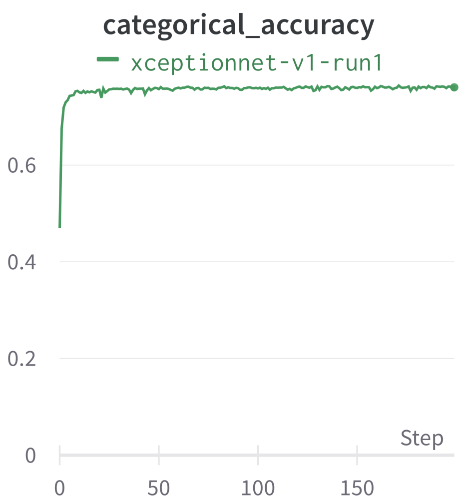
(b) DenseNet169



(c) InceptionResNetv2



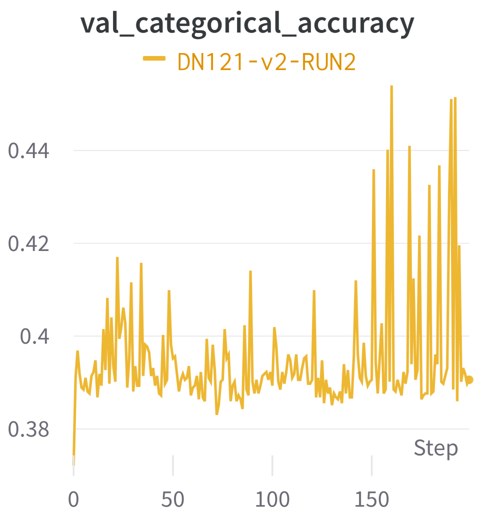
(d) MobileNetv2



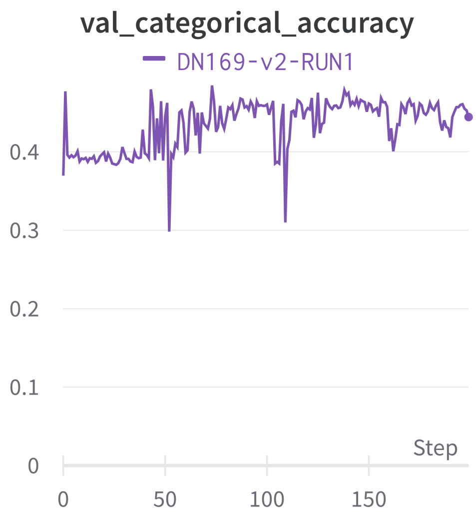
(e) XceptionNet

Fig. 5: Categorical Accuracy Comparison : Accuracy v/s Epoch

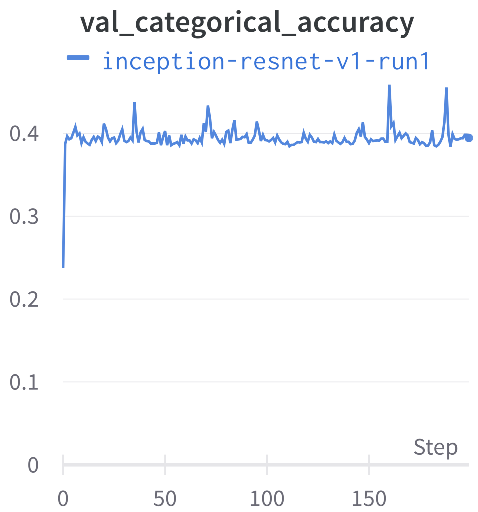
4.2 Validation Categorical Accuracy



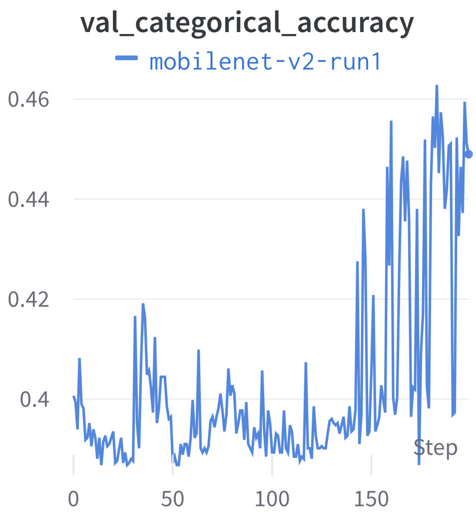
(a) DenseNet121



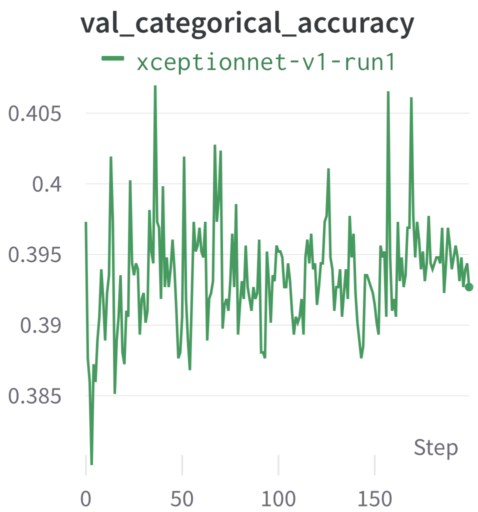
(b) DenseNet169



(c) InceptionResNetv2



(d) MobileNetv2

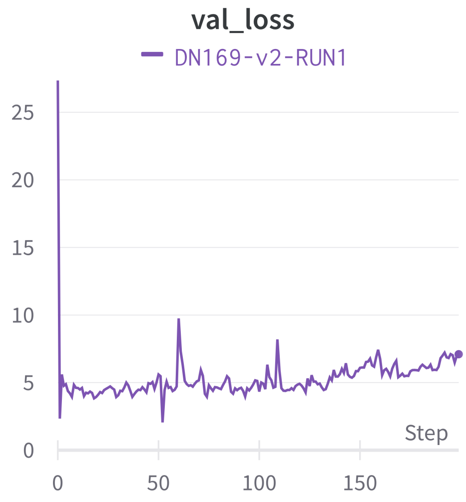


(e) XceptionNet

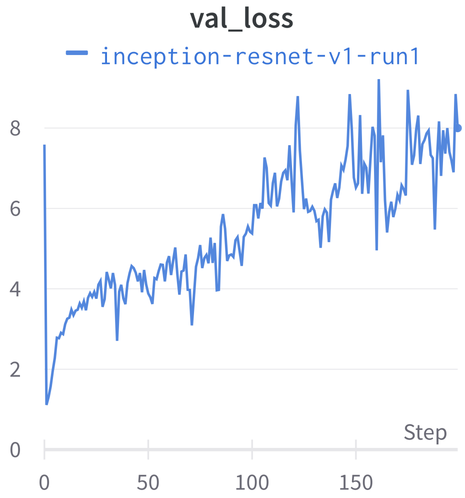
Fig. 6: Validation Categorical Accuracy Comparison : Accuracy v/s Epoch

4.3 Loss

(a) DenseNet121

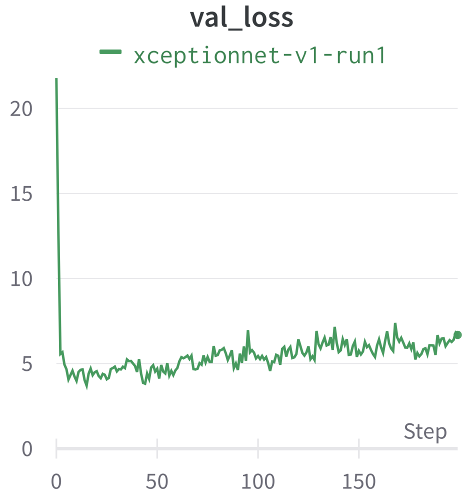


(b) DenseNet169



(c) InceptionResNetv2

(d) MobileNetv2



(e) XceptionNet

Fig. 7: Validation Loss Comparison : Loss v/s Epoch

4.4 Dirichlet Ensemble Modelling

To understand and compare the individual performance of the models in ensemble model, 3 Model Combinations and 4 Model combination were employed wherein Combinations involving DenseNet models gave the best results and for Combination involving InceptionResNetv2, MobileNetv2 and XceptionNet, the model accuracy was lowest compared to other ensemble combinations. It is also observed that in both 3-Model and 4-Model combinations InceptionResNet has the lowest weightage in the Ensemble model, hence, inferring that the model efficiency is low compared to other transfer learning model. It is also observed that DenseNet121 and XceptionNet has highest weightage in the Ensemble Model combinations with Ensemble model accuracy ~80%. Table II tabulates the results obtained.

