Breast Cancer Detection: A Deep Learning Approach

Samruddhi Raut- *raut.sa@northeastern.edu*

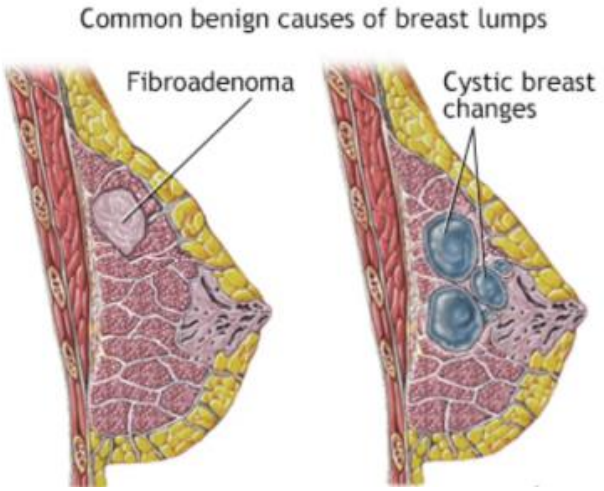
Vamsi Konyala - *konyala.v@northeastern.edu*

***Abstract*—** This research paper addresses the significance of early detection and treatment in reducing breast cancer mortality, considering it as the second leading cause of death in women. Mammography, a pivotal tool in breast cancer screening, aids in identifying early breast lumps or calcification regions. However, challenges arise in detecting cancer masses within exceptionally dense breast tissue, highlighting a limitation of traditional mammography. To overcome this, the paper employs deep learning modules to enhance detection accuracy, as manual methods are prone to human error. The study focuses on utilizing Convolutional Neural Networks (CNNs) to extract features for the detection of cancer cells. Specifically, the research involves training and classifying the INbreast and MIAS datasets into benign or malignant categories using a CNN model developed from scratch. Additionally, a comparative analysis is conducted with various pre-trained convolutional neural networks, including ResNet50, MobileNet-v2, Inception-v3, and Densenet121. The findings reveal that the CNN model developed in the study outperforms the pre-trained models, achieving a remarkable prediction accuracy of up to **97.45% for INbreast and 92.02% for MIAS** dataset. This suggests the effectiveness of the proposed CNN model in breast cancer detection, emphasizing its potential contribution to improving the accuracy and efficiency of breast cancer screening methods.

Keywords - *CNN, transfer learning, pretrained models, mammography, benign, malignant, maxpooling, batch normalization*

I. INTRODUCTION

Breast cancer, a pervasive health concern affecting one in eight women globally, arises from uncontrolled cell growth and typically initiates in the milk-carrying ducts. The potential for metastasis underscores the urgency for early detection and intervention. Screening mammograms, notably full-field digital mammography (FFDM), serve as the gold standard, offering a vital means to identify cancerous breast tumors during their incipient stages. However, the challenge lies in discerning these minute cancerous regions within comprehensive breast images, often measuring 227 x 227 pixels, with potential cancerous areas occupying as little as 50 x 50 pixels. This research project addresses the imperative of accurate breast cancer detection using advanced image analysis techniques. With a substantial dataset comprising 24,518 FFDM images, our objective is to develop a model capable of autonomously discerning the presence of cancerous tissues within a patient's breast patch. The motivation for this endeavor stems from the high prevalence of incorrect mammography diagnoses, which can lead to unnecessary treatments or, conversely, a lack of necessary interventions, thereby jeopardizing patient well-being.



The implementation of this model carries profound implications for breast cancer diagnosis and treatment. By providing physicians with a tool for early detection, we aim to maximize patient survival rates. Additionally, our model seeks to minimize the impact of misinterpretations by reducing the number of "untrained eyes" involved in the screening process, thereby enhancing accuracy. Ultimately, the model has the potential to prevent delayed treatments and unnecessary interventions by refining the precision of breast cancer screenings. This research thus contributes to the broader mission of improving healthcare outcomes and ensuring a more reliable and efficient breast cancer screening process.

A. Motivation

The personal journey of my friend's mother battling early-stage breast cancer has served as a poignant reminder of the profound impact this disease can have on individuals and their families. Witnessing her transition from rural to urban areas for treatment underscored the logistical challenges that many face in accessing specialized care. While she has shown remarkable resilience and is now undergoing radiation therapies, the experience illuminated the critical importance of early detection and effective diagnostic tools.

As a woman, I advocate for periodic mammogram screenings every five years to ensure the health of breast ducts. Despite sustained research efforts, the accuracy of risk models employed in clinical practice remains constrained. Pathologists, engaged in the laborious task of reviewing mammography test results to identify Invasive Ductal Carcinoma (IDC), encounter a manual and time-intensive process that is contingent upon their expertise and equipment. This scenario serves as the impetus for our research. We are driven to develop an efficient model capable of discerning the likelihood of masses and the probability of breast cancer in patients. This initiative aims to alleviate the dependence on pathologist expertise, particularly in regions lacking specialized professionals, by introducing a reliable and automated diagnostic solution.

