BREAST CANCER DETECTION

(UDP)

A PROJECT REPORT

Submitted by

DOSHI JAINESH. K	170090107005
SHAH DEVANSH. V	170090107049
SHAH DHRUVIL. N	170090107050
SHAH KENEEL .C	170090107051

Guided by

Prof. Neelam Surti

Associate Professor

In partial fulfillment for the award of the degree of

BACHELOR OF ENGINEERING

In

Department of Computer Engineering

C. K. Pithawala College of Engineering and Technology, Surat

Gujarat Technological University, Ahmadabad

Academic year
(2020-2021)

Index

Acknowledgements

Project work upgrades the vision of a student; it helps in confronting the difficulties, manage senior individuals, and furthermore enhances the communication skills and in addition knows the market request of different things and what is inadequate in the market. Our project would not have been at this phase without the direction of some special people. One by one we might want recognize them.

Warmly, we might want thank to our consultant PROF. NEELAM A. SURTI for the continuous support for guiding us on this Project task and for her understanding, inspiration, and excitement. She was there to help us at each progression of this project. She helped us widen our knowledge in regards to the subject and we apart from our project work came to think about numerous other interesting things relating to our field. She was always with an answer at whatever point we were stuck anyplace. Genuinely Ma`am we couldn't have imagined having a superior guide and mentor for our project. We heartily thank you Ma`am for your direction and keeping faith in us.

We are additionally appreciative to all the faculty members and additionally non-teaching staff of COMPUTER ENGINEERING DEPARTMENT, CKPCET who have helped us directly or indirectly in our project work. We might want to thank our HOD, Dr. Ami T. Choksi for giving us the opportunity to take this project. Finally, we might want to say thanks to GOD almighty who is always with us, our PARENTS for their all-time support and believing in us and to our FRIENDS for their assistance and guidance to finish the project.

Certificate

C. K. Pithawala College of Engineering and Technology, Surat

Department of Computer Engineering

Year – 2020

Date: 29/10/2020

This is to certify that the project entitled "**BREAST CANCER DETECTION**" has been carried out by following students under my guidance in partial fulfillment of the degree of Bachelor of Engineering in Department of Computer Engineering (7^{th} Semester) of Gujarat Technological University, Ahmadabad during the academic year 2020-21. The work done by them is found satisfactory.

DOSHI JAINESH. K	170090107005
SHAH DEVANSH. V	170090107049
SHAH DHRUVIL. N	170090107050
SHAH KENEEL .C	170090107051

Prof. Neelam A.Surti
Guide

Jury

Dr. Ami T. Choksi
Department In-Charge

Abstract

The subject disclosure presents systems and computer-implemented methods for assessing the risk of cancer recurrence in a patient based on the basis of IDC: invasive ductal carcinoma vs non-IDC: invasive ductal carcinoma images, on Mammography images like Benign vs Normal vs Malignant and on Mammography images Fatty vs Fatty-Glandular vs Dense-Glanular together and on Numeric based data for benign vs malignant. These classifications can be done on the basis of the method selection for breast cancer detection. Our aim is to classify cancerous cells from the IDC images, Mammography images, and numeric data so that different types of models can be trained which can provide the different types of facilities in breast cancer detection. Within the opening move, we analyzed the photographs and appearance at the distribution of the pixel intensities as well as numeric data. Then, the pictures were normalized, and we attempted some very well-developed transfer-learning Convolutional Neural Networks like VGG16, VGG19, InceptionV3, ResNet50, and InceptionResNetV2 on IDC and Mammography based images and Linear, Logistics Regression, SVM-Linear, SVM-RBF, Decision-Tree Classifier, Random-Forest Classifier and GaussianNB for Numeric data. We had then validated and compared each of those models for the finalization of the model for respective methods. These all-neural networks were implemented as a python class and therefore the complete TensorFlow session is often saved to or restored from a file. We then also implement tensor summaries, which were used for the visualizations with TensorBoard. The output layer of the IDC model gave the result in form of IDC-Positive vs IDC-Negative, whereas the output layer of the Mammography model gave the result in form of Benign vs Normal vs Malignant and Fatty vs Fatty-Glandular vs Dense-Glanular together and Benign vs Malignant for Numeric model. So as to stop over-fitting the training data we will generate new images by rotations, translations, and zoom or retrain the whole model for both numeric and image-based datasets.

List of Tables

List of Figures

1. Introduction

1.1.Problem Definition

Breast cancer is one of the foremost common cancers among women. These cancers can be classified into different forms like benign, malignant, normal, Fatty, Fatty-Gladular and Dense-Glandular to determine the more logical and understandable cancer detection. The etiological reason behind the cancer is because of the generation of superoxides or free radicles. X-Ray, MRI, Autopsy, Biopsy tests are accustomed to detect diseases nowadays. It takes a protracted time for diagnosis. In our study, data processing techniques of machine learning are very advantageous to detect disease. During this experiment, we use different transfer learning methods to provide cancer detection models for tissue and mammography-based images and also for numeric based data. All different type of datasets is available on the **Kaggle**.

1.2.Idea Background: -

- Breast cancer is a significant and crucial reason behind death in women, and therefore, the second leading explanation for cancer deaths worldwide.
- Primary prevention within the early stages of the disease becomes complex because the causes remain almost unknown.
- However, some typical signatures of this disease, like masses and microcalcifications appearing on autopsy, biopsy or on mammography image, are often accustomed to improve early diagnostic techniques, which is critical for women's quality of life.
- Numeric data of the cancer and its location in the breast can also give many traces which can help to classify the cancer in form of benign and malignant.

1.3.Reasons for selecting the Definition: -

- Breast cancer is the most common form of cancer in women. Accurately identifying breast cancer and categorizing its subtypes is very crucial clinical task, and thus automated methods like neural networks and machine learning can be used to save time and reduce errors.

1.4.Project Objectives: -

- To manage and arrange every dataset in such a way that dataset can be used in proper proportion for the model training and the for-model validations.
- To Remove unwanted noise, borders, non-linearity from images using image processing
- There are two primary methods to coach this machine learning model on which the task is often performed.
    - Method 1: Deep Neural Network (NN)
    - Method 2: Convolution Neural Network (CNN)
- Our project is going to be implemented on basis of Convolution Neural Network (CNN) for tissue and mammography-based images so that important data can be extracted and those data can be feed into the Deep Neural Network (DNN) to generate the final answer on cancer detection on the given respective images.
- Our project also has direct feed of numeric data to Neural Network so for numeric datasets we will directly use the Deep Neural Network (DNN).
- Validate the accuracy, sensitivity, specificity, and efficiency of our model using statistical methods so that the model can be finalized for cancer detection.

1.5.Breast Cancer incidence by age: -

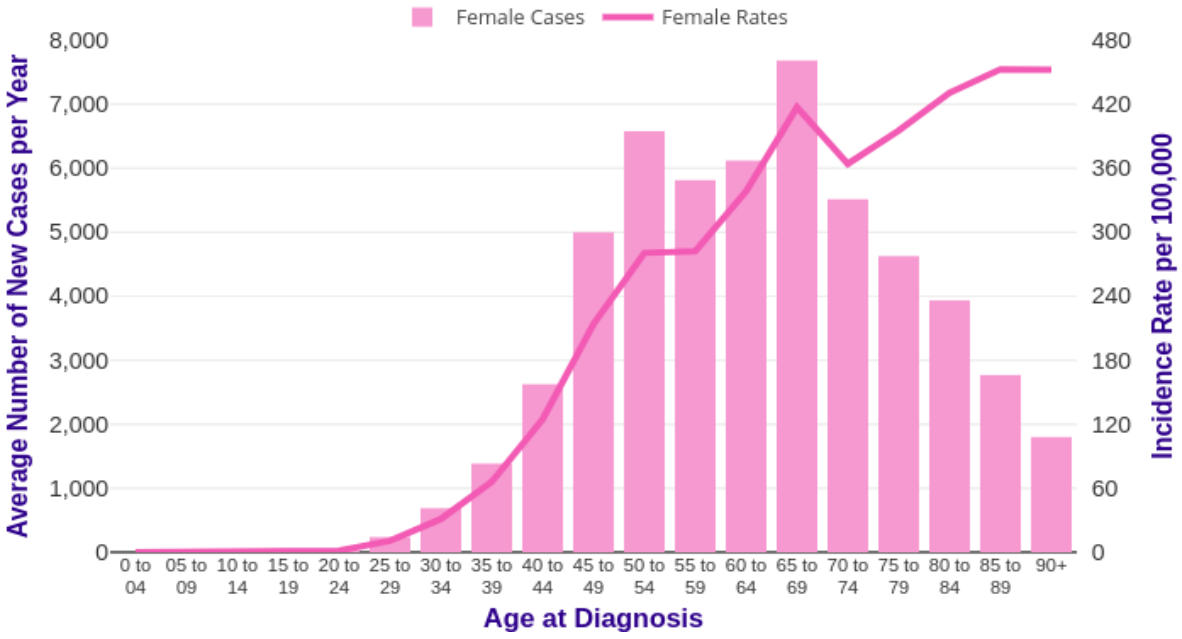

Figure 1: Breast Cancer incidence by age

- Breast cancer incidence is strongly associated with age, with the largest incidence rates being in older people. within the UK in 2015-2017, on average every year around 1 / 4 of the latest cases (24%) were in people aged 75 and over.[1-4]
- Age-specific incidence rates rise steadily from age 25-29, more steeply from age 35-39 in females and from age 60-64 in males. the largest rates are within the 85 to 89 age brackets for females and males.
- Incidence rates are significantly higher in females than males during several (mainly older) age groups. The gap is widest at age 45 to 49 when the age-specific incidence rate is 470 times higher in females than males.

Breast Cancer (C50): 2015-2017

Number of Average New Cases Every Year with Age-Oriented are Shown Below: Incidence: - Rates as per 1,00,000 Females, in United Kingdom.

Age Range	Female Cases	Female Rates
0 to 04	0	0.0
05 to 09	0	0.0
10 to 14	0	0.0
15 to 19	1	0.1
20 to 24	31	1.5
25 to 29	237	10.6
30 to 34	688	31.1
35 to 39	1,385	65.9
40 to 44	2,625	124.6
45 to 49	4,999	213.9
50 to 54	6,579	280.6
55 to 59	5,813	282.2
60 to 64	6,119	338.3
65 to 69	7,683	417.3
70 to 74	5,515	364.1
75 to 79	4,627	395.4
80 to 84	3,935	430.6
85 to 89	2,769	452.7
90+	1,800	452.5
All Ages	54,805	168.4

Table 1: Breast Cancer incidence by age

1.6.Summarized of studied literature: -

Sr. no.	Title	Author	Datasets	Results
1	A Novel Approach for Breast Cancer Detection Using Data Mining Techniques	Vikas Chaurasia, Saurabh Pal	Breast Cancer data available from the Wisconsin dataset from UCI machine learning	To compare three classification techniques in Weka software and comparison results show that Sequential Minimal Optimization (SMO) has higher prediction accuracy i.e., 96.2% than IBK and BF Tree methods.
2	Breast cancer detection by leveraging Machine Learning	Anji Reddy Vaka, Badal Soni, Sudheer Reddy K.	Well-annotated and large-scale Dataset is taken from the M. G Cancer Hospital & Research Institute, Visakhapatnam, India	The authors introduced the new method of DNNS for detecting Breast Cancer. Contrary to other methods, based on a deep neural network. To meet the better outcomes, efficiency, , accuracy and quality of images, a normalization methods has been used. Experiments proved that the newly proposed DNNS is relatively better than the existing standard methods.
3	Predictive Machine Learning Techniques for Breast Cancer Detection	S.Kharya, D. Dubey, and S. Soni	Multiple datasets were used as it is literature related to the review on multiple research papers which had researched on different methods in order to develop the mode which can detect the breast cancer	Review paper show the in-detail comparison of algorithms like decision tree, naïve bayes, neural networks and support vector machine and apart from this paper also show the result that how much hard and easy these algorithms are to implement and how much hard to understand them.
4	A Comparative Analysis of Nonlinear	Ali Al Bataineh	Wisconsin Breast Cancer	primary objective is to evaluate the performance in

	Machine Learning Algorithms for Breast Cancer Detection		Diagnostic (WBCD) dataset	classifying data with respect to efficiency and effectiveness of each nonlinear machine learning algorithms like Multilayer Perceptron (MLP), K-Nearest Neighbors (KNN), Classification and Regression Trees (CART), Gaussian Nave Bayes (NB) and Support Vector Machines (SVM) in terms of classification test accuracy, precision, and recall.
5	Analysis of Breast Cancer Detection Using Different Machine Learning Techniques	Siham A. Mohammed, Sadeq Darrab, Salah A. Noaman, Gunter Saake	Wisconsin Breast Cancer (WBC) and Breast Cancer dataset	In this paper, authors propose an approach that improves the accuracy and enhances the performance of three different classifiers: Decision Tree (J48), Naïve Bayes (NB), and Sequential Minimal Optimization (SMO)

