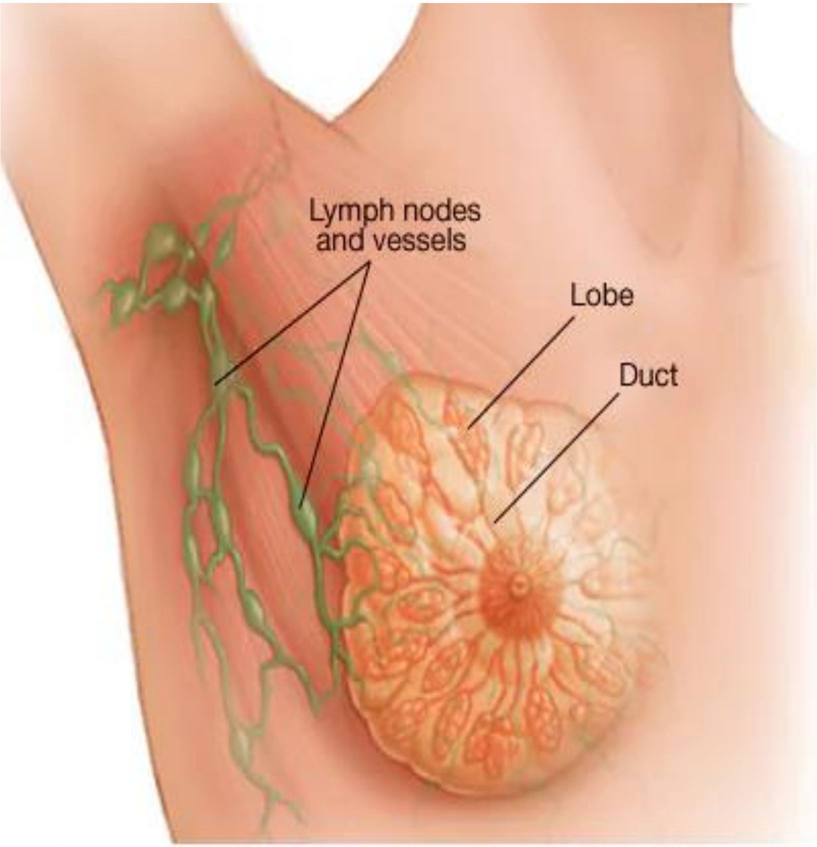
Breast Cancer Detection from Mammographic Images using Deep Learning

By:

Aswini Pusuluri

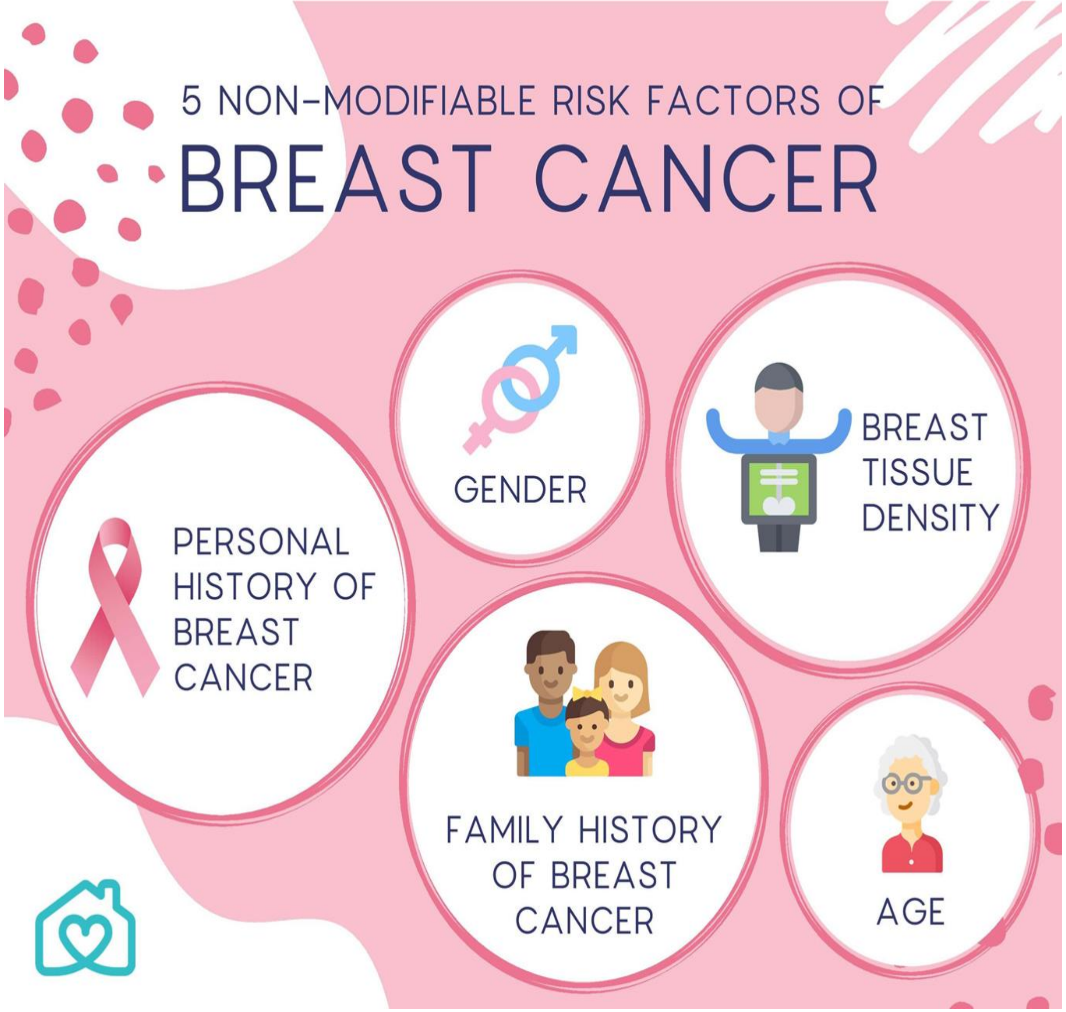
Project Background

- Breast cancer is a medical condition which originates in the breast, where the breast cells grow abnormally.
- Ducts, Lobules, and connective tissue are the three main parts that constitute the breast.
- Places of origin - Ducts or Lobules
- Diagnosis of breast cancer is primarily based on whether the tumors are malignant or benign.
- It is considered as positive if the tumors are malignant and negative if the tumors are benign in nature.