B. Approach

This paper presents the Convolutional Neural Network (CNN) methodology. The CNN architecture, as delineated in [5], incorporates two principal transformations. The initial transformation involves convolution, wherein pixels undergo convolution with a filter or kernel. The second pivotal transformation is sub-sampling, which may take various forms such as max-pooling, min-pooling, and average-pooling, depending on specific requirements. The methodological framework encompasses image processing, feature extraction, and classification. Additionally, a comparative analysis is conducted, juxtaposing our model with pretrained models to assess its performance on the Inbreast and MIAS datasets. The evaluation of the system's performance is based on metrics including Area Under the Curve (AUC), specificity, accuracy, sensitivity, and F-1 score. The dataset to be used in this project is put together by Mendeley Data consisting of INbreast and MIAS dataset. It is free to use and open source. The dataset contains mammography with benign and malignant masses.

II. BACKGROUND

A. Current State-of-Art

A team of researchers affiliated with MIT's Computer Science and Artificial Intelligence Laboratory (CSAIL) has developed an anticipatory analytics model designed to project a patient's susceptibility to breast cancer across various temporal intervals. Although this model is still in the developmental phase, healthcare practitioners from institutions such as Novant Health in North Carolina, Emory in Georgia, Maccabi in Israel, TecSalud in Mexico, Apollo in India, and Barretos in Brazil have invested in its potential for early detection. By considering diverse factors including age, hormonal influences, genetic predispositions, and breast density, the model assesses the necessity for additional examinations and establishes optimal screening frequencies for women. Despite sustained research endeavors, the accuracy of risk models employed in clinical settings has encountered limitations, as indicated by pertinent published data. While predictive analytics and other artificial intelligence techniques exhibit promise in forecasting cancer risk, they often face challenges in adapting to novel patient demographics, particularly in the intricate task of identifying masses within dense breast tissue..

B. Literature Review

Authors in [13] utilized four-fold cross-validation on X-ray mammograms from the INbreast dataset to estimate a full-resolution convolutional network (FrCN). It resulted in an F1 score of 91.24%, an accuracy of 92.96%. In another study [14], the BDR-CNN-GCN approach was proposed by combining a graph-convolutional network (GCN) with a basic 8-layer CNN that includes batch normalization and dropout layers. The final BDR-CNN-GCN model was formed by integrating the two-layer GCN with the CNN. This method was tested using the MIAS dataset, and successful results were obtained with a 89.10% accuracy level. In [15], a deep-CNN model that utilized transfer learning (TL) was introduced to prevent overfitting when working with small datasets. DDSM, MIAS, BCDR, and INbreast were used to assess its performance. INbreast dataset achieved an accuracy of 91.5%, the DDSM dataset achieved an accuracy level of 88.35%, and the BCDR database achieved a 91.67% accuracy level.

Authors in [16] presented a hybrid classification model using Mobilenet, ResNet50, and Alexnet with an accuracy level of 91.6%. In [17], four different CNN architectures (VGG19, InceptionV3, ResNet50, and VGG16) were utilized for model training using 5000 images, while prediction models were evaluated on 1007 images. Authors in [18] utilized alpha, geostatistics, and diversity analyses forms in their proposed breast cancer detection method. They employed the SVM classifier on MIAS and DDSM databases, which resulted in a detection accuracy level of 89.30%. The SVM classifier and gray level co-occurrence matrix (GLCM) were employed by [19] for detecting breast cancer abnormalities in the MIAS data set. Their method achieved an accuracy of 91.88% and surpassed the performance of the k-nearest neighbor (kNN) algorithm.

Although prior breast cancer detection and classification systems have improved information extraction, several issues still need attention, such as low contrast in tumor location, high memory complexity, long processing time, and the need for a large amount of training data for deep learning

approaches. In response to these problems, we propose a new approach to breast cancer detection and classification, which will be discussed in detail in the following section.

III. APPROACH

This study evaluates the effectiveness of a convolutional neural network (CNN) architecture in the detection of breast cancer. The model is trained using labeled data consisting of both benign and malignant input images, with 2480 benign and 5429 malignant samples from the RGB color model. CNN is trained separately for the MIAS and INbreast datasets. Furthermore, a comparative analysis is conducted, benchmarking our CNN model against various pre-trained convolutional neural networks such as Densnet121, ResNet50, MobileNet-v2, Inception-v3 to assess its performance relative to pre-existing models.