Table 2: Summarized of Studied Literature

2. Design

2.1. Convolutional Neural Network System architecture: -

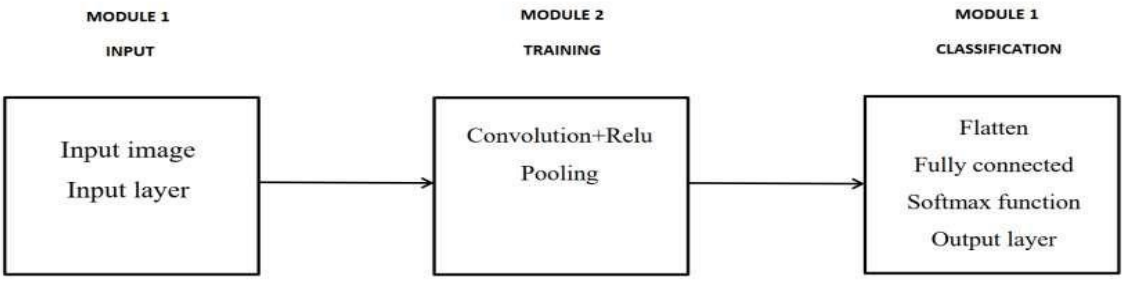

Figure 2: Overview of the Convolutional Neural Network System architecture

First image input: (Tissue based / Mammography based) images were passed through the input layer once image is inputted as shown in **figure 2.**

Second image training: Once the image is inputted in the input layer now the convolution layers will be applied. To retain the non-linearity of the image the "RELU" activation was applied with Convolutional layers so that important features can be extracted out. Once the process of Convolution + Relu gets over the pooling will be applied which can extract out the more important features from the images. This process is carried out by 3*3 filters in convolution and 2*2 filter in pooling. Thus, this process was kept on repetition till the given layer in convolution gets over as shown in **Figure 2**.

Third image main features will go for (NN): Finally, when the model finish all the layers of the convolution, all the important features of the images were then first flattened then they went through the deep neural network. After that process we finally provided the softmax function due to which we achieve the desired output as shown in **Figure 2**.

2.1.1. Convolutional Neural Network Layers: -

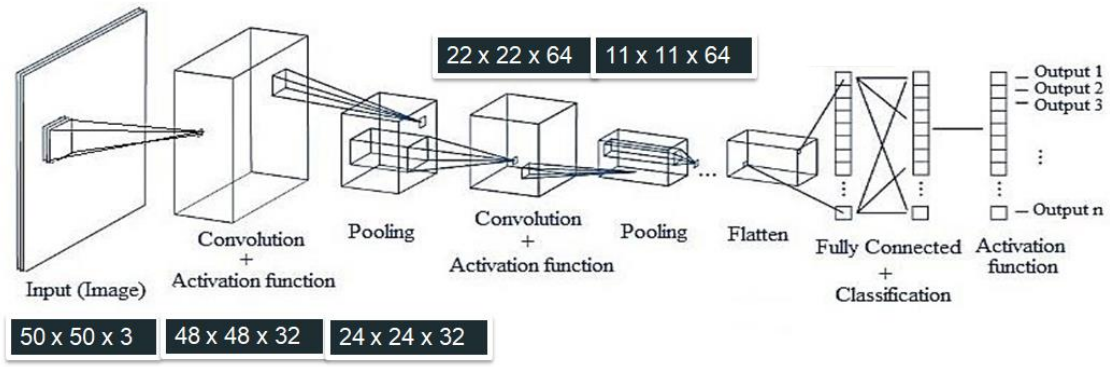

Figure 3: In Depth System Architecture of the CNN with it comprehensive layers

As we can see in the above **Figure 3** that for Example, first, we have to input the image of pixels which are provided to us in datasets of (IDC – Invasive Ductal carcinoma) or in the dataset of (Mammography). Let's take an example to understand in a better way, that 50*50*3 pixels of the image is available with us then it will pass through the first convolution layer where the Relu activation function will be applied to retain the non-linearity of the images. After that, the output of convolution can extract the features of images. Due to this less important feature of the image will be removed and due to that the image is now converted to the 48*48*32 pixels. After that, we will apply the pooling on 48*48*32 pixels so that we can reduce the number of parameters when the images are too large or block the input data and simply pass on maximum value or it reduces the size of the input and requires no added parameters. This particular process will be repeated several times and finally, we will able to achieve the main featured image with 11*11*64 pixels. Lastly, all main feature of the images will be gathered in form of parameters and feed to the deep neural network where different classification algorithms and the "SOFTMAX" activation function will be applied on it and then on the basis of that we can see the binary classified output which has the ability to predict patient has cancer or not.

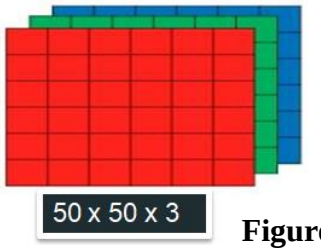

Figure 4

- An image of 50 x 50 x 3 array of matrix of RGB (3 refers to RGB values)
- Computers sees an input image as array of pixels and it depends on the image resolution.

- Based on the image resolution, it will see h x w x d(h = Height, w = Width, d = Dimension)

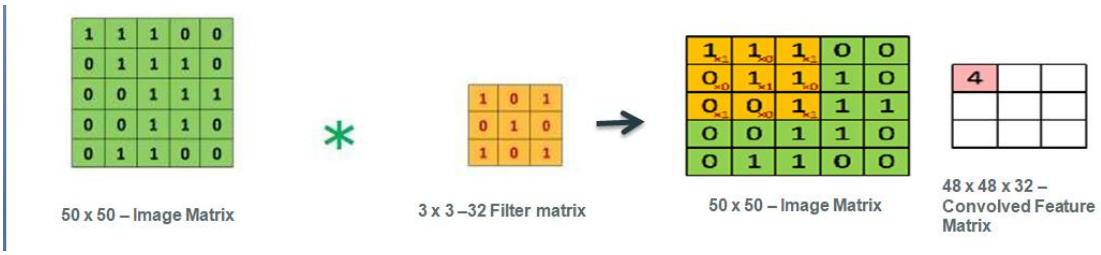

Figure 5: Feature Extraction Process from image using CNN

CONVOLUTIONAL LAYER

- Convolution **is that the** first layer to extract features from an input image.
- Convolution preserves **the connection** between pixels by learning image features using small squares of input data.
- It **could be a mathematical process** that takes two inputs like as image matrix and a filter or kernal.
- The convolution of 5 x 5 image matrix multiplies with 3 x 3 filter matrix which is termed "Feature Map" .

PADDING

- Sometimes filter doesn't perfectly fit the input image. We have two options:
  - Pad the image with zeros (zero-padding) so it fits
  - Pad the image with a neighbor(neighbor-padding) so that it fits

NON-LINEARITY (RELU)

- ReLU stands for the Rectified Linear Unit operation. The output is $f(x) = \max(0, x)$.
- Why ReLU is important: ReLU's purpose is to introduce non- linearity in our ConvNet.

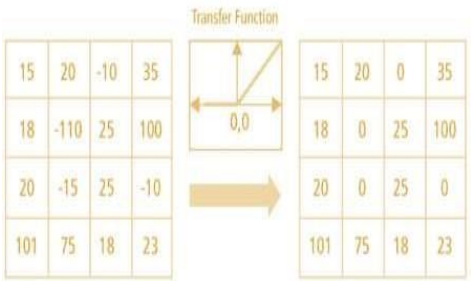

Figure 6: RELU Activation Function Process

Since the real-world data would want our ConvNet to be told would be non-negative linear values.

POOLING LAYER

- The pooling layers section would cut back the number of parameters when the photographs are overlarge.
- The pooling layer consider a block of the input file and easily die the most value
- Hence it reduces the dimensions of the input and requires no added parameters

FLATTENING

- After finishing the previous two steps, we're alleged to have a pooled feature map by now. because the name of this step implies, we are actually will went through flattening our pooled feature map into a column like within the image below.

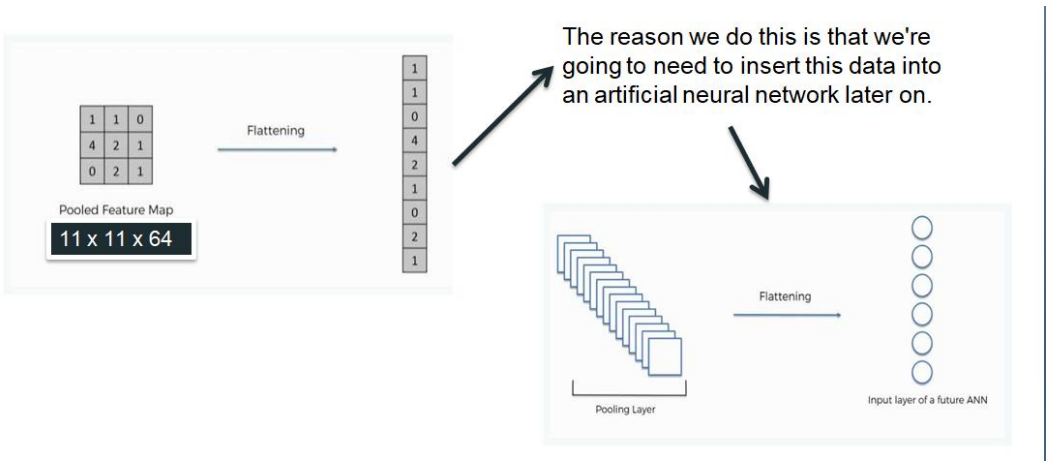

Figure 7: Last Process in CNN – Flattening all Parameters

FULLY CONNECTED LAYER

- The layer we call because the FC layer, we flattened our matrix into a vector and feed it into a completely connected layer sort of a neural network.

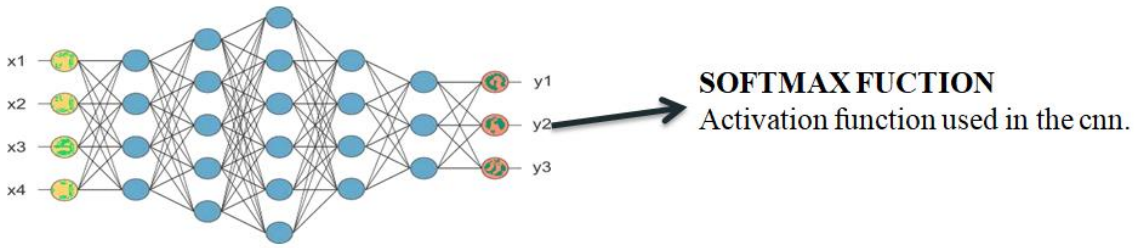

Figure 8: Feeding all Parameters into Deep Neural Networks

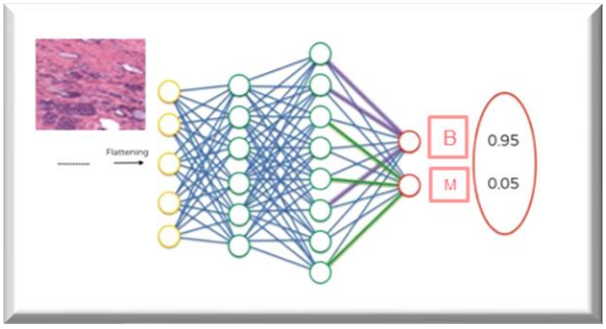

Figure 9: Tissue based image Fully-Connected layer

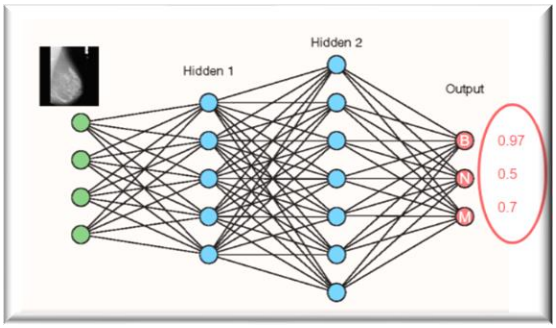

Figure 10: Mammography based image Fully-Connected layer Normal, Benign and Malignant

- As we can see in figure 9 once the important features are extracted and flattened this data will now feed to the fully connected layer and now after the training it will able to give the respective output of the respective inputted images like showed in Figure 9, Figure 10 and Figure 11.
    - I. **Figure 9 :** It shows how the tissue based images data (which are in the form of the array) will be flatten and then how this data will be feed to the fully connected layers in order to get the binary classified output that to what classification image belongs to **(Benign Vs Malignant).**
    - **II. Figure 10 :** It shows how the Mammography based images data (which are in the form of the array) will be flatten and then how this data will be feed to the fully connected layers in order to get the binary classified

output that to what classification image belongs to **(Benign Vs Normal VS Malignant).**

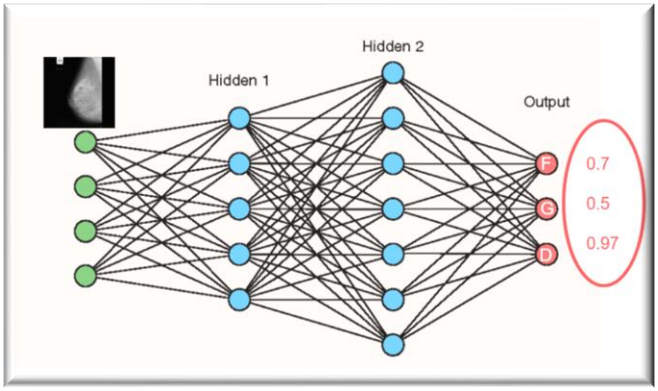

Figure 11: Mammography based image Fully-Connected Fatty, Fatty-Glandular and Dense-Glandular layer

III. **Figure 11 :** It shows how the Mammography based images data (which are in the form of the array) will be flatten and then how this data will be feed to the fully connected layers in order to get the binary classified output that to what classification image belongs to **(Fatty vs Fatty-Glandular vs Dense-Glandular).**

- So this was the Common system architecture of the Convolutional Neural Networks and their respective Deep Neural Networks (Fully-Connected Layer) as shown in the Figure 9, 10 and 11.