Risk Factors:

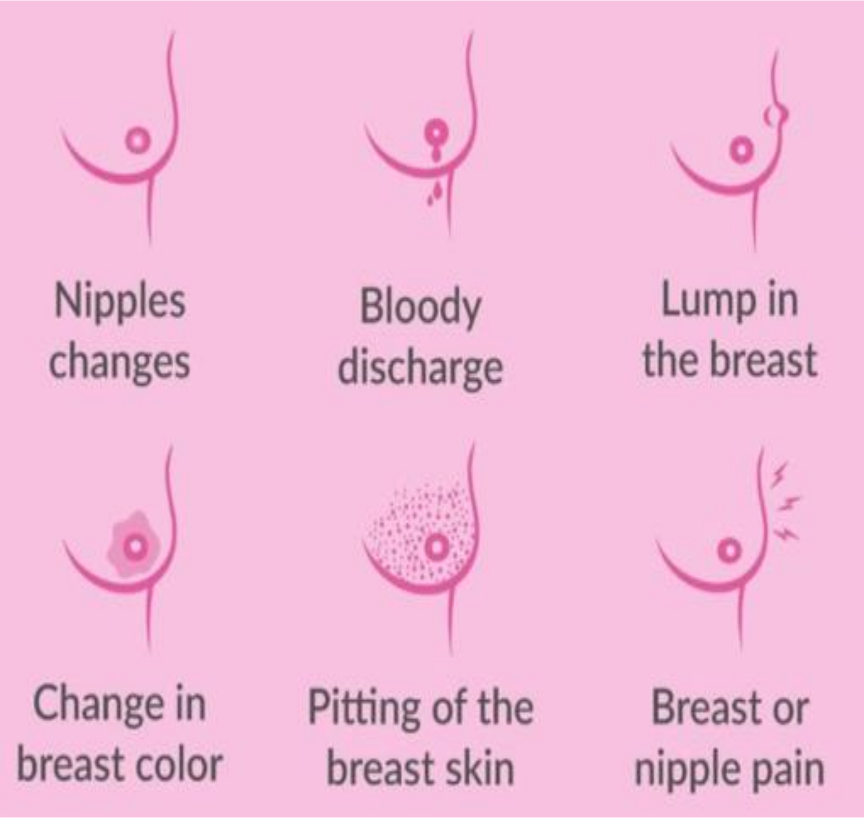
- Personal or family history
- Radiation therapy to chest or upper body
- Age
- Diets high in saturated fat



From https://www.researchgate.net/publication/346054458_Breast_cancer_PPT

Symptoms:

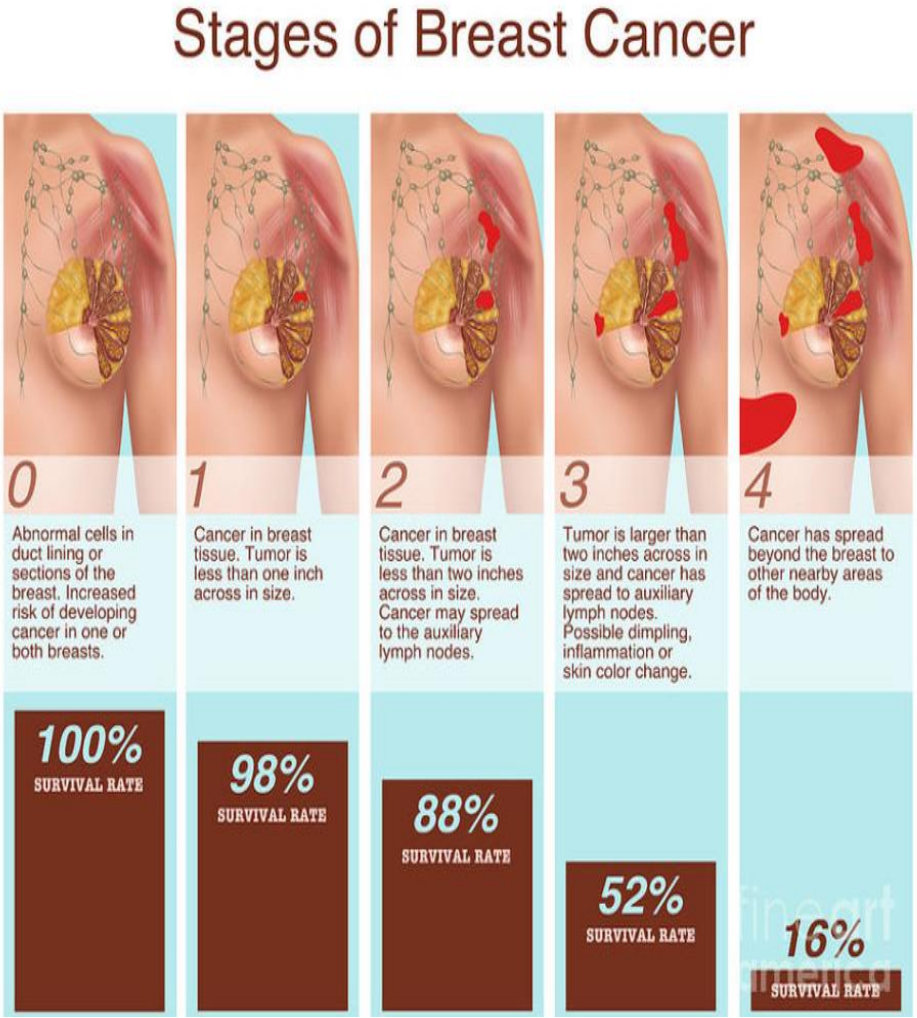
- Early breast cancer has little or no symptoms. It is not painful.
- Discharge, especially if only from one breast
- Sunken nipple
- Redness and changes in texture.
- Lumps on or around breast



From World Health Organization (WHO)
https://www.facebook.com/WHO/

Motivation:

- Breast cancer is one of the predominant medical conditions that has been claiming significant number of lives every year
- According to U.S. Breast Cancer Statistics, it is expected that of all the cancers that will be newly diagnosed in women 30% will be breast cancers.
- There were as many as 2.3 million women that were affected by breast cancer and of which 685,000 were dead, as it is usually too late when specialists identify the disease
- Early detection and diagnosis is very crucial to prevent the progressive damage caused by the disease



From https://www.researchgate.net/publication/346054458_Breast_cancer_PPT

Need for Project

Existing Detection Techniques:

- Mammography
- Breast Ultrasound scan
- Magnetic resonance imaging (MRI)
- Photoacoustic Imaging, Computed Tomography (CT) Scan
- Nuclear Magnetic Resonance Imaging, etc.