A. Convolutional Nueral Network(CNN)

CNNs are deep neural networks used for image recognition and classification. In recent years, CNNs have become a crucial tool in image analysis, especially for identifying faces, text, and medical imaging. CNNs have a long history of success in image classification and segmentation, first developed in 1989. CNNs replicate the human brain's visual information processing by incorporating layers of "neurons" that only respond to their local surroundings. These networks can understand the topological aspects of an image through a combination of convolutional, pooling, and fully connected (FC) layers.

a. Image Processing

A substantial portion of image pixels is often redundant and does not significantly contribute to image information. To mitigate unnecessary computational overhead, compression techniques are employed, aiming to eliminate redundancy in the input data that only adds to the network's computational complexity. Additionally, image augmentation is implemented on the datasets to prevent overfitting of the CNN model. This involves various transformations applied to the image datasets, including rotation, reshaping, shifting, shearing, and zooming. The objective is to augment the dataset, providing the model with a diverse set of images for training and learning various features. This diverse set of images with different perturbations enhances the model's ability to generalize effectively.

b. Feature Extraction

Feature extraction plays a pivotal role in neural networks, specifically in processing image patterns and textures [8]. In the context of detecting malignant tumors, identifying characteristics like large and irregular nuclei or multiple nuclear structures is crucial. In supervised learning, images, along with their assigned class labels (Benign or Malignant), are input into an architecture such as Convolutional Neural Network (CNN). During the training process, the CNN extracts computational features from filter values. To clarify, the features comprise the filter weights of a given CNN architecture, which are utilized during testing for model evaluation. CNN takes the raw pixels of an image and generates output in the form of learned filter weights. These weights are then used as input for the dense architecture in the deep neural network during the final prediction. This paper introduces a CNN model with a combination of two types of hidden layers, involving convolutional and pooling layers, as detailed in Table.1. The sequential model's architecture initiates with the first two layers defined as Convolution - Max pooling – ReLu activation function. The output layer adopts the softmax activation function. The final step involves classification, where an image is assigned to its respective class (benign or malignant). This classification can be achieved using a support vector machine or a fully connected layer employing an activation function like Softmax.

c. Classifiaction

Classification involves taking the flattened weighted feature map, acquired from the final pooling layer, and utilizing it as input for the fully connected network. This network is responsible for computing the loss and adjusting the weights of the internal hidden nodes. The layer attributes are clearly defined. Within the classification layer, pre-processed data is fed to the layers and the output of the last layer serves as the final result. The last layer outputs 2 nodes corresponding to benign and malignant classifications. The softmax function is employed to normalize the CNN output within the range of 0 to 1, especially suitable for multi-classification problems. The model is compiled using the Adam optimizer, with categorical cross-entropy as the loss function. Despite experimenting with other optimizers like Stochastic Gradient Descent (SGD) and RMSprop, no improvement in the model's performance was observed, given the two classes used for image data categorization.

B. Transfer Learning

Transfer learning is a machine learning approach that involves adapting a model initially trained on one task to perform a related task. Here, the DenseNet121, InceptionV3, MobileNetV2, and ResNet50 models, pre-trained on the extensive ImageNet dataset for image classification, serve as base neural networks for

