Published by :
http://www.ijert.org

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 9 Issue 05, May-2020

Breast Cancer Prediction using ML Techniques

Prof. Mounica B[*1] , Sudarshan C[2] , Pranav Pandhi[3] , Somya Singh[4] , Ashwini Holla[5]
[*1] Professor, Information Science, New Horizon College of Engineering, Bangalore, Karnataka, India.
[2,3,4,5]Information Science, New Horizon College of Engineering, Bangalore, Karnataka, India.

***Abstract*—Breast cancer is one amongst the foremost dangerous kinds of cancer exists among ladies. The breast cancer is diagnosed using microscopic anatomy pictures . The aim of this paper is to classify different kinds of breast cancer using histology images. The classification of histology images can be effectively done by image process techniques. Among totally different image process algorithms, deep learning offers the most effective performance for image classification applications. There are different convolutional neural network(CNN) architectures used for classification purpose like AlexNet, Inception-Net, ResNet etc. Since conventional convolutional neural network.**

***Index Terms*—Breast Cancer, Mammography, ResNet50, Histological images, CNN, ROC.**

I. INTRODUCTION

CANCER is one amongst the foremost common diseases found in each men and ladies.. Since the treatments are in the advanced stage, the death rate for cancer has been considerably reduced in past few years [1]. Although advanced treatments are being found for cancer, one amongst the most challenges during this field is that the early detection of disease. There are different types of cancer supported the body part affected. Among this, breast cancer is one amongst the deadliest kinds of cancer in ladies and it's the foremost common explanation for death in ladies aged between twenty and fifty nine [1-12].

To study the human breast, diagnostic technique is wide used as a diagnostic and a screening tool that uses Xrays. The target of diagnostic technique is that the premature revealing of breast cancer, usually through detection of characteristic microcalcifications[20] and/or masses. Mammography is that the solely only effective and viable technique to detect breast cancer in particular in the case of minimal tumors [13][18][21]. Benign lesions represent changes in traditional structures of breast parenchyma that aren't directly connected with progression to malignancy [18][19].

In this paper, the breast cancer classification using histopathology images is done using deep learning architectures. Among the present ways, Convolutional Neural Networks (CNN) are the backbone for image classification applications.

The purpose of Convolutional layer is to receive a feature map. Usually, we have a tendency to begin with low range of filters for low-level feature detection. The deeper we have tendency to enter the CNN, the additional filters we use to find high-level options. Feature detection relies on 'scanning' the input with the filter of a given size and applying matrix computations so as to derive a feature map.

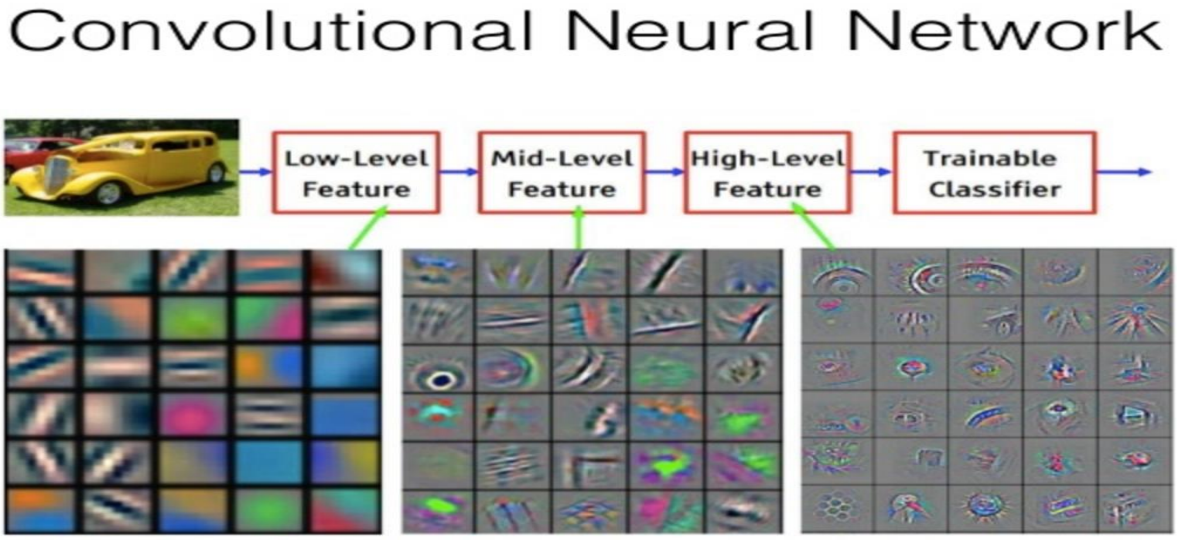


Fig 1.Convolution Operation.