Challenges with Existing Techniques:

- False negatives rate for preliminary Mammogram screenings is as high as 30%.
- External factors of error such as distractions, fatigue and manual errors while examining Mammograms.
- Incorrect interpretation of Mammogram images by radiologists can lead to decisions that are harmful to the patients.

Aim

- To detect cases of breast cancer in mammograms using different Convolution Neural Network (CNN) models.
- Perform a comparative study on which model gives best results based on performance matrix which gives more information like Accuracy, Precision, Recall and F-Score.

Deliverables

- The final deliverable of this project involves reporting the proposed deep learning pipeline which helps improving the accuracy and performance of breast cancer detection.

Project Requirements

Project Requirements

Data

- Mammographic images which are classified as Normal, Malignant & Benign
- Pre-processing and transformation of data is required before feeding it to the model

Cloud Services

- Data Storage and Backup
- Identity and access management for secured management of data
- Integrated Development Environment with all machine learning, Deep Learning packages and minimum GPU memory of 16 GB.
- Deployment and Versioning.

Task Partition

Task 1
- Business & Data Understanding

Task 2
- Data Collection

Task 3
- Data Pre-processing

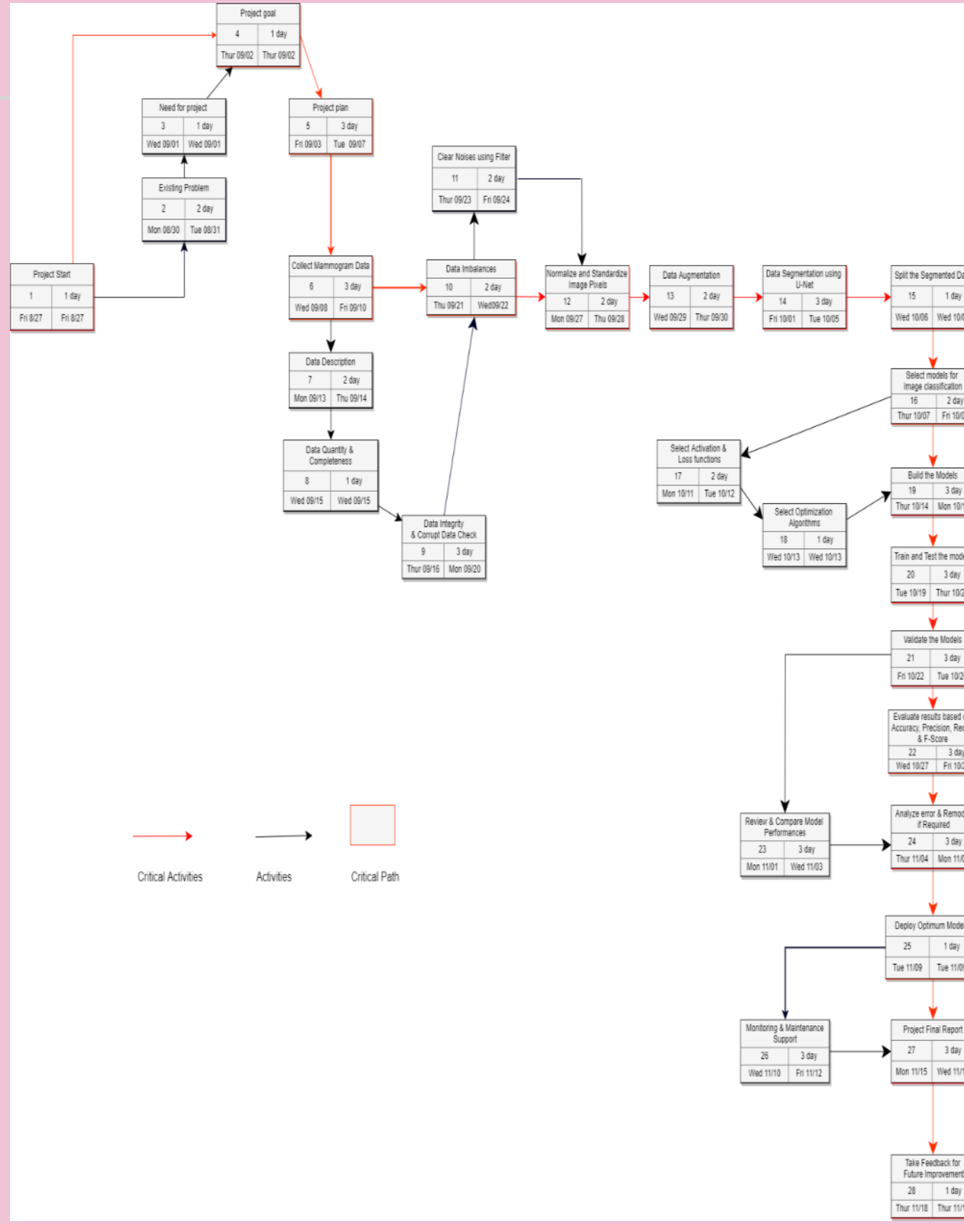
Task 4
- Data Segmentation

Task 5
- Models Creation

Task 6
- Model Validation and Evaluation

Task 7
- Deployment

Task Organization



Technology and Literature Survey

Breast Cancer Detection

Title	Datasets	Summary	Results
Breast Cancer Diagnosis by Different Machine Learning Methods Using Blood Analysis Data [3]	Breast Cancer Coimbra dataset taken from UCI ML Repository with features Age, BMI, Glucose ,Insulin, Leptin, Adiponectin etc.	Four different machine learning algorithms ANN, Extreme Learning Machine (ELM), SVM and KNN were used for classifying the target data as healthy and unhealthy. By using the hyper parameter optimization methodology, the hyper parameter values producing the least errors for all the four models were obtained. These values were used to calculate accuracy rates and training times.	Standard Extreme Learning Machine has the best accuracy rate of 83.87% and the shortest training period compared to other three models.
Machine Learning Classification Techniques for Breast Cancer Diagnosis [4]	Wisconsin Diagnostic Breast Cancer (WDBC) Dataset from the UCI Machine Learning repository. With features Radius texture, smoothness, compactness, concavity, perimeter, area, concave points, symmetry, and fractal dimension	A fusion strategy to identify breast tumor that involved utilizing Linear Discriminant Analysis (LDA) to trim the high feature dimensionality and apply the new trimmed feature dataset to different ML algorithms to classify benign and malignant tumors. The ML algorithms like SVM, ANN, and Naïve Bayes (NB) are used for classification.	Support Vector Machine-Linear Discriminant Analysis (SVM-LDA) and Artificial Neural Networks-Linear Discriminant Analysis (ANN-LDA) together had excellent results, a sensitivity of 98.4%, precision of 98.4%, and accuracy of 98.82 %.