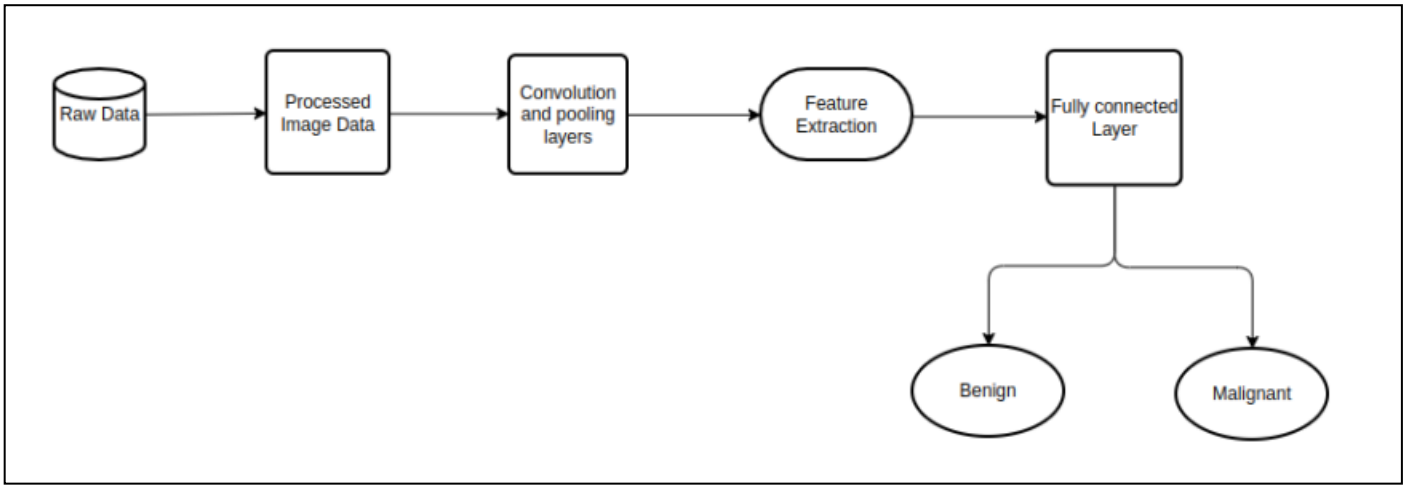


Figure.1. Classification Process

Layer	Configuration
Input	Shape: (92, 140, 3)
Conv2D 1	Filters: 32, Kernel: (3, 3), Activation: 'relu', Padding: 'same'
MaxPooling2D 1	Pool Size: (2, 2)
Conv2D 2	Filters: 32, Kernel: (3, 3), Activation: 'relu', Padding: 'same'
MaxPooling2D 2	Pool Size: (2, 2)
Conv2D 3	Filters: 64, Kernel: (3, 3), Activation: 'relu', Padding: 'same'
MaxPooling2D 3	Pool Size: (2, 2)
Dropout 1	Rate: 0.5
Flatten	-
Dropout 2	Rate: 0.5
Dense 1	Units: 64, Activation: 'relu'
Dropout 3	Rate: 0.5
Dense 2	Units: 64, Activation: 'relu'
Dropout 4	Rate: 0.5
Dense (Output Layer)	Units: num_classes, Activation: 'softmax'
Compilation	Loss: 'categorical_crossentropy', Optimizer: 'adam'

Table.1 CNN Architecture

a binary classification task specific to breast cancer detection (distinguishing between benign and malignant tumors). This technique is especially valuable when the pre-trained model has already acquired significant features from a large dataset, allowing these features to be repurposed for a different yet related task, even when the new dataset is comparatively smaller. In this research, a comparative analysis is conducted to evaluate the performance of a custom-built Convolutional Neural Network (CNN) in contrast to the aforementioned pre-trained models. The architecture details of each pre-trained model are outlined for comprehensive understanding and comparison.

Layer	Configuration
Input	Shape: (224, 224, 3)
Base Neural Network	Pre-trained model (DenseNet121, ResNet50, MobileNetV2, InceptionV3)
Flatten	-
Batch Normalization 1	-
Dense 1	Units: 256, Kernel Initializer: 'he_uniform'
Batch Normalization 2	-
Activation (ReLU)	-
Dropout	Rate: 0.5
Dense 2 (Output Layer)	Units: 2, Activation: 'softmax'
Trainable Layers	None (All layers in the base neural network are set to non-trainable)
PlotLossesKeras	Callback for live loss plotting during training

Table.2 Pre-trained Models Architecture

a. Pre-trained Models

This research focused on constructing a binary classification model for the detection of breast cancer

(INbreast and MIAS dataset) utilizing the **DenseNet121, ResNet50, MobileNetV2**, **InceptionV3** architecture pretrained on the ImageNet dataset. The sequential model is structured with several key layers - the pre-trained model base layer, a flattening layer to transform the output to a one-dimensional tensor, batch normalization for normalizing activations, a dense layer with 256 units and He uniform kernel initializer, additional batch normalization, ReLU activation for introducing non-linearity, a dropout layer to mitigate overfitting, and a final dense layer with two units and softmax activation to facilitate binary classification (benign or malignant). The pre-trained ResNet50 layers are frozen to retain the learned features. The model is compiled using the Adam optimizer, categorical crossentropy loss, and accuracy. It undergoes training on the provided training data for 100 epochs with a batch size of 16, and the training process is monitored using specified callbacks. You can find the details in Table.2. Subsequently, the model is evaluated on both the validation set and a test set. The evaluation includes the calculation of various performance metrics such as accuracy, precision, recall, F1 score.

IV. Experimental Results

A. *Dataset*

The dataset to be used in this project is put together by Mendeley Data. It is free to use and open source. The dataset contains mammography with benign and malignant masses.

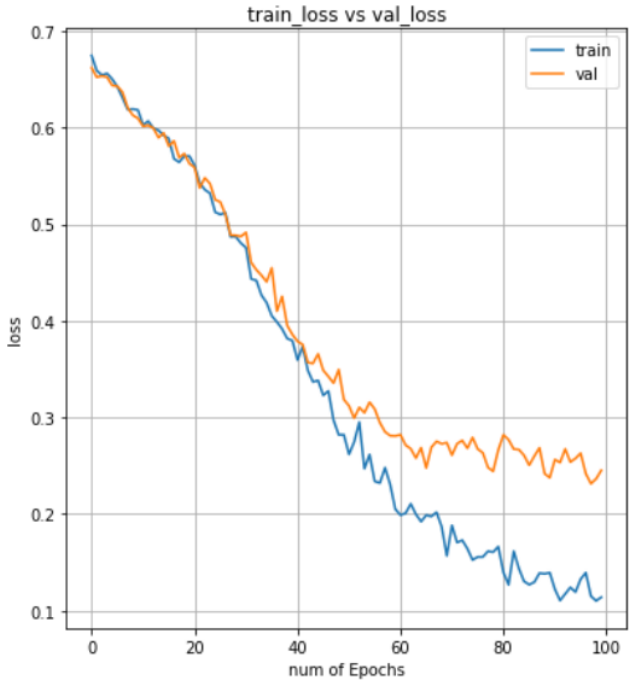
The dataset comprises mammography images encompassing both benign and malignant masses. Initially, 106 mass images were extracted from the INbreast dataset, 53 from the MIAS (Mammographic Imaging Analysis Society) dataset, and 2188 from the DDSM (Digital dataset for screening Mammography) dataset. Subsequently, data augmentation techniques and contrast-limited adaptive histogram equalization were employed for image preprocessing. Post augmentation, the Inbreast dataset encompasses 7632 images, the MIAS dataset comprises 3816 images, and the DDSM dataset includes 13128 images. All the images were resized to 227*227 pixels.