2.2.Transfer Learning Methods for Convolution Neural Networks: -

2.2.1. Definition and List of Different Transfer Learning Methods

- It is a unique learning method that can help utilize the previously learned knowledge and skills of different datasets on completely new problem solving or learning and understanding the datasets for model training. It plays a very crucial role in training any Convolutional neural network-related model and achieving outstanding accuracy so that the prediction of any input can be trustworthy. It is important to use in CNN because it can find the similarities and analogies between previous and actual learning content. There are many "CNN transfer learning methods" available but we used the most famous methods to derive the prediction results as accuracy stays important in our project, and we also had made the comparisons for different transfer learning CNN architectures result to derive that which architecture is most trustworthy for prediction. This transfer learning CNN architectures were decided based on F1-Score.

    I. VGG – 16
    II. VGG – 19
    III. InceptionV3
    IV. ResNet50
    V. InceptionResNetV2

1) VGG - 16:

Definition and Architecture:

VGG16 convolution neural net (CNN) architecture won ILSVR(ImageNet) competition in 2014. it's considered to be one amongst the wonderful vision model architecture till date. Most interesting thing about VGG16 is that rather than having an oversized number of hyper-parameters they focused on having convolution layers of 3x3 filter with a stride 1 and it used the padding and MaxPool layer of 2x2 filter with stride 2 everytime. It follows this arrangement of convolution and max pool layers consistently throughout the full architecture. Within the end it's 2 FC (fully connected layers) followed by a SoftMax for output. There is total 16 layer available in the VGG16 and these 16 layers consists of weights inside it. This network may be a pretty large network and it's about 138 million (approx.) parameters.

Figure 12: Architecture of VGG-16

3*3 filters are applied to the first two layers are convolutional layers in VGG-16 and 64 filters in total applied to those layers which result in 224*224*64. Every time filters which are used are of 3x3 with a stride 1. Now, a pooling layer was used with a max pool of 2*2 size and stride 2 which reduces the height and width of a volume from 224*224*64 to 112*112*64. Every time it is followed by 2 convolution layers with 128 filters as shown in the figure. This results in the new dimension of 112*112*128. After the pooling layer is used, the volume is reduced to 56*56*128. Two more convolution layers are added with 256 filters each followed by a down-sampling layer that reduces the size to 28*28*256. Two more stacks each with 3 convolution layers is separated by a max-pool layer. After the final pooling layer, 7*7*512 volume is flattened into a Fully Connected (FC) layer with 4096 channels and softmax output of 1000 classes.

2) VGG - 19:

Definition and Architecture:

VGG19 convolution neural net (CNN) architecture. VGG19 is have an oversized number of hyper-parameters, it focusses on having convolution layers of 3x3 filter with a stride of 1 and every time it used the padding and MaxPool layer with 2x2 filter of 2 strides. It follows this arrangement of convolution and max pool layers consistently throughout the full architecture. Within the end it's 3 FC (fully connected layers) followed by a SoftMax for output. There is total 19 layer available in the VGG16 and these 16 layers consists of weights inside it

Figure 13: Architecture of VGG-19

VGG-16 and VGG-19 are similar architecture only difference is that VGG-19 has 3 extra layers like one extra fully connected layer and it is of bigger in size.

3) ResNet50:

Definition and Architecture:

ResNet was the winner of the 2015 ILSVRC, it is a very deep network with 34–152 layers. The residual network is a full form of the ResNet model. This deep CNN model not only avoids the problem of model degradation but also achieves better accuracy. In ResNet-50, a structure called Bottleneck is used to reduce calculations and parameter quantities. In Bottleneck architecture, there are three layers which are 1 × 1, 3 × 3, and 1 × 1 convolution, where the two 1 × 1 layers play the role of reducing and then increasing dimensions, which gives the 3 × 3 layers the smallest input/output dimensions. It is similar to the VGG model 3 × 3 filters which are mostly used in this network. However, ResNet-50 has fewer filters and less complexity as compared to VGG. Many researchers have applied this model and received trustworthy results.

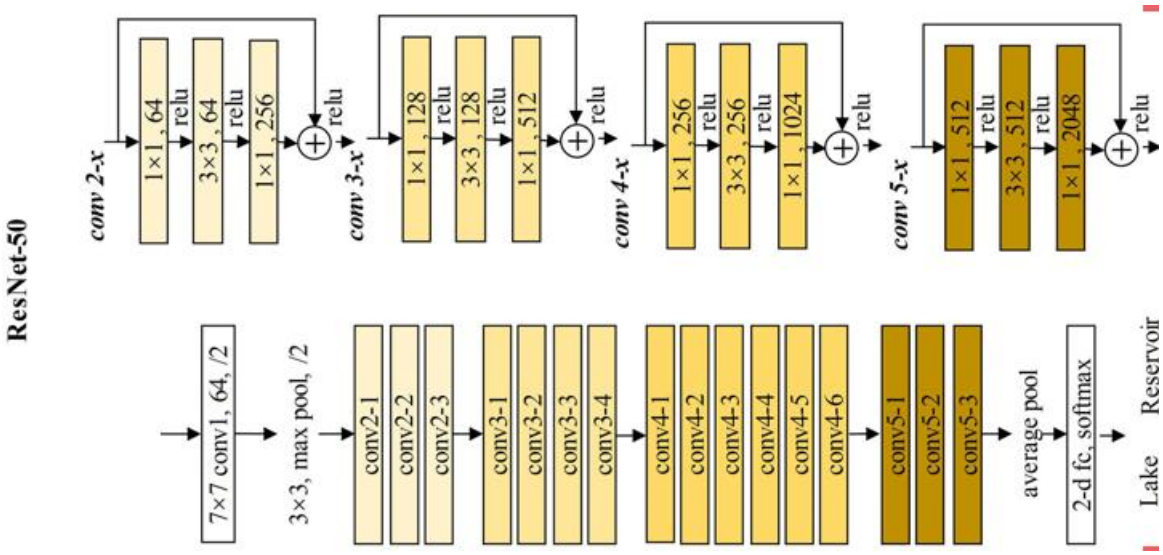

Figure 14: Architecture of ResNet50

ResNet was previously designed to classify the 1000-class RGB-band ImageNet database, while in this study we changed the last layer of the model to fit the two and three-category problem. ResNet50 50 layers consist of 23 million parameters.

4) Inception-V3:

Definition and Architecture:

Inception v3 is a Convolutional neural network architecture, it was used in this study as the network stacks 11 inception modules where each module consists of pooling layers and convolutional filters with RELU as activation function. There are 3 fully connected layers of size 1024, 512, and 3 that are added to the final concatenation layer. A dropout with the rate of 0.5 is applied before the fully connected layers as means of regularization. It has 42 layers due to which it is also known as a deep learning network with fewer parameters. To reduce the parameters, factorizing convolutions are used. For example, a 5 x 5 filter convolution can be done by two 3 x 3 filter convolutions. The parameters in this process reduces from 5 x 5 = 25 to (3 x 3) + (3 x 3) = 18. Thus, it brings a 28% reduction in the number of parameters. Due to the fewer number of parameters, chances of overfitting will be reduced and thus help to increase the accuracy. This model became 1st runner-up in ILSVRC-2015. We used the Inception V3 pretrained model to perform transfer learning on our 2 and 3 class tissue and mammography-based breast cancer detection problem.

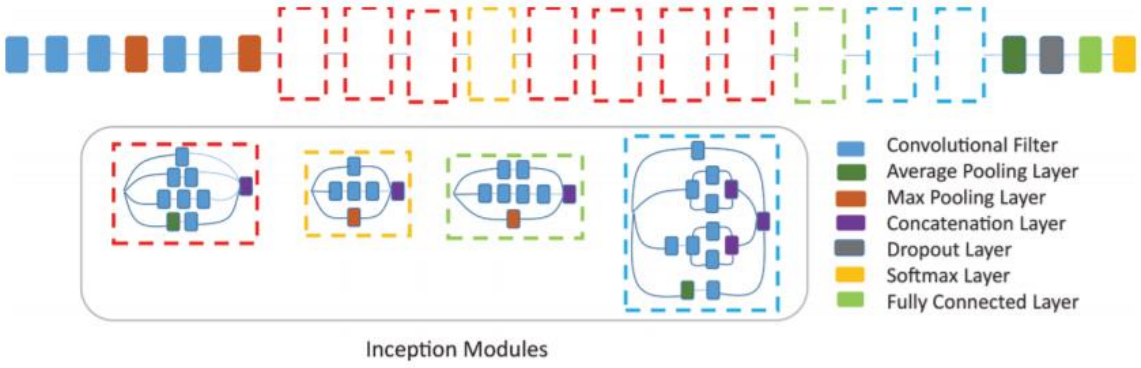

Figure 15: Architecture of InceptionV3

5) InceptionResNet V2:

Definition and Architecture:

InceptionResNet-V2 is a type of earlier version of Inception-V3 model which contains some algorithm methods from Microsoft's ResNet. It has 164 total layers due to which it consists of 54 million parameters

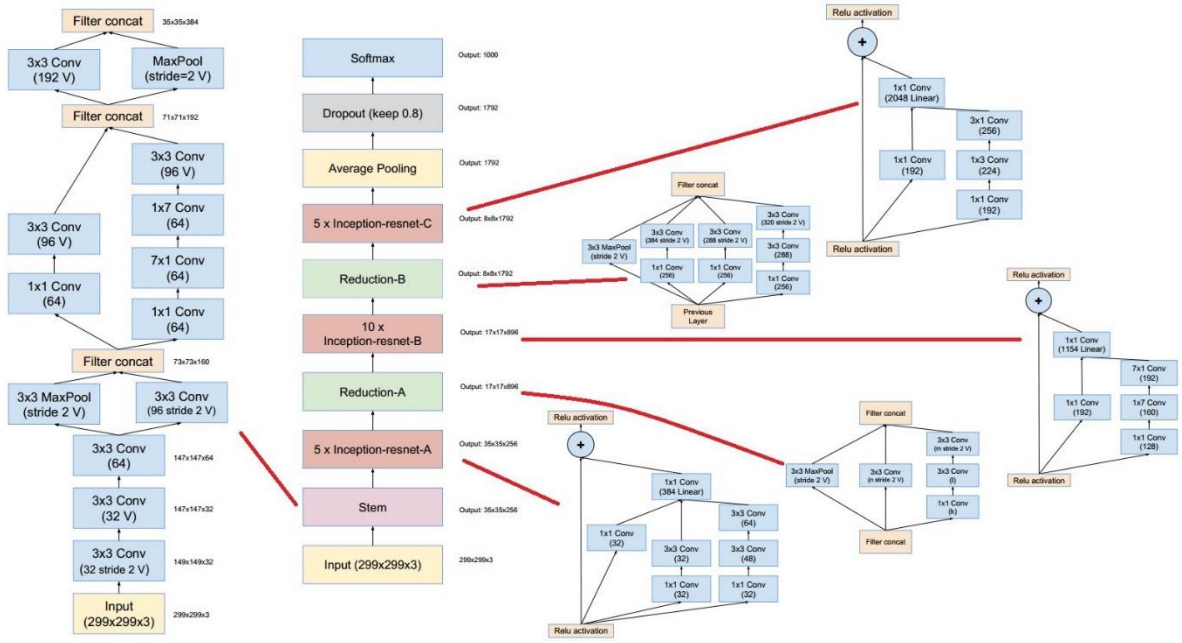

Figure 16: Hybrid of Inception net and Residual net

1. Each Inception block is followed by a filter expansion layer of 1×1 convolution without RELU activation which is used for increasing the dimensionality of the filters before the addition so that the depth of the input become similar to the output of the previous layers.