Breast Cancer Detection

Title	Datasets	Summary	Results
Boosting Breast Cancer Detection Using Convolutional Neural Network [1]	Kaggle 162 H&E dataset was used. The data set consists of both benign and malignant images.	CNN method to improve the automated breast cancer detection based on the analysis of affected tissue regions in Whole-Slide Images (WSIs) with ductal carcinoma. Architecture was driven by 50 x 50-pixel RGB image patches of large datasets which are typically about 2,75,000 in number. Proposed CNN Model is comprised of 2 convolution layers each of which has 32 kernels and 64 kernels, respectively. The overfitting problem is prevented using dropout regularization.	The proposed model is effective and efficient with 87% accuracy.
Deep learning in mammography images segmentation and classification: Automated CNN approach [5]	Three datasets MIAS, DDSM and CBIS-DDSM, which has benign and malignant mammogram images.	Breast cancer image classification and segmentation based on Mediolateral Oblique (MLO) and Crania Caudal (CC) views to improve system performance. Breast cancer is detected and diagnosed using the CC view and MLO view. To classify the data into malignant data or benign data, a variety of practices like VGG16, DenseNet121, MobileNetV2, and InceptionV3 are used.	The proposed approach with InceptionV3 and an altered U-Net model yields the results with 98.870% accuracy, 98.980% sensitivity, 97.990% F1 score, 98.790% precision, and 1.21340s computing time.

Breast Cancer Detection

Title	Datasets	Summary	Results
YOLO Based Breast Masses Detection and Classification in Full-Field Digital Mammograms [2]	INbreast mammogram dataset	This paper proposed CAD system design based on YOLO. Three YOLO architectures V1, V2, V3 were used for detection of masses and their classification. If the anchors related to the dataset are built using k-means clustering and when used in YOLO-V3, it significantly improved accuracy of detection. To evaluate and compare the performance with YOLO Inception and ResNet are used.	Around 89% of INbreast masses of mammograms were classified as benign with 94.2% average precision and malignant with 84.6% average precision.

Technical Solutions

Model	Advantages	Disadvantages
Conventional CAD systems	Served as second pair of eyes for clinicians during analysis of abnormal features.	This approach required feature extraction on the basis of manually defined descriptors.
Machine Learning Models	This models discover hidden patterns in mammography data that radiologists couldn't see	They were unable to operate on raw data such as mammography images. Needs feature extraction
Deep CNN	No need for manual intervention far faster than classic machine learning approaches. Works directly on image data	High Computation cost. Requires huge volume of data to train the model

Project Resource Requirements

Resources required for this Project

Function	Resource Type	Resource	Cost Estimation	Justification
Cloud Service Management Tool	Software	Amazon CLI	0	Provides the functionality to communicate with different AWS resources and services using the command window in the local machine
Data Storage	Hardware	Amazon S3	$2.07 for 30 GB for three months	Scalable, Reliable and Secured data management.
Disaster Recovery	Hardware	Amazon S3 Replication	$2.52 for three months	Would ensure high availability in case of any contingencies
Data Archival	Hardware	Amazon S3 Glacier	$0.36 for three months	Data is stored for a longer life cycle without much need to retrieve for low cost.
Build, train, test and deploy	Software	Amazon SageMaker ml.p3.2xlarge instance	$153 for forty hours	It has integrated set of tools for process automation, error correction and it is also cost effective. Amazon SageMaker is flexible and scalable. The p3.2xlarge instance is complemented with one Tesla V100 Graphical Processing Unit (GPU), 16 GB GPU memory, 8vCPUs, 61 GB Memory and up to ten Gbps of bandwidth

Resources required for this Project

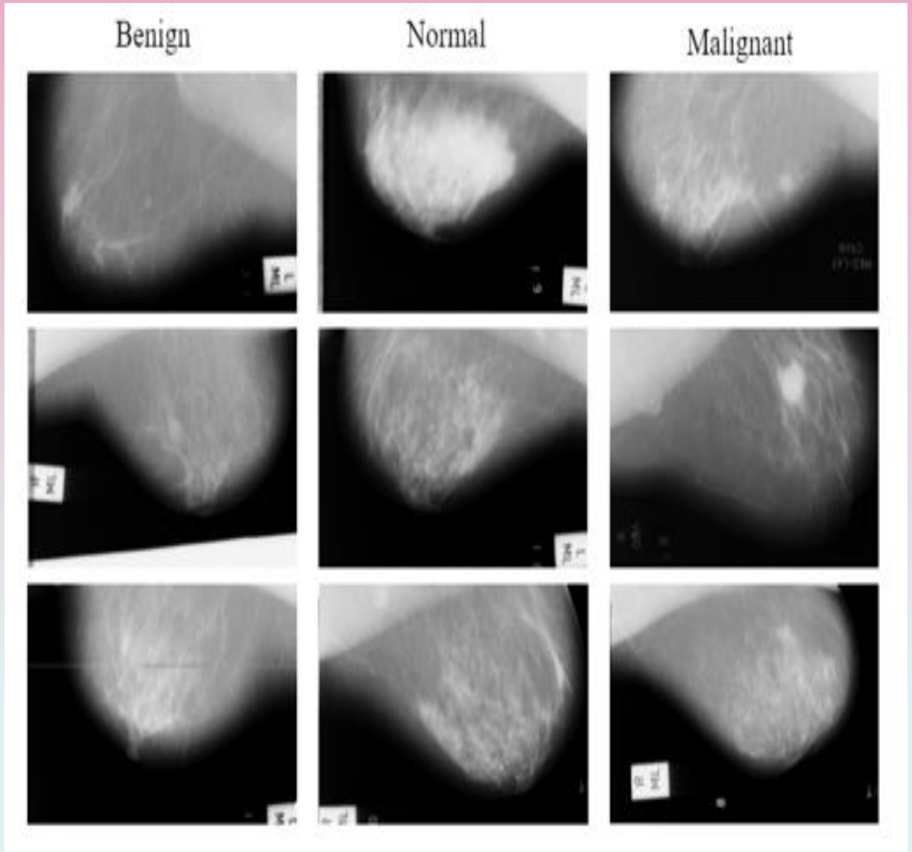
Function	Resource Type	Resource	Cost Estimation	Justification
ML Frameworks	Software	TensorFlow	0	TensorFlow and Keras are the machine learning frameworks that are utilized to build, train, test and evaluate the models.
Software Development	Software	Python 3.9	0	cv2, NumPy, json, matplotlib, seaborn, pandas are some of the noticeable libraries used for this project
Image Annotation	Software	VGG Image Annotator	0	Using the polygon shaped region select tool, identify, and draw the borders of the masses in the image manually
Diagram Tool	Software	Draw.io	$45 for three months	Used to create the work breakdown structure, PERT chart ,different diagrams and process flow charts.