B. *Experiments and Performance Evaluation*

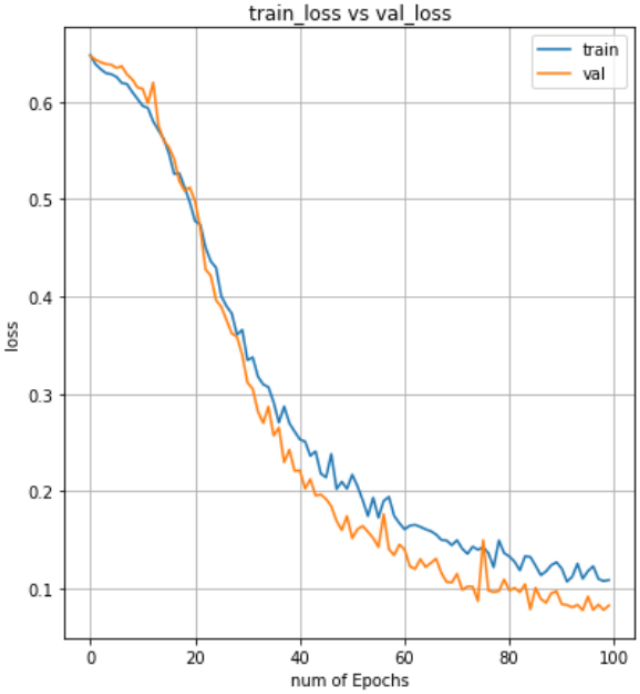
Convolutional Nueral Network(CNN) :

In light of the aforementioned configuration, we achieved a training accuracy of 92.02% & 97.45% with a test-train split ratio of 0.2 for the MIAS and INbreast dataset.

As anticipated in a neural network, the depicted losses in Figure.2 initiate with elevated values and progressively decrease during the training process. This pattern aligns with the conventional training protocol in deep learning. The discernible contrast in saturation levels between Training Loss and Validation Loss is 0.2, a margin that falls within an acceptable range, mitigating the risks of underfitting or overfitting for the network.



MIAS Dataset

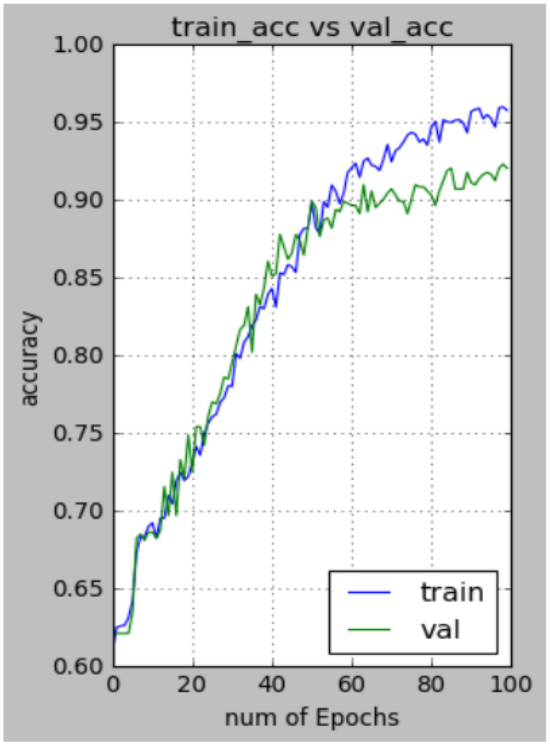


INbreast Dataset

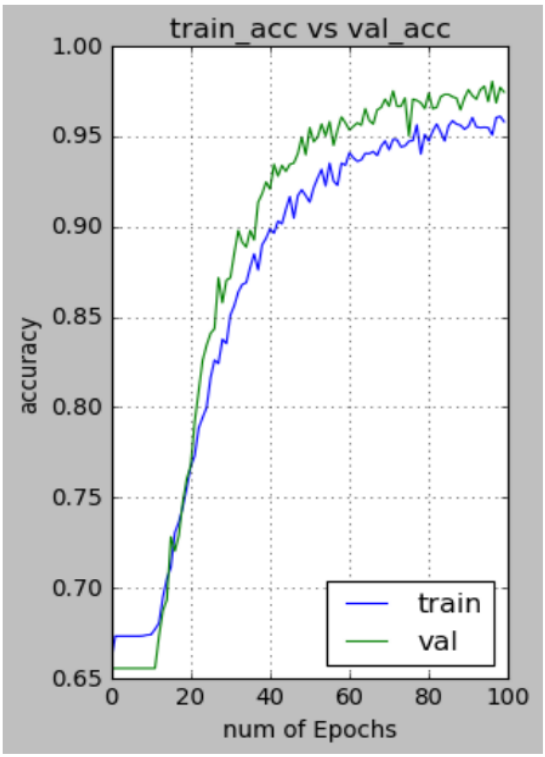
Figure.2 -Training and Validation loss.