2. In the case of Inception-ResNet, only batch-normalization is used on topmost of the traditional layers, but not on top part of the summations. As shown in Figure 16 and Figure 17.

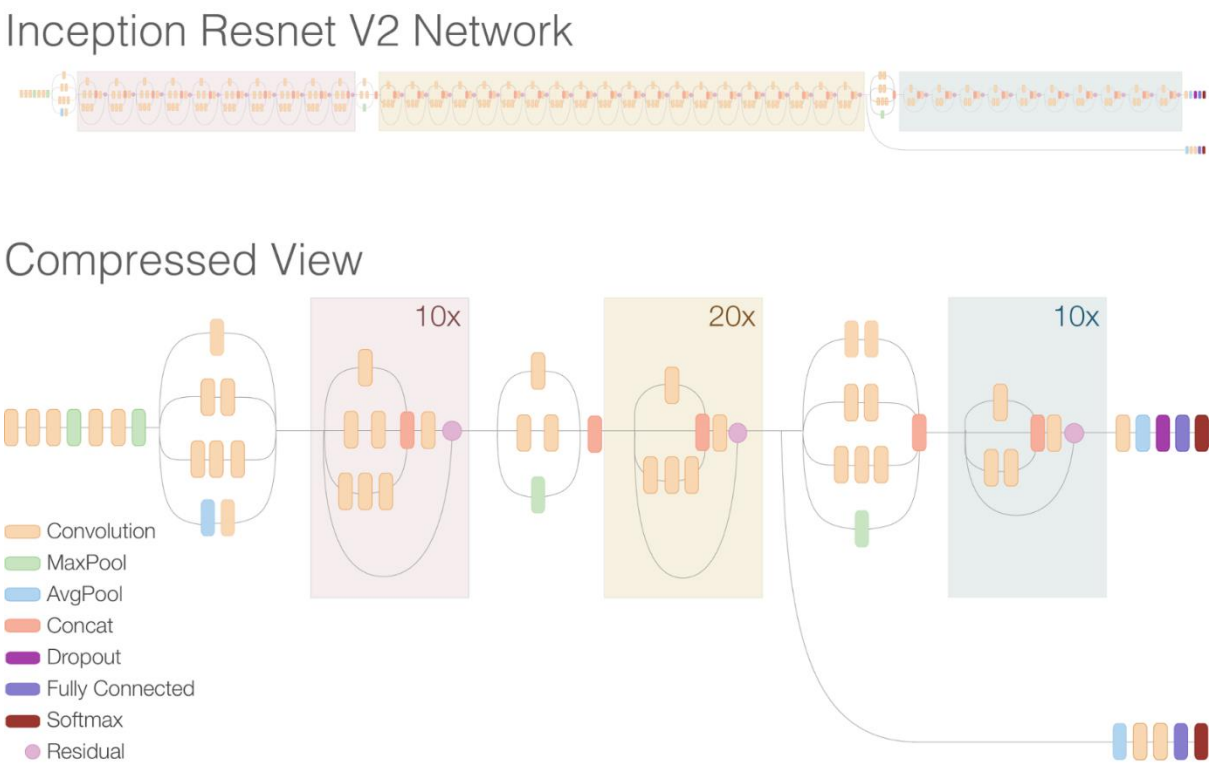

Figure 17: Architecture of InceptionResNetV2

2.2.2. Results & Comparisons on Tissue and Mammography based Images after Using Transfer Learning Methods of Convolutional Neural Networks

- After understanding the transfer learning methods, we will now compare the final results of various transfer learning model and their results on the basis of their F1-scores and after comparisons we will finalize the model for making prediction on the images like tissue based or mammography based.
    i. **Recall:** It is used to calculate the True Positive Rate (TPR), maximizing the **recall** will minimize the number of false negatives: **(TP/(TP+FN))**
    ii. **Precision:** Maximizing precision will minimize the number false positives: **(TP/(TP+FP))**
    iii. We will require the Recall and precision together so that model can have proper accuracy, so that we will be required to find the **F1-Score: (((1 + β^2) (Precision)(Recall)) / (Precision + Recall)) where β=1 as we are calculating F1-Score**
    iv. **Accuracy = (TP+TN)/(TP+TN+FP+FN).**
    v. **Precision Macro Average = (Precision of IDC + Precision of NON-IDC) / 2**
    vi. **Recall Macro Average = (Recall of IDC + Recall of NON-IDC) / 2**
    vii. **Weighted Average = ((support of NON-IDC / total test images) * Precision of NON-IDC) + ((support of IDC / total test images) * Precision of IDC)**

- Those comparisons and model selection are shown below depending on the different classifications.

1) Tissue Based Images (Benign Vs Malignant) Model:

Models	Recall	Precision	F1-Score	Accuracy	Selection
VGG-16	B-0.84 & M-0.88	B-0.88 & M-0.84	B-0.86 & M-0.86	0.86 ~ 86%	
VGG-19	B-0.93 & M-0.62	B-0.71 & M-0.90	B-0.81 & M-0.73	0.78 ~ 78%	
ResNet50	B-0.94 & M-0.71	B-0.76 & M-0.92	B-0.84 & M-0.80	0.82 ~ 82%	
Iception-V3	B-0.83 & M-0.92	B-0.91 & M-0.84	B-0.87 & M-0.88	0.87 ~ 87%	
InceptionResNet-V2	B-0.87 & M-0.91	B-0.91 & M-0.87	B-0.89 & M-0.89	0.89 ~ 89%	**Finalized**

Table 3: Comparisons and Selection of Models on the Basis of F1-Score

(B: Benign & M: Malignant)

- From the above table we can see that InceptionResNet-V2 has performed exceptionally well as compare to the other models in all terms like Recall, Precision and F1-Score, due to which we selected that model for tissue-based image for **benign vs malignant prediction.**
- Dataset of tissue-based image contains 277524 images of **benign and malignant** tissue and this dataset belongs to the 279 patients which is available on the Kaggle. This dataset images are scanned at 50x50 pixels which we had increased to 224x224 pixels during model training.
- **Results on InceptionResNet-V2 Finalized Model:**

  Confusion matrix, without normalization
  [[6838 1041]
  [690 7189]]

  Total Testing Images: 15758

  1. True Positive: 6838 → accurate IDC prediction
  2. False Positive: 1041 → NON-IDC predicted as IDC
  3. False Negative: 690 → IDC predicted as NON-IDC
  4. True Negative: 7189 → accurate NON-IDC prediction

- As we can see in the figure 18 and 19 that Training and validation loss is very low and almost same whereas training and validation accuracy are very high and same as well, this transfer learning method gave the 95.56% ROC Curve, which was the highest among other methods due to this InceptionResNet-V2 is finalized for tissue-based image dataset to do prediction on **Benign cells and Malignant cells**-based **images**.

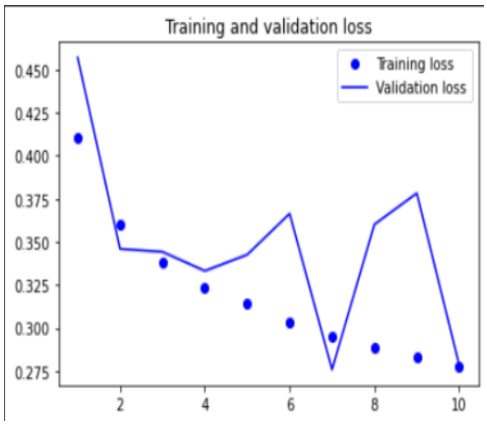

Figure 18: Training and Validation Loss of InceptionResNet-V2

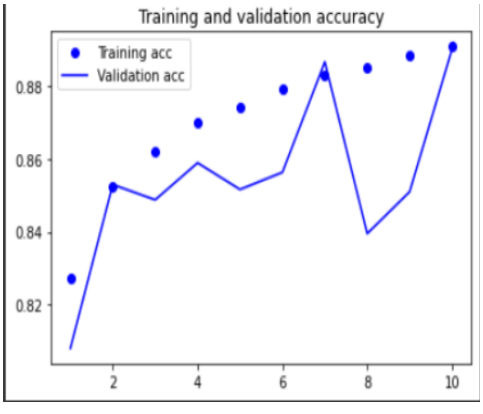

Figure 19: Training and Validation Loss of InceptionResNet-V2

2) Mammography Based Images (Normal Vs Benign Vs Malignant) Model:

Models	Recall	Precision	F1-Score	Accuracy	Selection
VGG-16	N: 0.85, B: 0.84 & M: 0.92	N: 0.82, B: 0.88 & M: 0.93	N: 0.83, B: 0.86 & M: 0.93	0.90 ~ 90%	
VGG-19	N: 0.82, B: 0.87 & M: 0.93	N: 0.85, B: 0.88 & M: 0.92	N: 0.83, B: 0.87 & M: 0.93	0.90 ~ 90%	**Finalized**
ResNet50	N: 0.84, B: 0.82 & M: 0.92	N: 0.81, B: 0.84 & M: 0.92	N: 0.83, B: 0.83 & M: 0.92	0.89 ~ 89%	
Iception-V3	N: 0.39, B: 0.46 & M: 0.88	N: 0.57, B: 0.64 & M: 0.75	N: 0.47, B: 0.54 & M: 0.81	0.71 ~ 71%	
InceptionResNet-V2	N: 0.07, B: 0.15 & M: 0.97	N: 0.61, B: 0.59 & M: 0.66	N: 0.12, B: 0.24 & M: 0.79	0.66 ~ 66%	

Table 4: Comparisons and Selection of Models on the Basis of F1-Score

(N: Normal, B: Benign & M: Malignant)

- From the above table we can see that VGG-19 and VGG-16 has performed exceptionally well as compare to the other models in all terms like Recall, Precision but F1-Score of VGG-19 for Benign image is 1% greater than VGG-16, due to which we selected the VGG-19 model for mammography-based image for **normal vs benign vs malignant prediction.**
- Dataset of mammography-based image contains 322 images of **normal, benign and malignant** images, due to which this dataset was increased using augmentation method to 14490 images so that more training can be done. This dataset images are trained at 224x224 pixels.
- **Results on VGG-19 Finalized Model:**

Confusion matrix, without normalization

```
   N   B   M
N [[217, 6, 31],
B [ 8, 193, 24],
M [ 44, 32, 750]]
```

- As we can see in the figure 20 and 21 that Training loss is very low whereas training accuracy are very high, this transfer learning method gave the 96.51% ROC Curve, which was the highest among other methods due to this VGG-19 is finalized for

mammography-based image dataset to do prediction on **Normal, Benign and Malignant based images.**

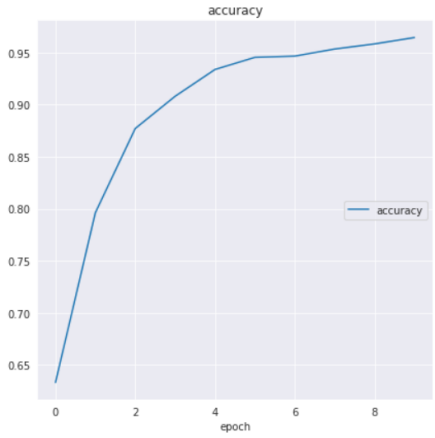

Figure 20: Training Accuracy of VGG-19

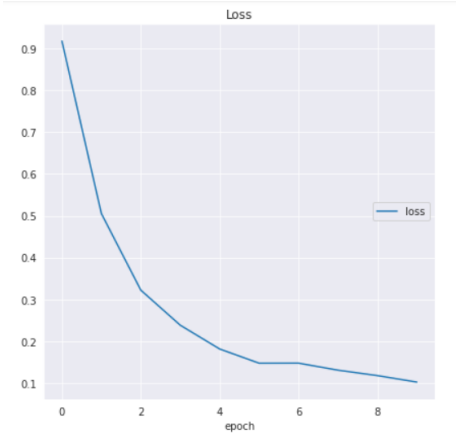

Figure 21: Training Loss of VGG-19

3) Mammography Based Images (Fatty Vs Dense-Glandular Fatty-Glandular) Model:

Models	Recall	Precision	F1-Score	Accuracy	Selection
VGG-16	F: 0.93, D: 0.95 & G: 0.97	F: 0.98, D: 0.97 & G: 0.91	F: 0.95, D: 0.96 & G: 0.94	0.95 ~ 95%	
VGG-19	F: 0.95, D: 0.94 & G: 0.94	F: 0.96, D: 0.91 & G: 0.96	F: 0.95, D: 0.93 & G: 0.95	0.94 ~ 94%	
ResNet50	F: 0.98, D: 0.97 & G: 0.92	F: 0.96, D: 0.95 & G: 0.96	F: 0.97, D: 0.96 & G: 0.94	0.96 ~ 96%	**Finalized**
Iception-V3	F: 0.89, D: 0.83 & G: 0.69	F: 0.83, D: 0.80 & G: 0.78	F: 0.86, D: 0.81 & G: 0.73	0.80 ~ 80%	
InceptionResNet-V2	F: 0.70, D: 0.78 & G: 0.65	F: 0.74, D: 0.74 & G: 0.67	F: 0.72, D: 0.76 & G: 0.66	0.71 ~ 71%	

Table 5: Comparisons and Selection of Models on the Basis of F1-Score

(F: Fatty, D: Dense-Glandular & G: Fatty-Glandular)

- From the above table we can see that ResNet50 has performed exceptionally well as compare to the other models in all terms like Recall, Precision and F1-Score. Due to this we selected the ResNet50 model for mammography-based image for **Fatty, Fatty-Glandular and Dense-Glandular prediction.**
- Dataset of mammography-based image contains 322 images of **Fatty, Fatty-Glandular and Dense-Glandular** images, due to which this dataset was increased using augmentation method to 14490 images so that more training can be done. This dataset images are trained at 224x224 pixels.
- **Results on ResNet50 Finalized Model:**

  Confusion matrix, without normalization

	F	D	G
F	[[384,	1,	7],
D	[3,	440,	18],
G	[15,	15,	422]]

- As we can see in the figure 22 and 23 that Training loss is very low whereas training accuracy are very high, this transfer learning method gave the 99.54% ROC Curve, which was the highest among other methods due to this ResNet50 is finalized for mammography-based image dataset to do prediction on **Fatty, Fatty-Glandular and Dense-Glandular based images**.

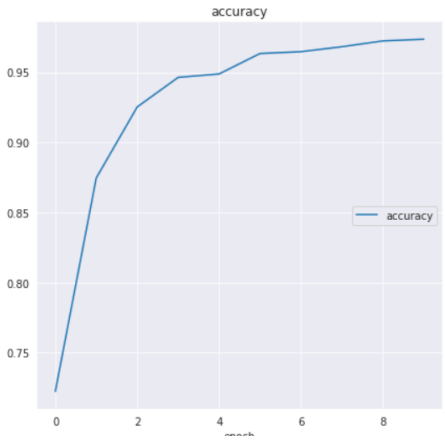

Figure 22: Training Accuracy of ResNet50

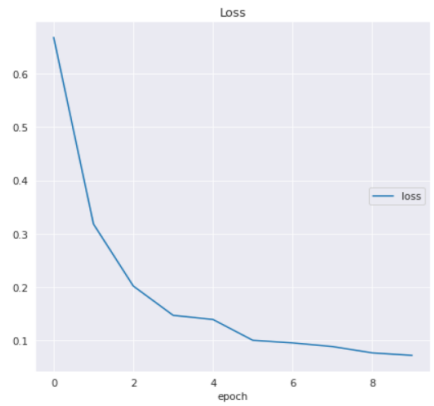

Figure 23: Training Loss of ResNet50

2.3. Neural Networks System Architecture: -

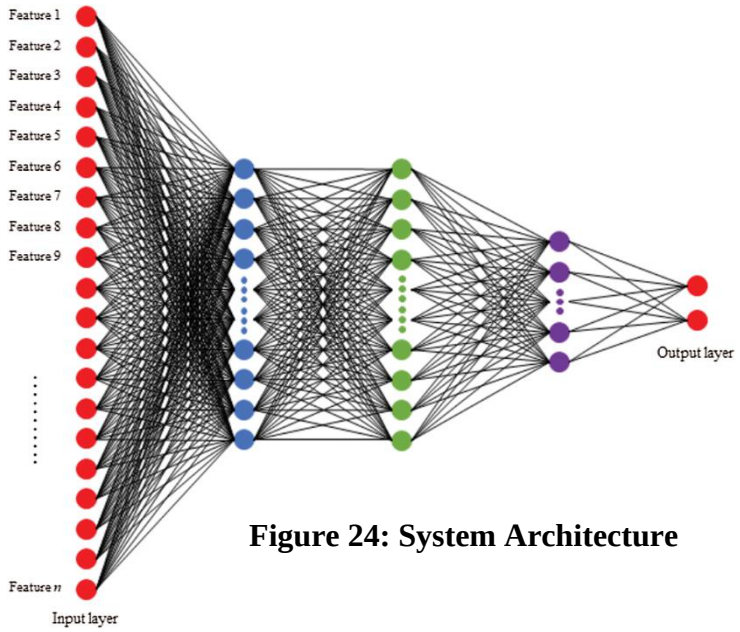

Figure 24: System Architecture

- In above Figure 24 we can see that Feature are first of all inputted to the input layers and then using supervised algorithms we can make separate the different classes on the basis of the feature to determine the classification. In this hidden layer plays very crucial role for calculation, distribution and for model training purpose. With the help of this technique, we can train any model for the numeric datasets. This architecture of fully-connected layer majorly use the softmax or sigmoidal activation function to produce the output for the output layer.
- In this input and output are given together to the neural networks due to which machine train itself on the basis of the inputted data and it can validate the model predictions with the actual output. If the prediction and actual output are not equivalent to each other than backtracking method is used to rearranged the weights of layers connection and this process will be continued till the difference between predictions and actual outputs is minimal.
- In our numeric dataset we have total **30 features** and we had used the various type of supervised algorithms so that model can be trained in very accurate manner.

2.4.Supervised Learning Algorithms for Artificial Neural Networks: -

2.4.1. Definition and List of Different Supervised Learning Algorithms

- **Supervised algorithm**: - It consists of an outcome or dependent variable which is to be predicted from a given set of independent variables or features. Using these set of features, we generate a method or use the algorithms or function that connects the inputs to actual outputs. The training process continues until the difference between predictions and actual outputs is minimal and model achieves a desired accuracy on the training features.

- **List of Supervised Learning Algorithms Used in this Project:**

I. Logistic Regression
II. Decision Tree Classifier
III. SVM - Linear
IV. SVM - RBF
V. Gausian NB
VI. KNN
VII. Random Forest Classifier

1) Logistic Regression:

- It is a classification algorithm. It is used to estimate discrete values like Binary (0/1, yes/no, true/false) values. In this set of independent variables or features are inputted. In simple, it predicts the probability of occurrence of an event by fitting variables to a logit function or sigmoidal function. It is also known as **logit regression**. Since, it predicts the probability, whose output values lie between 0 and 1.

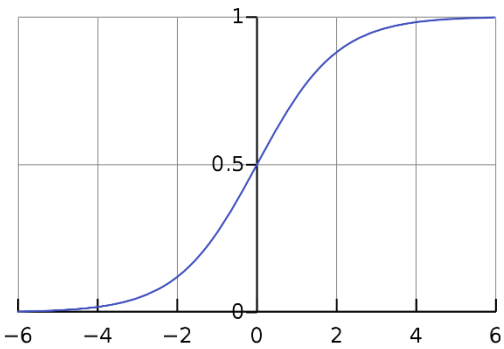

Figure 25: Logistic Regression Function

- As can be seen from the above figure 25 that algorithm will distribute two classes “0” and “1” on the basis of the features that mean below the 0.5 the prediction will belong to the “0” in our case “benign” and above the 0.5 prediction will belong to the “1” in our case “malignant.

2) Decision Tree Classifier:

- Decision Trees is a Supervised Machine Learning in which inputs with is corresponding output of the training data are explained from the beginning. The data gets split according to a certain parameter continuously. The tree can be properly explained with the help of two entities, decision nodes and decision leaves.

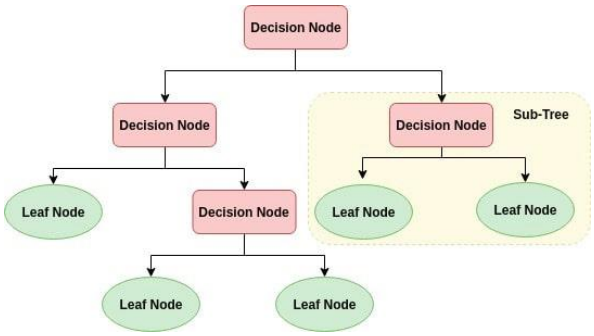

Figure 26: Basic Overview of Decision Tree Classifier

3) Support Vector Machine – Linear Kernel:

- In SVM Linear algorithm we need to find the points closest to the line from both the classes. These points are known as support vectors. After this we need to compute the distance between the line and the support vectors as shown in the figure. This distance is known as margin. Our main goal is to increase the margin. The optimal hyperplane is decided on the basis of the maximum margin.

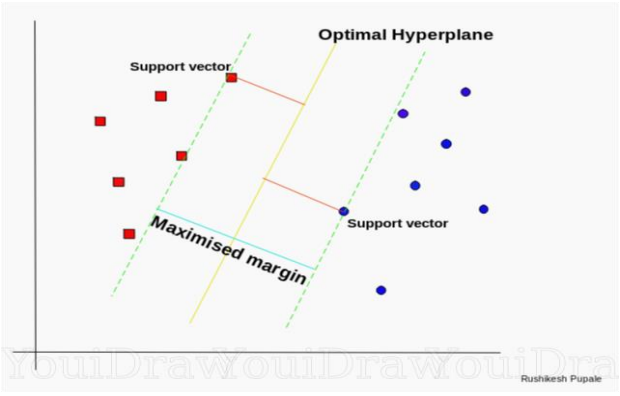

Figure 27: Basic Overview of SVM – Linear

- SVM tries to make a decision boundary in such a way that the separation between the two classes is as wide as possible.

4) Support Vector Machine – RBF Kernel:

- RBF (**Radial Basis Function)** is the type of SVM kernel which default form of SVM classification algorithm and can be described with the help of following formula:

$$K(x, x') = e^{-\gamma||x-x'||^2}$$

- A **kernel is a type of a function** which takes the original non-linear task or problems and converts it into a linear form within the space of higher dimensions.
- where gamma can be set manually and has to be >0. In sklearn, the default value of gamma in SVM classification algorithm is shown below:

$$\gamma = \frac{1}{n\,features * \sigma^2}$$

5) Gaussian NB:

- **Gaussian NB (Naive Bayes)** is a type of Naive Bayes that uses Gaussian normal distribution and **supports continuous data**. It is a supervised machine learning classification algorithm based on the Bayes theorem. It is also a simple classification technique with high functionality. It is used for the high input dimensionality.
- In below shown formula of "Gaussian NB" we can see we are calculating the Standard Deviation, Variance and Mean all together at the same time for distributions. In below figure 28 we can see that in graph when we draw a line exactly from the between of the bell curve, we can find a point exactly at the intersection of it horizontal line which is known as "μ"

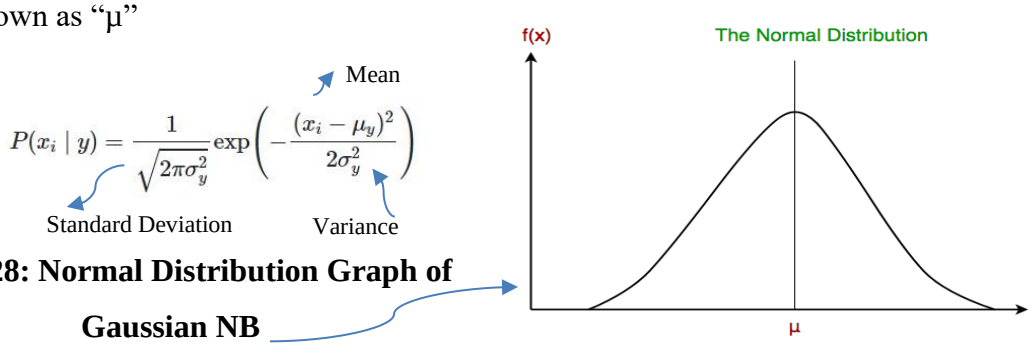

Figure 28: Normal Distribution Graph of Gaussian NB

6) KNN:

- K nearest neighbors is useful for both classification and regression problems. However, its wider application in classification problems in the industry. It is a simple type algorithm that stores and carry all available cases and classifies new cases by a majority vote of its k neighbors. The case being assigned to the class is measured by a distance function.
- First number of neighbors are selected which gets stored in K, then Euclidean distance of given feature is calculated with every feature available in the dataset after that select the K-nearest neighbors with the help of Euclidean distance formula. Among these k neighbors, count every number of the data points in each category, then assign the new data points to that category for which the number of the neighbor is maximum. Following this help, us to develop our final model.
- Suppose there are two categories Category A = Benign and Category B = Malignant, and we have a new data point x, so in which of these categories it will lie. To solve this, we need a K-NN algorithm. With its help, we can identify class or category of a particular dataset. Consider the below diagram:

$$d(\mathbf{p},\mathbf{q}) = d(\mathbf{q},\mathbf{p}) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \cdots + (q_n - p_n)^2}$$

$$= \sqrt{\sum_{i=1}^{n} (q_i - p_i)^2}.$$

Euclidean Distance Formula

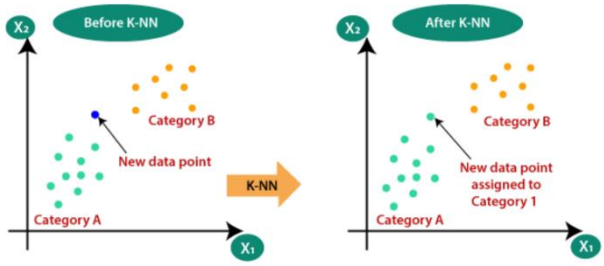

Figure 29: K - Nearest Neighbors Distribution

7) Random Forest Classifier:

- **Random** forest **classifier** creates a set of decision trees from **randomly** selected subset of training set. It then aggregates every vote from various decision trees to finalized the final class of the test object for the final output.
- Each node of the decision tree works on a random subset of features or data to calculate the output. The random forest then combines all the output of individual decision trees to generate the final output.