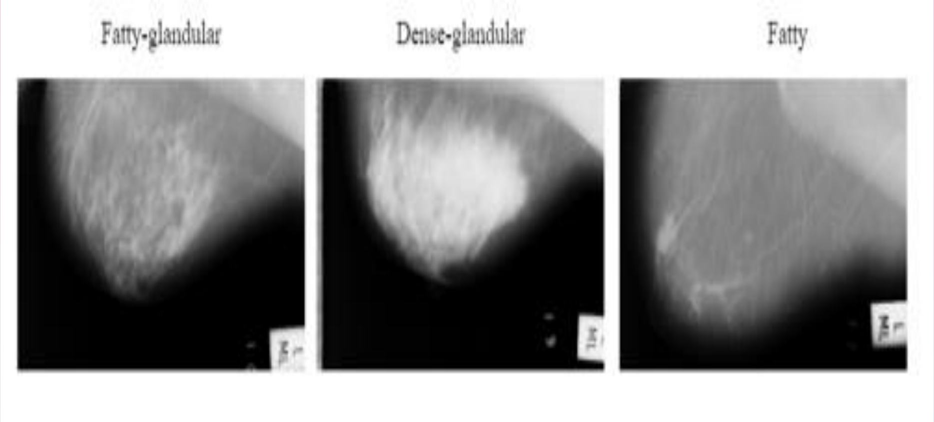
Data Preparation

Data Source

- Any data collection effort involving humans will have to undergo a series of regulations and compliances to avoid any unforeseen issues related to data misuse.
- Institutional Review Board (IRB) approval needed.
- Collaborative Institutional Training Initiative (CITI) Certificate.
- Data used in this project are mammographic images which are excerpted from Mini MIAS Database.
- Mammography Image Analysis Society (MIAS) is an UK based National Breast Screening Program.
- Total of 322 images in which 63 are benign, 208 are normal, and 51 are malignant that belong to both sides of breast sections of 161 patients.
- 1024 x 1024 pixels for each image

Samples of raw data





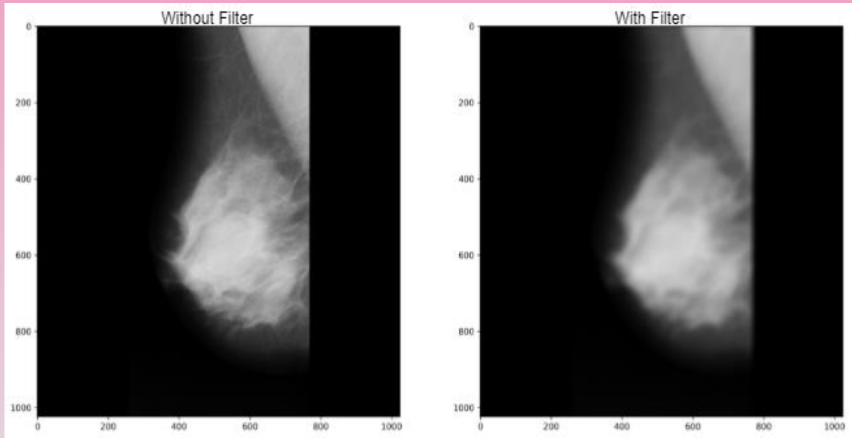
Metadata related to the mammographic images

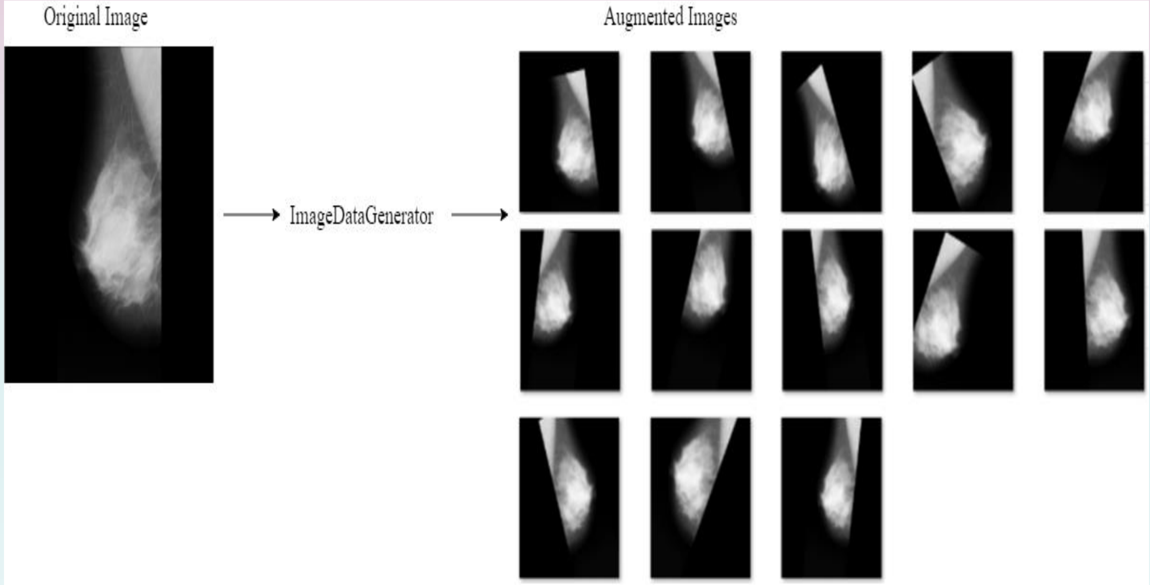
- Background tissue characteristics
- Types of abnormalities
- Severity of the tumor detected
- Coordinated location of micro-calcification
- Radius of circle surrounding the impacted region.