TABLE II: Dirichlet Ensemble Model Accuracy and Combination Model weightage

Ensemble Model	3 - Model Ensemble Weightage					Accuracy
	DenseNet121	DenseNet169	InceptionResNetv2	MobileNetv2	XceptionNet	
3-Model Combination 1	0.9598	0.0390	0.0012			77.94
3-Model Combination 2	0.9589	0.0373		0.0039		79.00
3-Model Combination 3	0.1806	0.0083			0.8111	78.04
3-Model Combination 4			0.0083	0.1729	0.8189	74.20

Ensemble Model	4 - Model Ensemble Weightage					Accuracy
	DenseNet121	DenseNet169	InceptionResNetv2	MobileNetv2	XceptionNet	
4-Model Combination 1	0.0002	0.3346	0.0129	0.6523		78.84
4-Model Combination 2	0.7349	0.0280		0.0025	0.2346	78.95
4-Model Combination 3	0.4223	0.0261	0.0000		0.5515	77.93
4-Model Combination 4	0.6850		0.0002	0.0014	0.3134	75.53
4-Model Combination 5		0.3509	0.0018	0.6464	0.0009	78.91

4.5 Stacking Ensemble Modelling

In stacking ensemble modelling, final meta estimator utilized is Linear Regression hence the final classification result can be inferred to possess equal weightage of all employed Models in the ensemble combinations. Hence the overall ensemble model result is comparatively less with the average accuracy for all ensemble models being ~65%. The reason for this is the fact that majority of the models utilized in the model have a low accuracy of ~75% with high bias which can lead to lowered predictive classification. It is observed that Ensemble Model combination involving DenseNet models had higher accuracy compared to Ensemble model incorporating InceptionResNetv2. Table III tabulates the results obtained.

TABLE III: Stacking Ensemble Model Accuracy

3 - Model Ensemble Weightage						
Ensemble Model	**DenseNet121**	**DenseNet169**	**InceptionResNetv2**	**MobileNetv2**	**XceptionNet**	**Accuracy**
3-Model Combination 1	**Yes**	**Yes**	**Yes**			62.71
3-Model Combination 2	**Yes**	**Yes**		**Yes**		62.37
3-Model Combination 3	**Yes**	**Yes**			**Yes**	63.38
3-Model Combination 4			**Yes**	**Yes**	**Yes**	58.86

4 - Model Ensemble Weightage						
Ensemble Model	**DenseNet121**	**DenseNet169**	**InceptionResNetv2**	**MobileNetv2**	**XceptionNet**	**Accuracy**
4-Model Combination 1	**Yes**	**Yes**	**Yes**	**Yes**		62.71
4-Model Combination 2	**Yes**	**Yes**		**Yes**	**Yes**	63.04
4-Model Combination 3	**Yes**	**Yes**	**Yes**		**Yes**	65.22
4-Model Combination 4	**Yes**		**Yes**	**Yes**	**Yes**	57.69
4-Model Combination 5		**Yes**	**Yes**	**Yes**	**Yes**	63.88

Table IV tabulates the Individual Model Validation accuracy for 4300 images. Highest classification performance was of MobileNet.

TABLE IV: Individual Validation Model Accuracy

Transfer Learning Model	Accuracy
DenseNet121	0.5867
DenseNet169	0.7370
InceptionResNetv2	0.7180
MobileNet	0.7392
XceptionNetv2	0.5226

5 Conclusion

Major focus of this project is to study and compare various Ensemble Learning technique on top of transfer learning, which can then be used to perform important and sensitive tasks such as diagnosing cancer in Breast using Mammogram as the diagnostic medium. In this project we have trained and tested 5 transfer learning models, DenseNet 121-169 – InceptionResNetv2-MobileNet- XceptionNet. Individually we achieved average training accuracy of ~80% for the transfer learning models and Ensemble Model accuracy of ~80%. However, individual model accuracies of the transfer learning models are comparatively low due to diminished ROI and image feature due to conversion of 5000x6000 pixel dimension mammogram to 224x224 pixel dimension input feature vector.