- This can be easily understood with the help of following diagram:

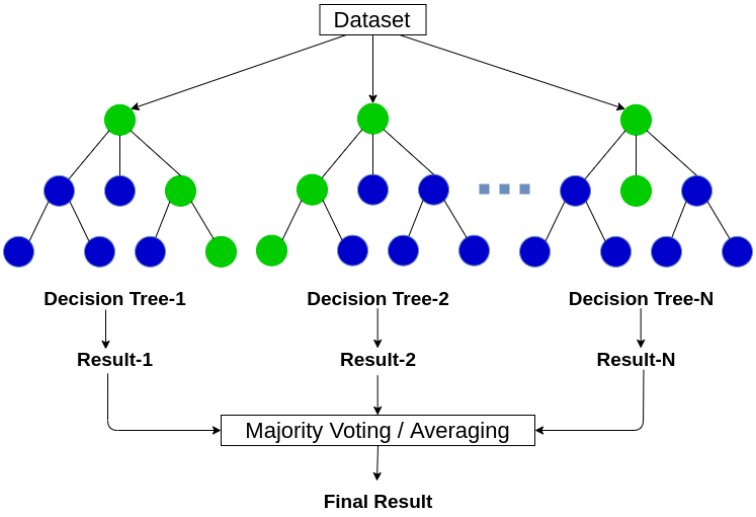

Figure 30: Overview of Random Forest Classifier

2.4.2. Results and Comparisons on Numeric Dataset after Using Various Supervised Learning Algorithms of Artificial Neural Networks: -

❖ **Numeric Based Data (Benign Vs Malignant) Model:**

Models	Recall	Precision	F1-Score	Accuracy	Selection
Logistic Regression	B: 0.89 & M: 0.97	B: 0.95 & M: 0.93	B: 0.92 & M: 0.95	0.94 ~ 94%	
Decision Tree Classifier	B: 0.94 & M: 0.93	B: 0.90 & M: 0.95	B: 0.92 & M: 0.94	0.93 ~ 93%	
SVM – Linear Kernel	B: 0.98 & M: 0.94	B: 0.92 & M: 0.98	B: 0.95 & M: 0.96	0.96 ~ 96%	**Finalized**
SVM – RBF Kernel	B: 0.85 & M: 0.99	B: 0.98 & M: 0.90	B: 0.91 & M: 0.94	0.93 ~ 93%	
Gaussian NB	B: 0.91 & M: 0.94	B: 0.91 & M: 0.94	B: 0.91 & M: 0.94	0.93 ~ 93%	
KNN	B: 0.94 & M: 0.94	B: 0.92 & M: 0.95	B: 0.93 & M: 0.95	0.94 ~ 94%	
Random Forest Classifier	B: 0.96 & M: 0.94	B: 0.92 & M: 0.97	B: 0.94 & M: 0.95	0.95 ~ 95%	

Table 6: Comparisons and Selection of Models on the Basis of F1-Score (B: Benign & M: Malignant)

- From the above table we can see that Support Vector Machine (SVM) – Kernel Linear has performed exceptionally well as compare to the other models in all terms like Recall, Precision and F1-Score. Due to this we selected the Support Vector Machine (SVM) – Kernel Linear model for Numeric based data for **Benign and Malignant prediction.**
- Dataset of Numeric data contains 30 features of **Benign and Malignant**, and using this various feature we had prepared various models using very famous supervised learning algorithms and from that we selected the Support Vector Machine (SVM) – Kernel Linear because of the very high accuracy of 96% as compare to the other supervised algorithms.
- **Results on Support Vector Machine (SVM) – Kernel Linear Finalized Model:**

  Confusion matrix, without normalization

```
        B    M
B     [[46,  1],
M     [ 4, 63]]
```

2.5.Business Model Canvas of Breast Cancer Detection: -

- BMC also known as Business Model Canvas, it is a type of a strategic as well as management model which is used to develop the new business models or record, report and update the existing ones. It offers a visualization with elements to describe the products Key Partners, Key Activities, Key Resources, Value Propositions, Customer Relationship, Channels, Customers Segments, Cost Structure, Revenue Streams

1) Key Partners

- Doctors
- Oncologist
- Lab Technicians
- Hospitals & Clinics
- MI / AI Developers
- Advertisers

2) Key Activities

- Datasets collections for implementation
- Tools & Technology selection
- Implementing various models for various task using “PYHTON”
- Updates & Results verifications

3) Key Resources

- Higher level graphic card
- Research to increase accuracy by increasing datasets
- MI / AI / PYTHON developers

4) Value Propositions

- Complex cancer detection
- Easy to use for oncologist
- Easy to update
- Reduce risk in false detection
- Trustworthy accuracy in cancer detection
- AI / ML based Algorithms
- Generate final report of result as “PDF”

5) Customer Relationship

- Automated Services
- Dedicated personal assistance
- 89% average cancer detection accuracy
- Self-Services

6) Channels

- Advertisements
- Direct Purchase
- Evaluation & Proofs
- Website

7) Customers Segments

- Private Hospitals
- Government Hospitals
- Laboratories
- Clinics
- Oncologist
- Lab Technicians
-

8) Cost Structure

- Architecture design
- Update datasets
- Algorithms & Models updates
- New higher level graphic cards
- Deployment & Maintenance

9) Revenue Streams

- On every new update
- On report generations

3. Implementation

3.1. Front-End

For front-end we had used the HTML, CSS, JAVASCRIPT, and FLASK framework, using all this front-end technology we develop the following website.

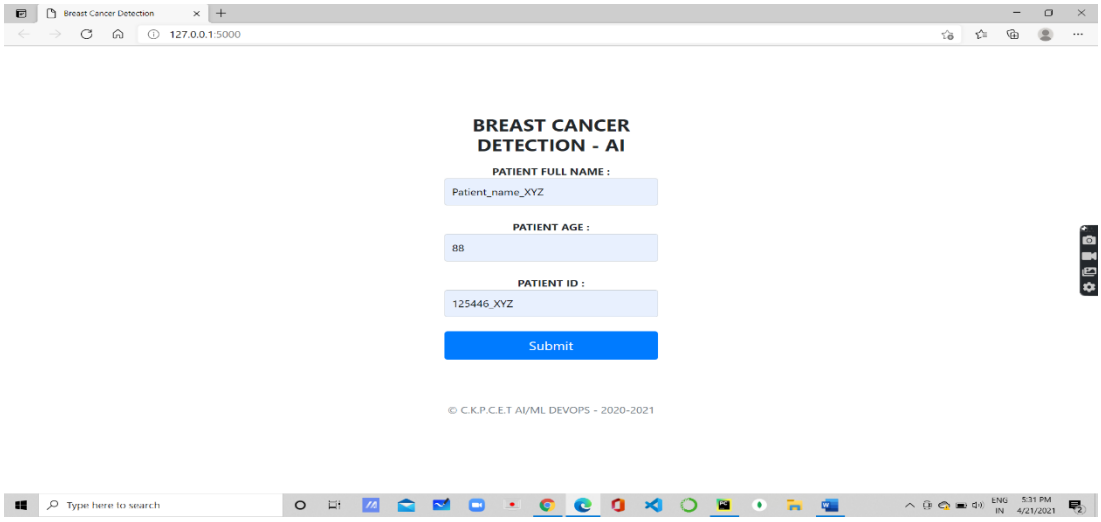

Figure 31: Index Page

- In above figure 31 we can see breast cancer specialist have to enter the information of the Patient like Patient full name, age and Unique Id.

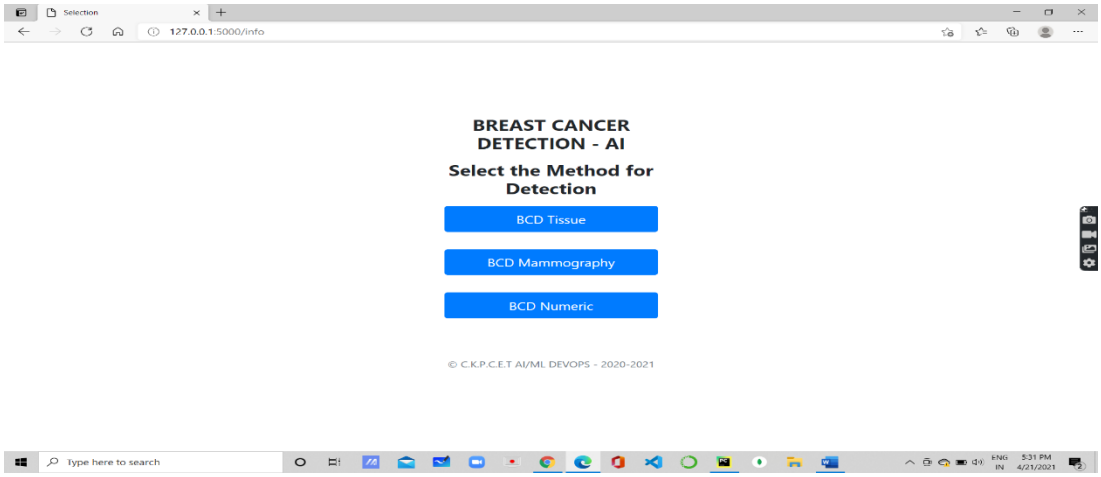

Figure 32: Method Selection Page

- In above figure 32 breast cancer specialist have to select the method for breast cancer detection and they have to provide the respective data during the prediction.

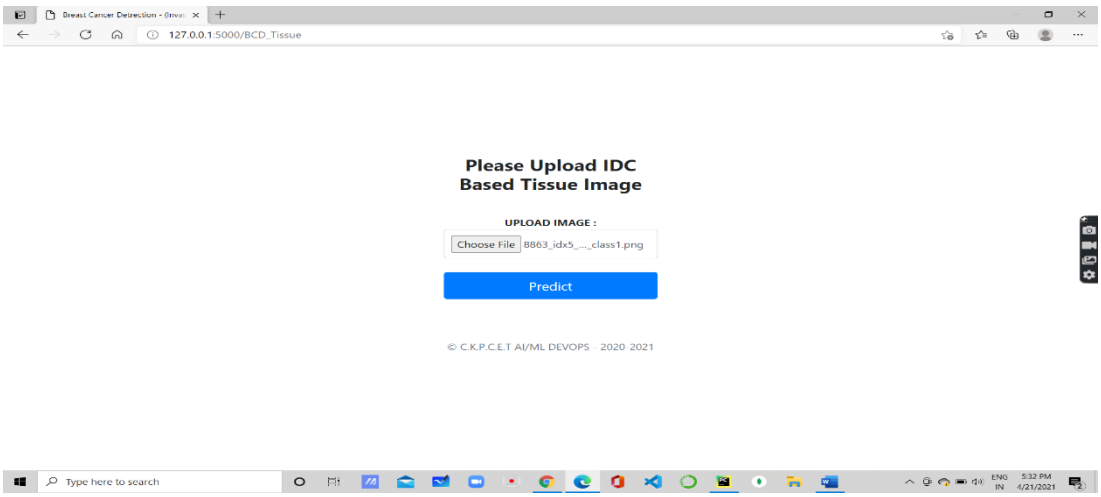

Figure 33: Tissue Based Method Selected

- Let say, breast cancer specialist has **tissue or biopsy-based image** then they will select the Tissue-based method and according to that they can carry out the process of prediction by just clicking the button "Predict" and they will able to see the predicted results shown in Figure 33.