The visual representation of the distribution of accuracy is illustrated in Figure.3. The accuracy exhibits a progressive increase with the number of epochs and eventually reaches saturation, indicating the completion of training on the dataset for the

specified network. Notably, a crucial inference from this graph is that the network undergoes training without displaying characteristics of underfitting or overfitting, as evidenced by the similarity in the distribution of validation accuracy and training accuracy curves.



MIAS Dataset



INbreast Dataset

Figure.3 - Training and Validation accuracy

Prediction results on test data:

MIAS Dataset – 92.02%

```
# Final evaluation of the model
scores = model.evaluate(X_test, y_test, verbose=1,batch_size=batchsize)
```

```
6/6 [==============================] - 2s 401ms/step - loss: 0.2450 -
accuracy: 0.9202
```

INbreast Dataset- 97.45%

```
# Final evaluation of the model
scores = model.evaluate(X_test, y_test, verbose=1,batch_size=batchsize)
```

```
12/12 [==============================] - 4s 306ms/step - loss: 0.0829
- accuracy: 0.9745
```

Testing on any random image:

```
***** TESTING ANY RANDOM IMAGE *****

Shape of test image 2:
(1, 92, 140, 3)
Predicted accuracies:
[[0.01409874 0.9859013 ]]
Predicted class:
[1]
```

As evident, an unidentified image was input into the loaded and trained model, yielding accurate predictions. The model assigned a probability of 98.59% for the image being malignant. Consequently, the predicted class is identified as malignant, denoted by the label 1.

Classification Report:

i. Precision serves as a probabilistic metric for assessing whether a positive case, as per our definition, truly pertains to the positive class.
ii. Recall acts as a probabilistic metric to ascertain whether an actual positive case is accurately classified within the positive class.
iii. The F1 score is computed as the geometric mean between precision and recall.

Formula:

$$F1=2(precision * recall) / (precision + recall)$$

	precision	recall	f1-score	support
0	0.92	0.95	0.94	467
1	0.92	0.87	0.89	285
accuracy			0.92	752
macro avg	0.92	0.91	0.91	752
weighted avg	0.92	0.92	0.92	752

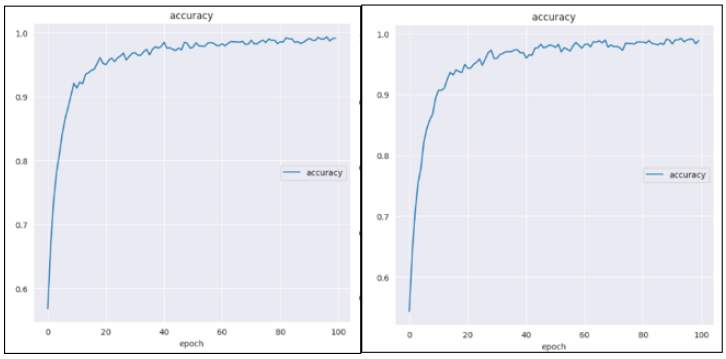
MIAS Dataset

	precision	recall	f1-score	support
0	0.95	0.98	0.96	526
1	0.99	0.97	0.98	1001
accuracy			0.97	1527
macro avg	0.97	0.98	0.97	1527
weighted avg	0.97	0.97	0.97	1527

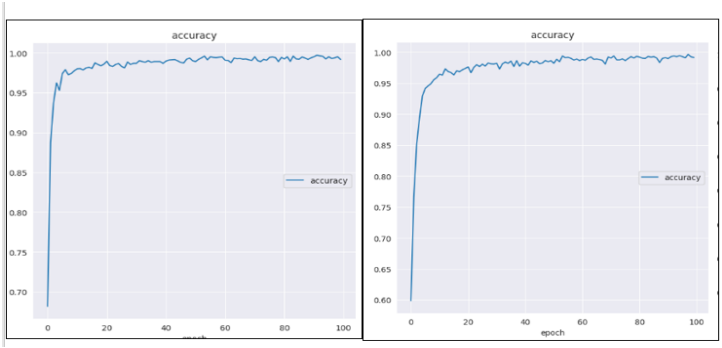
INbreast Dataset

Transfer Learning :