6 Future Work

Mammograms possess a large black area diminishing the actual breast x-ray when the mammogram is resized to a smaller dimension for processing. For better results in computer vision task and deep learning based convolutional model, cropping and area removal techniques must be employed. To overcome this limitation, we propose implementation of Image Processing technique such as grab-cut-mask based Contour cropping for extracting the ROI of the mammogram and SLIC Image Segmentation to incorporate ROI highlight enhancement.

7 References

[1] Malebary, S. J., & Hashmi, A. (2021). Automated breast mass classification system using deep learning and ensemble learning in digital mammogram. *IEEE Access*, *9*, 55312-55328.

[2] Kadam, V. J., Jadhav, S. M., & Vijayakumar, K. (2019). Breast cancer diagnosis using feature ensemble learning based on stacked sparse autoencoders and softmax regression. *Journal of medical systems*, *43*(8), 1-11.

[3] Sun, R., Meng, Z., Hou, X., Chen, Y., Yang, Y., Huang, G., & Nie, S. (2021). Prediction of breast cancer molecular subtypes using DCE-MRI based on CNNs combined with ensemble learning. *Physics in Medicine & Biology*, *66*(17), 175009.

[4] Ghiasi, M. M., & Zendehboudi, S. (2021). Application of decision tree-based ensemble learning in the classification of breast cancer. *Computers in Biology and Medicine*, *128*, 104089.

[5] Sun, R., Hou, X., Li, X., Xie, Y., & Nie, S. (2022). Transfer Learning Strategy Based on Unsupervised Learning and Ensemble Learning for Breast Cancer Molecular Subtype Prediction Using Dynamic Contrast-Enhanced MRI. *Journal of Magnetic Resonance Imaging*, *55*(5), 1518-1534.

[6] Abdar, M., & Makarenkov, V. (2019). CWV-BANN-SVM ensemble learning classifier for an accurate diagnosis of breast cancer. *Measurement*, *146*, 557-570.

[7] Moon, W. K., Lee, Y. W., Ke, H. H., Lee, S. H., Huang, C. S., & Chang, R. F. (2020). Computer-aided diagnosis of breast ultrasound images using ensemble learning from convolutional neural networks. *Computer methods and programs in biomedicine*, *190*, 105361.

[8] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2818-2826).

[9] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4700-4708).

[10] Kele ̧s, S.; Günlü, A.; Ercanli, I. Estimation of the carbon of the aboveground stand by combining Sentinel-1 and Sentinel-2 satellite data: A case study from Turkey. In Forest Resources Resilience and Conflicts; Elsevier: Amsterdam, The Netherlands, 2021; pp. 117–126

[11] Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009, June). Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition* (pp. 248-255). Ieee.

[12] Szegedy, C., Ioffe, S., Vanhoucke, V., & Alemi, A. A. (2017, February). Inception-v4, inception-resnet and the impact of residual connections on learning. In *Thirty-first AAAI conference on artificial intelligence*.

Minor Project Report: Rushir Bhavsar, Smit Shah

by smitaagrawal

General metrics

26,671	3,848	328	15 min 23 sec	29 min 36 sec
characters	words	sentences	reading time	speaking time

Score

83

This text scores better than 83% of all texts checked by Grammarly

Writing Issues

159	100	59
Issues left	Critical	Advanced

Plagiarism

4 %

13 sources

4% of your text matches 13 sources on the web or in archives of academic publications

Unique Words

25%

unique words

Measures vocabulary diversity by calculating the percentage of words used only once in your document

Rare Words

42%

rare words

Measures depth of vocabulary by identifying words that are not among the 5,000 most common English words.

Word Length

5.2

characters per word

Measures average word length

Sentence Length

11.7

words per sentence

Measures average sentence length