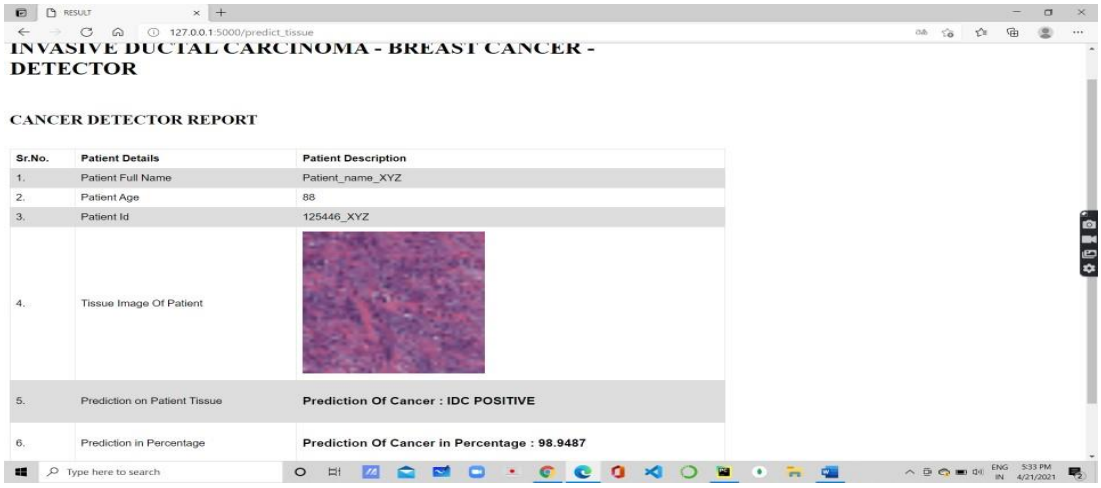

Figure 34: Tissue Based Method – Result Page

- In above figure 34 it is shown that how breast cancer specialist will see the final prediction result when they press the "predict" button to see the prediction. Breast cancer specialist can see the result like is the inputted image is Benign (IDC_NEGATIVE) or Malignant (IDC_POSITIVE) with percentage.

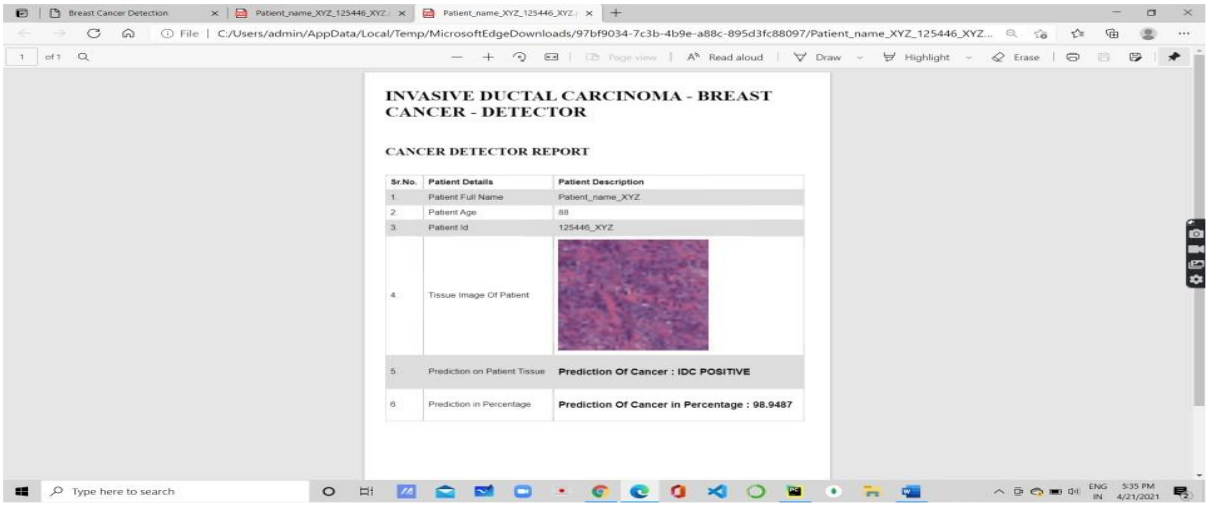

Figure 35: Report Generation in form of PDF

- In above Figure 35 it is shown that the project has also "**Generate PDF**" facility where breast cancer specialist can generate the whole final report of prediction report with the patient info. The PDF will be downloaded in form of "**Patient-name_Patien-id.pdf**".

- Let say now, breast cancer specialist has **Mammography-based image** then they will select the Mammography -based method and according to that they can carry out the process of prediction by just clicking the button "Predict" and they will able to see the predicted results shown in Figure 36.

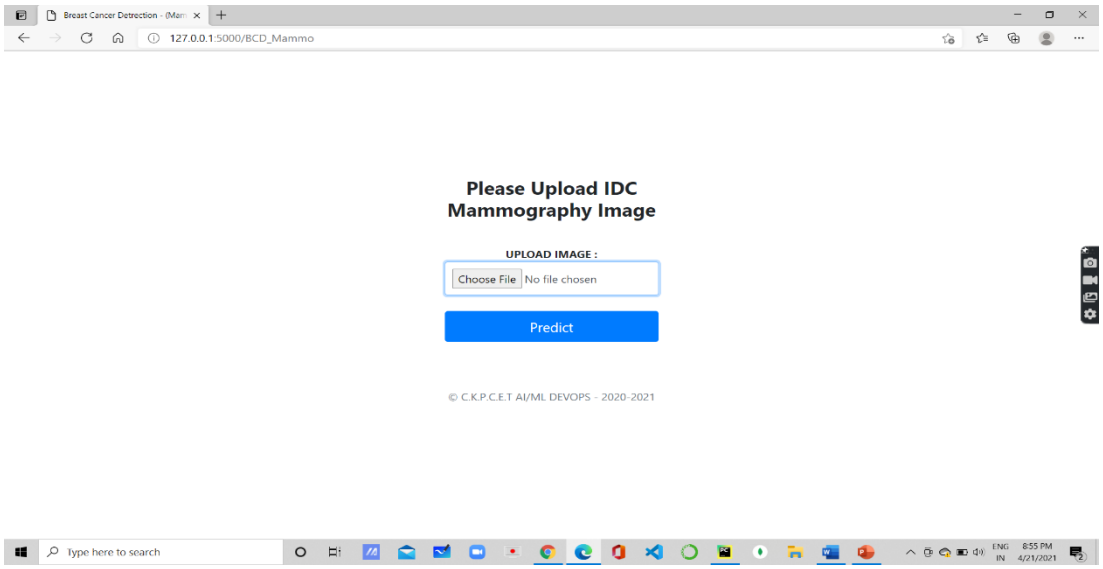

Figure 36: Mammography Based Method Selected

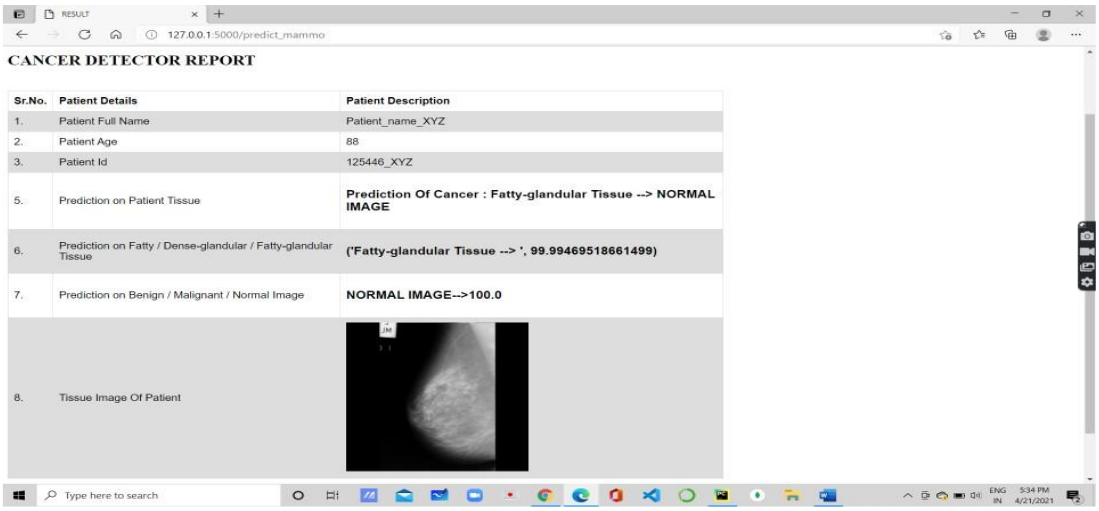

Figure 37: Mammography Based Method – Result Page

- In above figure 37 it is shown that how breast cancer specialist will see the final prediction result when they press the "predict" button to see the prediction. Breast cancer specialist can see the result like, is the inputted image belongs to Normal, Benign or Malignant as well as and Fatty, Fatty-Glandular or Dense-Glandular together with their respective percentages.

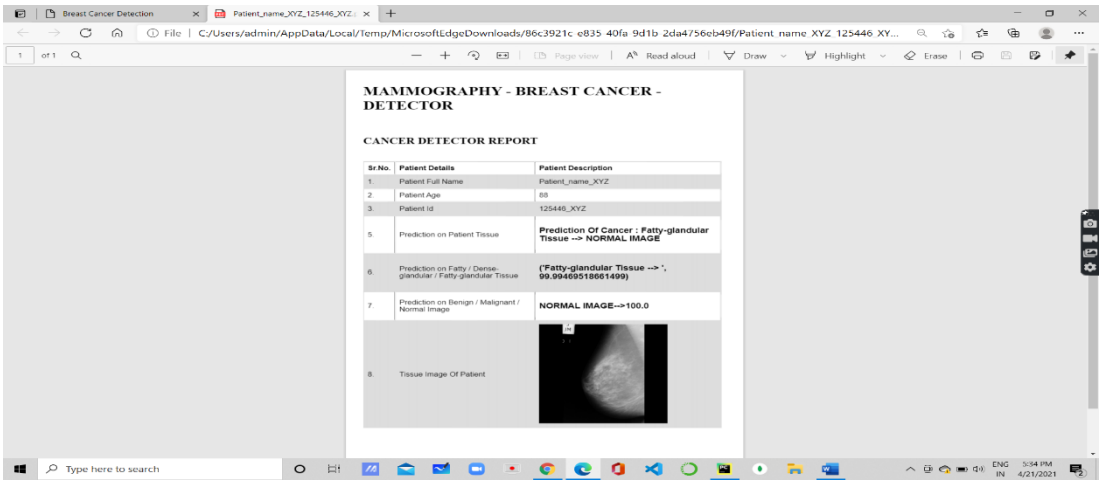

Figure 38: Report Generation in form of PDF

- In above Figure 38 it is shown that the project has also "**Generate PDF**" facility where breast cancer specialist can generate the whole final report of prediction report with the patient info. The PDF will be downloaded in form of "**Patient-name_Patien-id.pdf**".

- Finally, if breast cancer specialist has **Numeric data,** then they will select the Numeric-based method and according to that they can carry out the process of prediction by just clicking the button "Predict" and they will able to see the predicted results shown in Figure 39. In this method will demand 30 features for carrying out prediction. This 30 feature will be seen on the website when breast cancer specilist select the Numeric based method for breast cacncer detection.