Image Classification :- The entire image classification pipeline is formalized as : Our input maybe a training dataset that consists of *N* images, every label with one amongst two totally different categories.Then, we use this training set to train a classifier to learn what every one of the classes looks like. In the end, we evaluate the standard of the classifier by asking it to predict labels for a brand new set of images that it has never seen before. We'll then compare the actuality labels of those images to those predicted by the classifier.

The organization of the paper is as follows: the Related works are presented in section II. Section III consist of Literature survey ,Section IV deals with the Proposed Methodology followed by the Data description, experimental results and discussion in section V and section VI respectively. Section VII concludes the work presented in this paper.

II. RELATED WORKS

There are several works that exist for breast cancer classification. Also,most of the works are using large number of data for their experiments. But, within the present work, we are using only half of the dataset to reduce the computation.[3] applied capsule net architecture in breast cancer classification drawback and obtained accuracy around 87%. But, they used a total of 400 images for 4 class classification problem.

The next work is finished by [4] using pre-trained Inception Resnet V2 and achieved accuracy of 76%. In [5], CNN is used for feature extraction and support vector machines for classification. They obtained accuracy around 77.8%. In [6] Classification Of Breast Microcalcification- CAD System And Performance evaluation Using SSNE. In [8], breast cancer is classification as benign and malignant based on computer aided methods applied on the cytological images of fine needle biopsies.

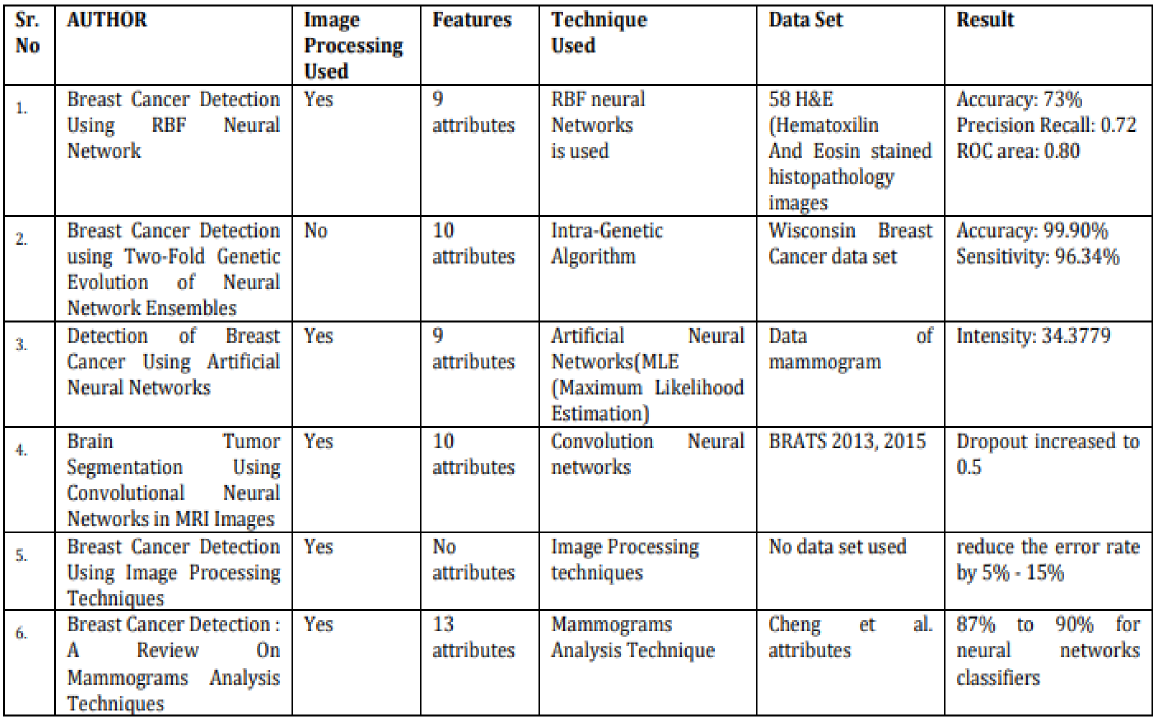
Alexander Brook [9] machine-driven the breast cancer diagnosis from microscopic biopsy images. They obtained high recognition rates by applying multi-class support vector machines on generic options, which are obtained from level set statistics of the images. In [10], skin cancer is diagnosed using deep convolutional neural networks.

In [14], Breast cancer analysis using Wavelet and Symmetric Stochastic Neighbor Embedding Based Computer Aided Analysis. In [15], Image mining for machine-driven detection of Retinal Defects.In[16] Analysis of various wavelets for brain image classification using support vector machine. In[17] The Performance analysis of the Breast Mass classification CAD System supported on DWT, SNE AND SVM.