REFNUM	BG	CLASS	SEVERIT	X	Y	RADIUS
mdb001	G	CIRC	B	535	425	197
mdb002	G	CIRC	B	522	280	69
mdb003	D	NORM				
mdb004	D	NORM				
mdb005	F	CIRC	B	477	133	30
mdb005	F	CIRC	B	500	168	26
mdb006	F	NORM				
mdb007	G	NORM				
mdb008	G	NORM				
mdb009	F	NORM				
mdb010	F	CIRC	B	525	425	33
mdb011	F	NORM				
mdb012	F	CIRC	B	471	458	40
mdb013	G	MISC	B	667	365	31
mdb014	G	NORM				
mdb015	G	CIRC	B	595	864	68
mdb016	G	NORM				

Data Pre-processing

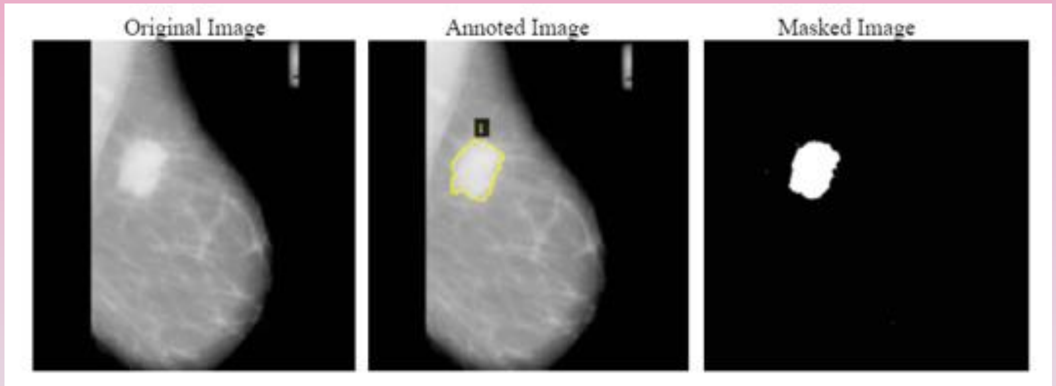
- Median filter is used to filter the different noises present in the images
- Data Augmentation to generate subsamples by varying degrees of 0, 90, 180 and 270 and mirroring
- Resized from 1024x1024 to 227x227 pixels

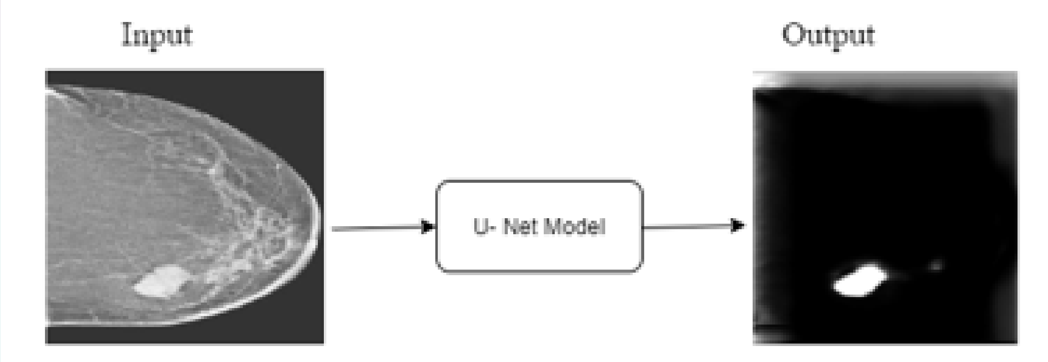




Data Transformation

- In the MIAS dataset ground truth data is available but not the annotated images
- VGG Image Annotator is used for Image annotations
- Data Segmentation using U-Net model
- Adam optimizer, Cross entropy loss function.
- Parameters used for implementing the U-Net model are 40 epochs with 300 steps per each epoch.
- Dice similarity coefficient (DSC) metric is used to determine the U-Net model efficiency
- The output of the U-Net model is the segmented map image





Data Split

- One-hot encoding technique is used to represent normal, benign, and malignant.
- Total of 3252 records in the dataset after transformation process.
- The resulting dataset from transformation process is immediately split into 80% and 20% ratio as training and testing datasets
- The training dataset is further divided into train and validation datasets using k-fold cross validation.

Source	Raw	Augmented	Train	Validation	Test
Mini-MIAS Database	322	3252	2341	260(10-Fold)	651

Samples of Train, Test and Validation Datasets

```
X_train.head(10)
```

	Ref_Num	ab_class	bg	Benign	Malignant	Normal	sample_path
258	mdb022_180_rotated	NORM	G	0	0	1	../samples\mdb022_180_rotated.png
232	mdb020_90_mirr_lr	NORM	G	0	0	1	../samples\mdb020_90_mirr_lr.png
33	mdb003_270_rotated	NORM	D	0	0	1	../samples\mdb003_270_rotated.png
157	mdb014_0_mirr_lr	NORM	G	0	0	1	../samples\mdb014_0_mirr_lr.png
148	mdb013_90_mirr_lr	MISC	G	1	0	0	../samples\mdb013_90_mirr_lr.png
93	mdb008_270_rotated	NORM	G	0	0	1	../samples\mdb008_270_rotated.png
37	mdb004_0_mirr_lr	NORM	D	0	0	1	../samples\mdb004_0_mirr_lr.png
139	mdb012_180_mirr_lr	CIRC	F	1	0	0	../samples\mdb012_180_mirr_lr.png
16	mdb002_90_mirr_lr	CIRC	G	1	0	0	../samples\mdb002_90_mirr_lr.png
217	mdb019_0_mirr_lr	CIRC	G	1	0	0	../samples\mdb019_0_mirr_lr.png

Ref_Num	ab_class	bg	Benign	Malignant	Normal	sample_path
mdb008_90_rotated	NORM	G	0	0	1	../samples\mdb008_90_rotated.png
mdb018_270_mirr_lr	NORM	G	0	0	1	../samples\mdb018_270_mirr_lr.png
mdb011_0_mirr_lr	NORM	F	0	0	1	../samples\mdb011_0_mirr_lr.png
mdb023_270_mirr_tp	CIRC	G	0	1	0	../samples\mdb023_270_mirr_tp.png
mdb002_180_mirr_tp	CIRC	G	1	0	0	../samples\mdb002_180_mirr_tp.png
mdb016_180_mirr_tp	NORM	G	0	0	1	../samples\mdb016_180_mirr_tp.png
mdb006_270_mirr_tp	NORM	F	0	0	1	../samples\mdb006_270_mirr_tp.png
mdb009_270_mirr_lr	NORM	F	0	0	1	../samples\mdb009_270_mirr_lr.png
mdb023_180_rotated	CIRC	G	0	1	0	../samples\mdb023_180_rotated.png
mdb009_180_rotated	NORM	F	0	0	1	../samples\mdb009_180_rotated.png