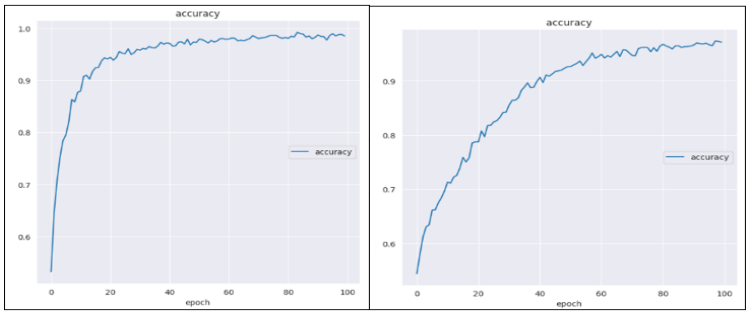
In the context of the outlined configuration detailed in Table.2, we have attained notable training accuracies, as presented in Table.3, for both the MIAS and INbreast datasets. Our validation strategy involved allocating 70% of the total data to training, 21% to validation, and 9% to testing. To mitigate overfitting, various augmentation techniques were employed. The accuracy trends across epochs for all pretrained models are visually represented in Figure.4. Upon thorough evaluation, it was discerned that among the four pretrained models utilized, ResNet50 exhibited superior performance, achieving an accuracy of 94.32% on the INbreast dataset and 90.39% on the MIAS dataset.



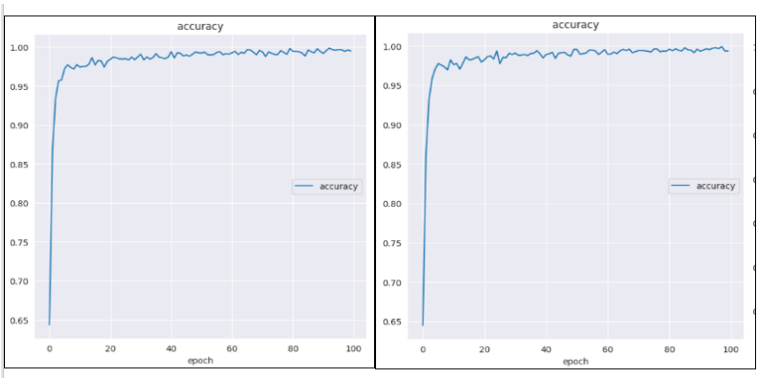
Densenet121



MobileNetV2



InceptionV3



ResNet50

Figure.4 Training accuracy for pre-trained model (Left-MIAS dataset, Right-INbreast dataset)

C. Discussion

After conducting an extensive literature review, we identified techniques outlined in Table.4 that have been previously utilized for breast cancer detection. Our primary objective was to construct a model capable of achieving the highest accuracy and precision compared to existing or past models. The results obtained from our Convolutional Neural Network (CNN) model were 92.02% for the MIAS dataset and 97.45% for the INbreast dataset. Additionally, we employed transfer learning techniques, leveraging pre-trained models such as DenseNet121, ResNet121, InceptionV3, and MobileNetV2 to evaluate and compare their performance.

Pre-trained Models	Accuracy	
	Inbreast	**MIAS**
Densenet121	93.98%	88.50%
MobileNetV2	91.03%	87.34%
InceptionV3	93.09%	88.92%
ResNet50	94.32%	90.39%

Table.3 Accuracy Matrix-Pretrained Models

Notably, our CNN model outperformed all other models. ResNet50 demonstrates a slightly superior performance among the other pre-trained models with accuracy of 94.32% Inbreast and 90.39% MIAS. The accuracy achieved by our CNN model represents a significant improvement over various state-of-the-art experiment. We systematically tested our model across various image resolutions, and the results demonstrated a remarkable insensitivity to resolution changes. Employing this automated process holds the potential to detect cancer in its early stages, contributing to an increased survival rate among breast cancer patients. To provide context, we compared our results (92.02% and 97.45% validation accuracy from the test set) with those reported in several published studies, as summarized in Table.4.

Year	Model Name	Accuracy Achieved
2017	K- Nearest Neighbor	**83%-89%**
2019	FrCN (Convolutional Network)	**90.96%**
2017	YOLOv5 Modified	**89.50%**
2019	Deep-CNN with Transfer Learning	**91.54**
2018	CNN Inception-v3	**87%**
2019	CNN and TL Classification (Eight Pretrained Models)	**85%-91%**
2021	Hybrid Model (Mobilenet, ResNet50, Alexnet)	**92.6%**
2020	Four Different CNN Architectures (VGG19, InceptionV3, ResNet50, VGG16)	**88.32%**
2019	AlexNet and SVM with Data Augmentation	**87.2%**

Table.4 Existing methods and respective Accuracy

V. Conclusion and Future scope

In conclusion, the experimental results affirm that our CNN model stands as the most effective solution, outperforming established pre-trained models, including DenseNet121, ResNet121, InceptionV3, and MobileNetV2. The application of deep learning in the realm of medical pathology, particularly in the detection of breast cancer using digital images, marks a significant milestone and presents ample research opportunities. Leveraging tools like deep learning proves instrumental in addressing such complex challenges. As part of the future scope, one can explore the implementation of autoencoders for image compression, offering an automated approach without losing key features, as autoencoders can regenerate up to 90% of the original image. Furthermore, fine-tuning the model architecture and hyperparameters holds potential for achieving even higher accuracy. This study successfully classified breast cancer tissues into benign and malignant categories, achieving an accuracy of 92.02% for the MIAS dataset and an impressive 97.45% accuracy for the INBreast dataset.