III. LITERATURE SURVEY

- The Prediction of recurrent events in breast cancer using the Naïve Bayesian Classification paper was published by Diana Dumitru. During this paper the Naïve Bayes Classifier has been applied to Wisconsin Prognostic Breast Cancer (WPBC) dataset regarding variety of patients. The testing identification accuracy was about 74.24%.
- Breast cancer is that the most typical cancer among women everywhere the world. Every thirteen minutes a woman dies with the diagnosis of breast cancer. These facts have led to continue, learning of a way to find breast cancer in women, especially older women, which are of higher risk.
- In 2016 the Weighted Naïve Bayes Classifier: A Predictive Model for Breast Cancer Detection paper was published by Shweta Kharya and Sunita Soni. This paper was about the use of traditional naïve bayes classifier with a novel approach of weights assignment to its attributes and to develop probabilistic classifier for Breast cancer detection system that can be used by experts in decision making. The software can imitate like human diagnostic expertise for treatment of cancer alignment.

Few additional surveys:-



Sr. No	AUTHOR	Image Processing Used	Features	Technique Used	Data Set	Result
1.	Breast Cancer Detection Using RBF Neural Network	Yes	9 attributes	RBF neural Networks is used	58 H&E (Hematoxilin And Eosin stained histopathology images	Accuracy: 73% Precision Recall: 0.72 ROC area: 0.80
2.	Breast Cancer Detection using Two-Fold Genetic Evolution of Neural Network Ensembles	No	10 attributes	Intra-Genetic Algorithm	Wisconsin Breast Cancer data set	Accuracy: 99.90% Sensitivity: 96.34%
3.	Detection of Breast Cancer Using Artificial Neural Networks	Yes	9 attributes	Artificial Neural Networks(MLE (Maximum Likelihood Estimation)	Data of mammogram	Intensity: 34.3779
4.	Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images	Yes	10 attributes	Convolution Neural networks	BRATS 2013, 2015	Dropout increased to 0.5
5.	Breast Cancer Detection Using Image Processing Techniques	Yes	No attributes	Image Processing techniques	No data set used	reduce the error rate by 5% - 15%
6.	Breast Cancer Detection : A Review On Mammograms Analysis Techniques	Yes	13 attributes	Mammograms Analysis Technique	Cheng et al. attributes	87% to 90% for neural networks classifiers

IV. PROPOSED ARCHITECTURE

The aim of the currentt work is to classify breast cancer using machine learning thus because it is used as an automatic tool to help doctor's diagnosis. In the proposed method, the histology images are fed as an input to the CNN architecture.Recent results on breast cancer detection show that Convolution Neural Networks (CNN) are able to do higher recognition rates than hand-crafted feature descriptors, however the price to pay is an increase in complexity to develop the system, requiring longer training time and specific expertise to fine-tune the architecture of the CNN.

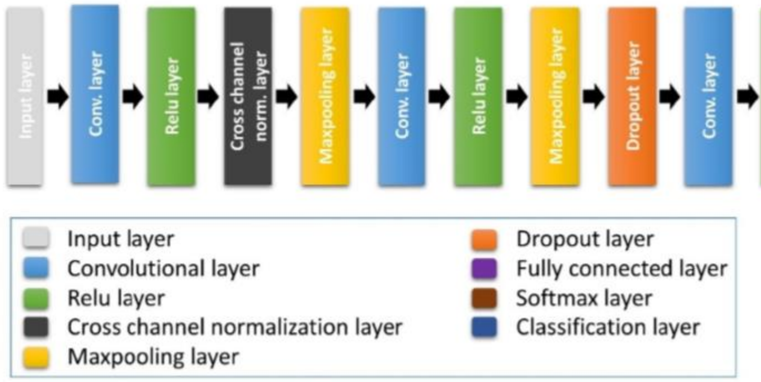


Fig. 2. The architecture of CNN.

The proposed CNN structure has 16 layers that starts from the input layer that hold the augmented images from the previous pre-processing step passing through the convolution layers and their activation functions that utilized in features choice and down-sampling (convolution, Rectified Linear Unit (ReLU), normalization and pooling layers).

To prevent overfitting, a dropout layer is used and followed by, a fully connected layer and a softmax layer to predict the output and finally a classification layer that produces the predicted class.