Ref_Num	bg	sample_path
mdb003_180_rotated	D	../samples\mdb003_180_rotated.png
mdb011_90_mirr_lr	F	../samples\mdb011_90_mirr_lr.png
mdb017_90_mirr_lr	G	../samples\mdb017_90_mirr_lr.png
mdb011_180_mirr_lr	F	../samples\mdb011_180_mirr_lr.png
mdb019_0_rotated	G	../samples\mdb019_0_rotated.png
mdb020_180_rotated	G	../samples\mdb020_180_rotated.png
mdb013_0_rotated	G	../samples\mdb013_0_rotated.png
mdb017_270_mirr_tp	G	../samples\mdb017_270_mirr_tp.png
mdb022_0_mirr_tp	G	../samples\mdb022_0_mirr_tp.png
mdb012_270_mirr_tp	F	../samples\mdb012_270_mirr_tp.png
mdb007_180_mirr_lr	G	../samples\mdb007_180_mirr_lr.png

Models

Model Selection

- Traditional CAD systems still require the intervention of medical professional or a clinician while identification of image features of impacted regions of the breast.
- Deep learning CAD systems like Convolutional Neural Network (CNN) proved to overcome the challenges by learning the features of the mammography images through different layers.
- This project uses CNN techniques with different training approaches to identify the best detection model.
- The different models adopted in this project are modified CNN trained from scratch, AlexNet and ResNet50.

CNN Training from Scratch

- The proposed CNN architecture has six convolutional layers, four average pooling layers and three fully connected layers. Every convolution layer is activated by the ReLU activation function.

- There is a SoftMax activation function that is applied on the final fully connected layer which has three neurons that correspond to three classifications.

- Categorical Cross Entropy is used to calculate the loss and Adam is used as an optimizer.

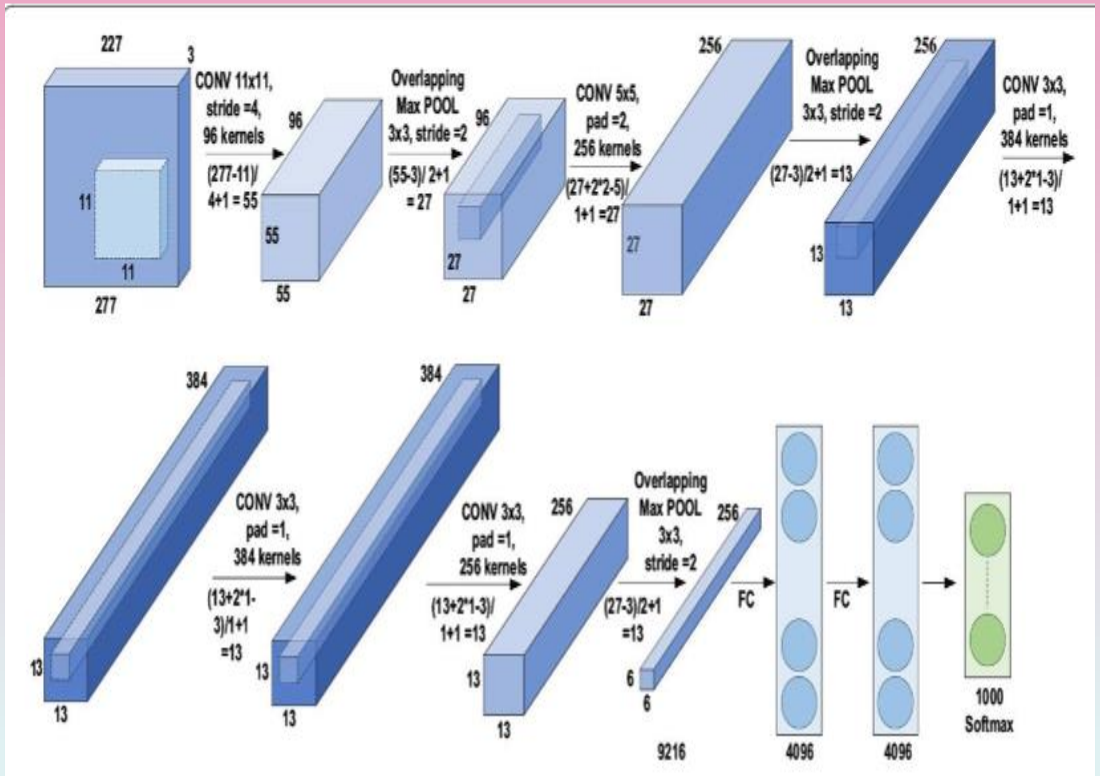
From https://ieeexplore.ieee.org/document/8346384

CNN Training from Scratch

Model	Strengths	Limitations	Justification
CNN	• Automatic detection of features • No need for manual intervention • Availability of large number of trainable parameters in various layers • Extraction of distinguishing features at different abstraction levels	• Requires huge volume of data to train the model • More computing power is needed • No definitive approach in determining layers • Not efficient when a smaller number of layers are involved	• The self-reliant learning patterns, efficient processing abilities and adaptability with different systems have made CNN outperform the existing machine learning models in the detection of breast cancer. • Building a model from scratch has the advantage of allowing you to find simpler model designs that can do the task better than a sophisticated model.

AlexNet

- AlexNet is comprised of different layers that involve five layers of convolution, three layers of max pooling, two layers of fully connected and a SoftMax layer.
- ReLu is used as an activation function in AlexNet.
- AlexNet uses momentum optimizer technique as it is focused on gradient with momentum optimizer.
- SoftMax layer with 1000 output channels which is replaced with three output channels for this research to categorize whether the tumor is normal, benign, or malignant.