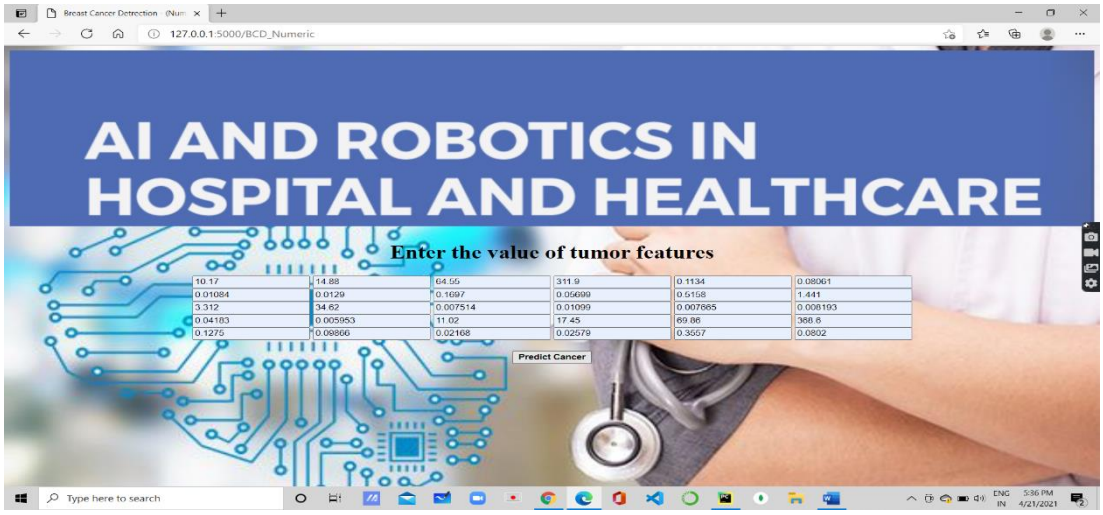

Figure 39: Numeric Based Method Selected

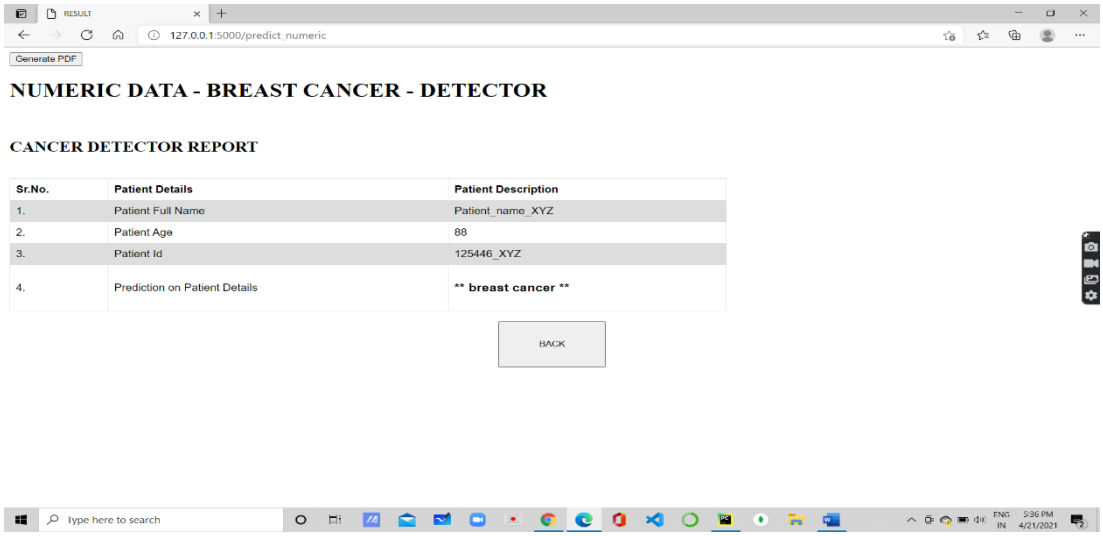

Figure 40: Numeric Based Method – Result Page

- In above figure 40 it is shown that how breast cancer specialist will see the final prediction result when they press the "predict" button to see the prediction. Breast cancer specialist can see the result like, is the inputted data belongs to Benign (Not a Breast Cancer) or Malignant (**Breast Cancer).

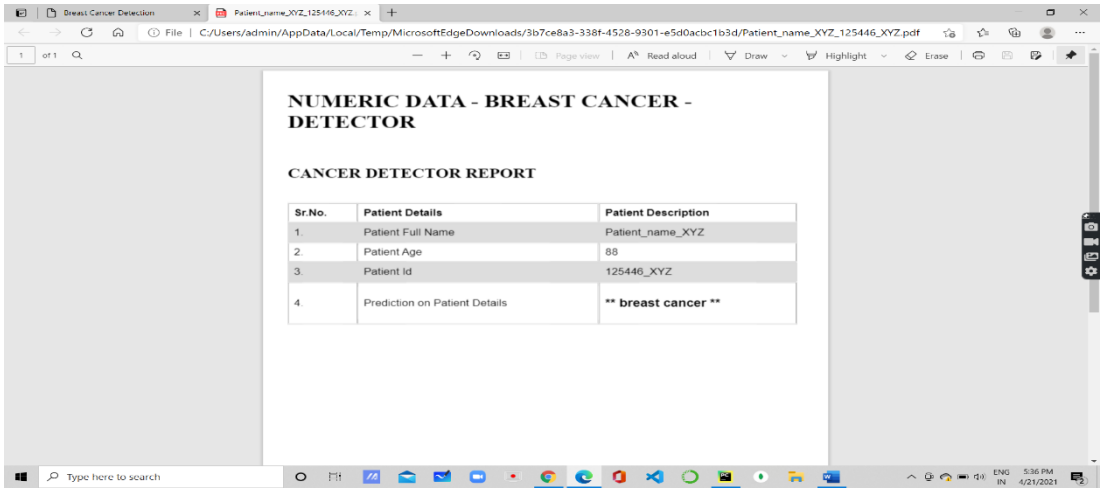

Figure 41: Report Generation in form of PDF

- In above Figure 41 it is shown that the project has also "**Generate PDF**" facility where breast cancer specialist can generate the whole final report of prediction report with the patient info. The PDF will be downloaded in form of "**Patient-name_Patien-id.pdf**".

3.2. Back-End

On the back-end side we had use the:

- **Tools: -**

    i. **Numpy** is a inbuilt library for the **Python** programming language, It helps in supporting the large, multi dimensions arrays or matrices. It can also handle the large collection of high level mathematical based functions to carry out various operations on these arrays.

    ii. **Pandas** is mainly **used** for data analysis

    iii. **TensorFlow** offers APIs that facilitates Machine Learning operations. **TensorFlow** also has a faster compilation time as compare to the other Deep Learning libraries such as Keras.

    iv. **Keras** is a very powerful and simple to use free open-source **Python** library for developing, validating and evaluating deep learning models and train deep neural network models in very efficient manner.

    v. **sklearn** library contains a lot of efficient tools to carry out machine learning and statistical modeling operations. It has also tools for classification, regression, clustering and dimensionality reduction operations.

    vi. **OpenCV (CV2)** is **used** for all sorts of image and video analysis.

    vii. **Matplotlib** is a graph plotting library in **Python** programming language and it's an extension of NumPy for numerical mathematics.

    viii. **Skimage & Imageio** is used to read images.

- **Technology Stack: -** Python 3.7 and 3.8, HTML & CSS, Javascript and Flask as a framework.
- **Platform for code writing & implementation: -** Anaconda, Jupyter notebook, Pycharm and Google collab.
- On back-end side we had implemented the different models using above listed tools for different type of datasets and after the model training, we had loaded the saved model in the flask application which was written in python. This all models then together were integrated with the help of python defined function under the ("app.route()") with their respective methods selection page and result page to see their final prediction results and also to generate the final result report in form of PDF.

4. Conclusion

- In conclusion, our project will assist doctors or Lab Technicians with a diagnosis of breast cancer. Breast cancer is one of the leading causes of death in women and there are many cases available where due to misdiagnosis, patients have lost their lives. So, our project will reduce the possibility of misdiagnosis by stating the possibility of breast cancer merely by the pictures of the breast's.
- Lastly, our project will help to detect the breast cancer using convolutional neural networks-Transfer learning methods and Artificial neural networks-Supervised learning methods in following way:
    i. Using **InceptionResNetV2 a convolutional neural networks-Transfer learning methods-based models** was developed in this project for carrying out prediction of Benign and Malignant **Tissue based images.**
    ii. Using **VGG-19 a convolutional neural networks-Transfer learning methods-based models** was developed in this project for carrying out prediction of Normal, Benign and Malignant **Mammography based images.**
    iii. Using **ResNet50 a convolutional neural networks-Transfer learning methods-based models** was developed in this project for carrying out prediction of Fatty, Fatty-Glandular and Dense-Glandular and Malignant **Mammography based images.**
    iv. Using **Support Vector Machine (SVM) – Linear Kernel an artificial neural networks-Supervised learning methods-based models** was developed in this project for carrying out prediction of Benign and Malignant **Numeric dataset.**
- In conclusion, we faced many difficulties related to the model training at first due to the lack of knowledge and low computational power but at last we were able to develop all the models and integrated them to the single website so that we can provide the facilities of breast cancer detection in every manner to the doctors with the good accuracy.

5. Reference

- Thermography-based breast cancer screening using a measure of symmetry: https://patents.google.com/patent/WO2018002861A1/en?oq=breast+cancer+detection+machine+learning **(Citations:** Patent citations (8), Non-patent citations (3), Cited by (3), Legal events, Similar documents, Priority and Related Applications**)**.
- Methods and systems for assessing risk of breast cancer recurrence: https://patents.google.com/patent/US20170091937?oq=breast+cancer+detection+machine+learning **(Citations:** Patent citations (29), Cited by (22): Legal events, Similar documents, Priority and Related Applications**)** .
- System and method for breast cancer detection in X-ray images: https://patents.google.com/patent/US10111632B2/en?q=breast+cancer+detection+machine+learning&oq=breast+cancer+detection+machine+learning **(Citations:** Patent citations (5), Non-patent citations (5), Cited by (5), Legal events, Similar documents, Priority and Related Applications**)**.
- Deep-learning-based cancer classification using a hierarchical classification framework: https://patents.google.com/patent/US20190183429A1/en?q=breast+cancer+detection+convolution+neural+country&oq=breast+cancer+detection+convolution+neural+country **(Citation:** Patent citations (10), Cited by (20), Legal events, Similar documents, Priority and Related Applications**).**
- Pattern recognition algorithm of early breast cancer detection in mammogram images and cyst"s size measurement in sonomammomgram images through physical characteristics using python: http://ipindiaservices.gov.in/publicsearch/
- Breast Cancer Detection and Diagnosis Using Mammographic Data: Systematic Review: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6688437/ **(Citation:** Cited by the Following 11 Articles(Cited by ... (nih.gov)) **).**

- Deep CNNs for microscopic image classification by exploiting transfer learning and feature concatenation: https://www.researchgate.net/publication/324961229_Deep_CNNs_for_microscopic_image_classification_by_exploiting_transfer_learning_and_feature_concatenation **(Citation: cited by total 72 articles).**
- A Deep Learning Model to Predict a Diagnosis of Alzheimer Disease by Using 18F-FDG PET of the Brain: https://pubmed.ncbi.nlm.nih.gov/30398430/ **(Citation: cited by total 50 articles).**
- Most Popular Convolutional Neural Networks Architectures: https://towardsdatascience.com/convolutional-neural-networks-most-common-architectures-6a2b5d22479d.
- ToLeD: Tomato Leaf Disease Detection using Convolution Neural Network: https://www.sciencedirect.com/science/article/pii/S1877050920306906 **(Citation: cited by total 4 articles).**
- Breast cancer detection by leveraging Machine Learning: https://www.sciencedirect.com/science/article/pii/S2405959520300801 **(Citation: cited by total 6 articles).**
- A Novel Approach for Breast Cancer Detection Using Data Mining Techniques: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2994932 **(Citation: cited by total articles).**
- Early Detection of Breast Cancer Using Machine Learning Techniques: https://journal.utem.edu.my/index.php/jtec/article/view/4706/3462 **(Citation: No Citations are available)**
- A Comparative Analysis of Nonlinear Machine Learning Algorithms for Breast Cancer Detection: https://www.semanticscholar.org/paper/A-Comparative-Analysis-of-Nonlinear-Machine-for-Bataineh/30e071cbb7d7a2df62f72f5ef04396d8cb15b388#paper-header **(Citation: cited by 16 articles,** [PDF] A Comparative Analysis of Nonlinear Machine Learning Algorithms for Breast Cancer Detection | Semantic Scholar**).**
- Analysis of Breast Cancer Detection Using Different Machine Learning Techniques: https://link.springer.com/chapter/10.1007/978-981-15-7205-0_10.
- Deep CNNs for microscopic image classification by exploiting transfer learning and feature concatenation: https://www.researchgate.net/publication/324961229_Deep_CNNs_for_microscopic_i

mage_classification_by_exploiting_transfer_learning_and_feature_concatenation **(Citation: cited by 3 articles,** SpringerCitations - Details Page (springernature.com)**).**

- Predictive Machine Learning Techniques for Breast Cancer Detection: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.643.9&rep=rep1&type=pdf **(Citation: cited by 33 articles,** https://scholar.google.co.in/scholar?cites=17097301280703163 64&as_sdt=2005&sciodt=0,5&hl=en**).**
- A Comparative Analysis of Nonlinear Machine Learning Algorithms for Breast Cancer Detection: http://www.ijmlc.org/vol9/794-L0201.pdf **(Citation: No Citations are available).**
- Platforms: Youtube Tutorials and Kaggle for various datasets.