Acknowledgment

We are highly grateful to Professor Jerome Braun for the guidance and support throughout the project. The course material was very useful and came absolutely handy while working on the breast cancer detection project.

References

[1] http://peipa.essex.ac.uk/info/mias.html

[2] A. Alias, B. Paulchamy, Detection of breast cancer using artificial neural network, International Journal of Innovative Research in Science 3 (3).

[3] A.F. Agarap On breast cancer detection: an application of machine learning algorithms on the Wisconsin diagnostic dataset, CoRR abs/1711.07831

[4] Y. LeCun, Y. Bengio, G. Hinton Deep learning, nature, 521 (7553) (2015), p. 436, 10.1038/nature14539 https://doi.org/10.1038/nature14539

[5] F. Xing, Y. Xie, L. Yang An automatic learning-based framework for robust nucleus segmentation IEEE Trans Med Imaging, 35 (2) (2016), pp. 550-566, 10.1109/TMI.2015.2481436 https://doi.org/10.1109/TMI.2015.2481436

[6] R.C. González, R.E. Woods, S.L. Eddins Digital image processing using MATLAB Pearson (2004)

[7] R. Geirhos, P. Rubisch, C. Michaelis, M. Bethge, F.A. Wichmann, W. Brendel Imagenet-trained CNNs are biased towards texture; increasing shape bias improves accuracy and robustness International conference on learning representations (2019) https://openreview.net/forum?id=Bygh9j09KX

[8] M. Sokolova, N. Japkowicz, S. Szpakowicz Beyond accuracy, f-score and ROC: a family of discriminant measures for performance evaluation AI 2006: advances in artificial intelligence, 19th Australian joint conference on artificial intelligence, hobart, Australia, december 4-8, 2006, proceedings (2006), pp. 1015-1021, 10.1007/11941439_114
https://doi.org/10.1007/11941439%5c_114

[9] J. Sun, A. Binder Comparison of deep learning architectures for h&e histopathology images 2017 IEEE conference on big data and analytics (ICBDA), IEEE (2017), pp. 43-48

[10] S.A. Adeshina, A.P. Adedigba, A.A. Adeniyi, A.M. Aibinu Breast cancer histopathology image classification with deep convolutional neural networks 2018 14th international conference on electronics computer and computation (ICECCO), IEEE (2018), pp. 206-212

[11] A.A. Samah, M.F.A. Fauzi, S. Mansor Classification of benign and malignant tumors in histopathology images 2017 IEEE international conference on signal and image processing applications, ICSIPA 2017, kuching, Malaysia, september 12-14, 2017 (2017), pp. 102-106, 10.1109/ICSIPA.2017.8120587
https://doi.org/10.1109/ICSIPA.2017.8120587

[12] M. Sokolova, N. Japkowicz, S. Szpakowicz Beyond accuracy, f-score and ROC: a family of discriminant measures for performance evaluation AI 2006: advances in artificial intelligence, 19th Australian joint conference on artificial intelligence, hobart, Australia, december 4-8, 2006, proceedings (2006), pp. 1015-1021, 10.1007/11941439_114
https://doi.org/10.1007/11941439%5c_114

[13] Al-Antari M. A., Al-Masni M. A., Choi M.-T., Han S.-M., Kim T.-S. A fully integrated computer-aided diagnosis system for digital x-ray mammograms via deep learning detection, segmentation, and classification. International Journal of Medical Informatics. 2018; 117:44–54. doi: 10.1016/j.ijmedinf.2018.06.003

[14] Zhang Yu-D., Satapathy S. C., Guttery D. S., Górriz J. M., Wang S. H. Improved breast cancer classification through combining graph convolutional network and convolutional neural network. Information Processing & Management.2021;58

[15] Kooi T., Litjens G., van Ginneken B., et al. large-scale deep learning for computer aided detection of mammographic lesions. Medical Image Analysis. 2017; 35:303–312

[16] Eroğlu Y., Yildirim M., Çinar A. Convolutional neural networks-based classification of breast ultrasonography images by hybrid method with respect to benign, malignant, and normal using mrmr. Computers in BiologyandMedicine.2021

[17] Zhang H., Han L., Chen Ke, Peng Y., Lin J. Diagnostic efficiency of the breast ultrasound computer-aided prediction model based on convolutional neural network in breast cancer. Journal of Digital Imaging. 2020;33(5):1218–1223.

[18] Braz Junior G., da Rocha S. V., de Almeida J. D. S., de Paiva A. C., Silva A. C., Gattass M. Breast cancer detection in mammography using spatial diversity, geostatistics, and concave geometry. Multimedia Tools and Applications. 2019;78

[19] Harefa J., Alexander A., Pratiwi M. Comparison classifier: support vector machine (svm) and k-nearest neighbor (k-nn) in digital mammogram images. Jurnal Informatika dan Sistem Informasi . 2017;2