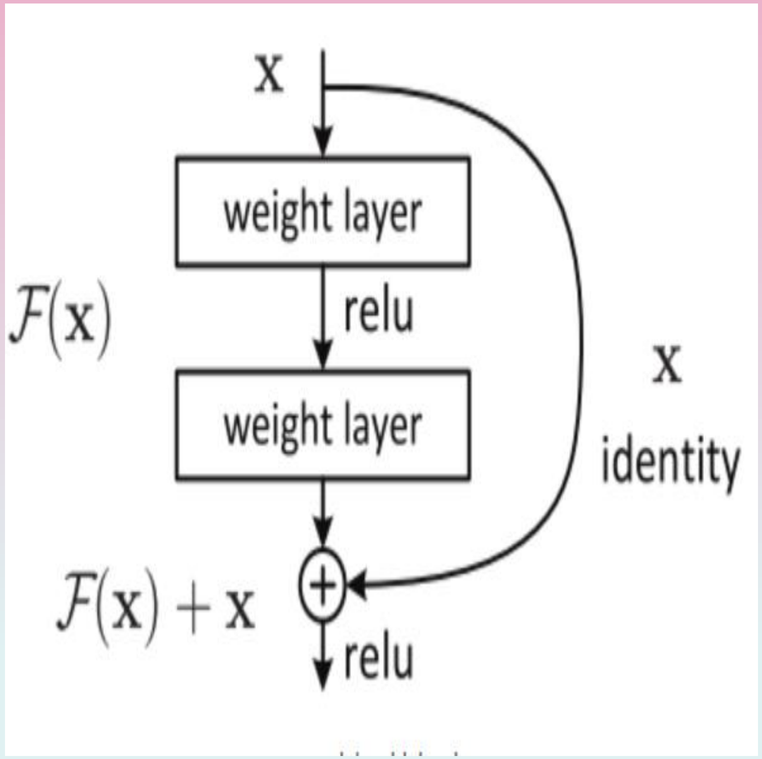
From https://www.researchgate.net/publication/350527503_Review_of_deep_learning_concepts_CNN_architectures_challenges_applications_future_directions

AlexNet

Model	Strengths	Limitations	Justification
AlexNet	• Pre-trained model based on transfer learning • Use of two GPUs result in faster computation • Use of Local Response Normalization to resolve unbound output problem • Use of Dropout regularization to address overfitting problem	• Limited depth when compared to other models • Use of convolution filters that are larger in size • Inferior performance when compared to complex models like GoogleNet, ResNet etc.	• AlexNet was developed to achieve better accuracy at a faster rate as the training speed is greatly improved in this model when compared to CNN model trained from scratch. • Transfer learning frequently produces better results than models that have been trained from the scratch.

ResNet50

- ResNet50 is a residual neural network with individual residual blocks in it.
- There are 49 convolution layers with the combination of different kernels sizes with varying numbers and one fully connected layer.
- ResNet50 uses Adam optimizer and Categorical cross entropy.
- ResNet50 uses both average and max pooling.
- SoftMax layer which categorize whether the tumor is normal, benign, or malignant.



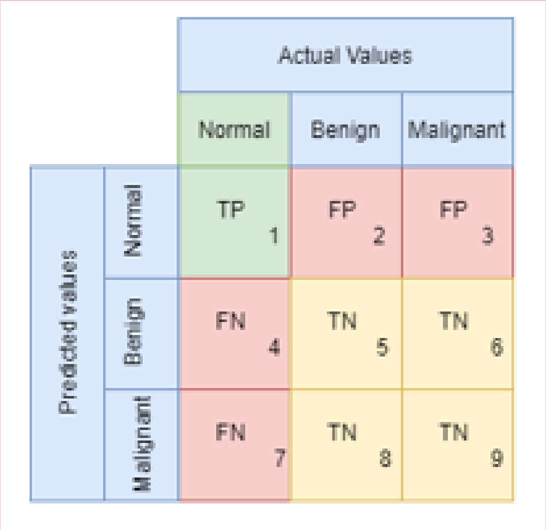
From https://towardsdatascience.com/an-overview-of-resnet-and-its-variants-5281e2f56035

ResNet50

Model	Strengths	Limitations	Justification
ResNet50	• More number of layers results in effective understanding of features • Degradation problem is resolved with skip connection mechanism • Saves time and results in improved performance by skipping already learned features • Low validation loss due to use of residual networks	• Complex architecture • Batch normalization requires more computational resources • Without transfer learning this model could result in larger training times	• Due to the increase in number of layers in deep CNN, the model fails to learn new features at certain point which is called degradation problem. This problem is addressed with the use of residual networks in ResNet50. • Skip connections are introduced to skip the already learned features in the intermittent layers. This approach will significantly improve the performance and reduce the training time.

Evaluation

- Classification report from confusion matrix for Evaluation
- AUC-ROC curve will be used to compare all proposed model's performance.
- Validation loss
- Validation accuracy




```
Classes: {'A': 0, 'B': 1, 'M': 2}

              precision    recall  f1-score   support

           0       0.75      0.60      0.67        55
           1       0.79      0.84      0.81       132
           2       0.84      0.85      0.84       133

    accuracy                           0.80       320
   macro avg       0.79      0.76      0.77       320
weighted avg       0.80      0.80      0.80       320
```

References for Literature Review

1. Alanazi, S. A., Kamruzzaman, M. M., Islam Sarker, M. N., Alruwaili, M., Alhwaiti, Y., Alshammari, N., & Siddiqi, M. H. (2021). Boosting breast cancer detection using convolutional neural network. *Journal of Healthcare Engineering*, *2021*, 1–11. https://doi.org/10.1155/2021/5528622

2. Aly, G. H., Marey, M., El-Sayed, S. A., & Tolba, M. F. (2021). YOLO Based Breast Masses Detection and Classification in Full-Field Digital Mammograms. *Computer Methods and Programs in Biomedicine*, *200*, 105823. https://doi.org/10.1016/j.cmpb.2020.105823

3. Aslan, M. F., Celik, Y., Sabanci, K., & Durdu, A. (2018). Breast cancer diagnosis by different machine learning methods using blood analysis data. *International Journal of Intelligent Systems and Applications in Engineering*, *6*(4), 289–293. https://doi.org/10.18201/ijisae.2018648455.

4. Omondiagbe, D. A., Veeramani, S., & Sidhu, A. S. (2019). Machine learning classification techniques for breast cancer diagnosis. *IOP Conference Series: Materials Science and Engineering*, *495*, 012033. https://doi.org/10.1088/1757-899x/495/1/012033

5. Shen, L., Margolies, L. R., Rothstein, J. H., Fluder, E., McBride, R., & Sieh, W. (2019). Deep learning to improve breast cancer detection on screening mammography. *Scientific Reports*, *9*(1). https://doi.org/10.1038/s41598-019-48995-4

Thank You