- Input layer - Input layer holds the augmented images.
- Convolution layer -The purpose of this layer is to receive a feature map. Usually, we have tendency to begin with low range of filters for low-level feature detection. The deeper we enter the CNN, the additional filters we use to find high-level options. Feature detection is relies on 'scanning' the input with the filter of a given size and applying matrix computations so as to derive a feature map.
- ReLu layer - Rectifier Unit, the foremost commonly deployed activation function for the outputs of the CNN neurons.
- Cross channel normalization layer - This layer performs a channel-wise local response normalization. It always follows the ReLU activation layer. This layer replaces every component with a normalized value it obtains using the elements from a explicit range of neighboring channels (elements in the normalization window).
- Max pooling layer -The goal of this layer is to provide spatial variance, that simply means that the system are going to be capable of recognizing an object even once its looks varies in a way. Pooling layer can perform a down sampling operation on the spatial dimensions (width, height), leading to output like [16x16x12] for pooling_size=(2, 2).
- Dropout layer -A dropout layer randomly sets input parts to zero with a given chance.
- Fully Connected layer - In a fully connected layer, we have tendency to flatten the output of the last convolution layer and connect each node of the present layer with the opposite nodes of consequent layer. Neurons in a fully connected layer have full connections to all or any activations within the previous layer, as seen in regular Neural Networks and add a similar way.
- Softmax layer - A softmax layer to predict the output .
- Classification layer -The classification layer produces the expected class.

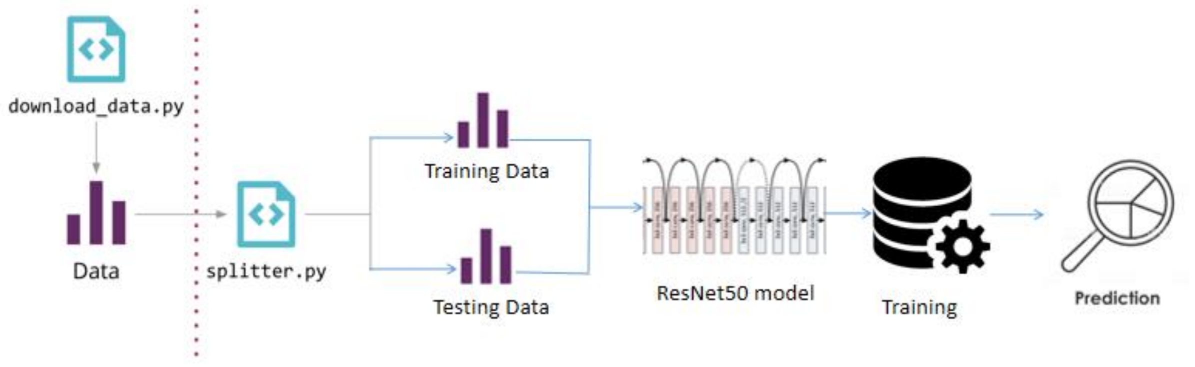


Fig. 3. Flow of proposed system.

Starting with loading all the libraries and dependencies, the images are loaded in the respective folders. After that a numpy array of zeroes is created for labeling benign images and similarly a numpy array of ones is created for labeling malignant images.Then, the dataset is shuffled and the labels are converted into categorical format. Then the data-set is split into two sets — train and test.

After performing data augmentation(The practice of data augmentation is an effective way to increase the size of the training set). The data generator gets the data from our folders and into Keras in an automated way. It provides convenient python generator functions for this purpose.

The next step is to build the model. DenseNet201 is used as the pre trained weights which is already trained in the Imagenet competition. The learning rate was chosen to be 0.0001.On top of it the globalaveragepooling layer is used followed by 50% dropouts to reduce over-fitting. Then the batch normalization and a dense layer with 2 neurons for 2 output classes is used which is benign and malignant with softmax as the activation function. The Adam is used as the optimizer and binary-cross-entropy as the loss function.

The model was trained for 20 epochs where Batch_Size is 16:-

```
history = model.fit_generator(
   train_generator.flow(x_train, y_train, batch_size=BATCH_SIZE),
   steps_per_epoch=x_train.shape[0] / BATCH_SIZE,
   epochs=20,
   validation_data=(x_val, y_val),
   callbacks=[learn_control, checkpoint]
)
```

V. DATA DESCRIPTION

The dataset BreaKHis images were collected through a clinical study from January 2014 to December 2014. All patients referred to the P&D Laboratory, Brazil[22].The dataset BreaKHis is split into two main groups: benign tumors and malignant tumors. Histological benign may be a term bearing on a lesion that doesn’t match any criteria of malignancy – e.g., marked cellular atypia, mitosis, disruption of basement membranes, metastasize, etc. Normally, benign tumors are comparatively “innocents”, presents slow growing and remains localized. Malignant tumor is a synonym for cancer: lesion can invade and destroy adjacent structures (locally invasive) and spread to distant sites (metastasize) to cause death.[22]

Table 1 Dataset Details

Cancer type	No. of Images
Benign	328
Malignant	1343

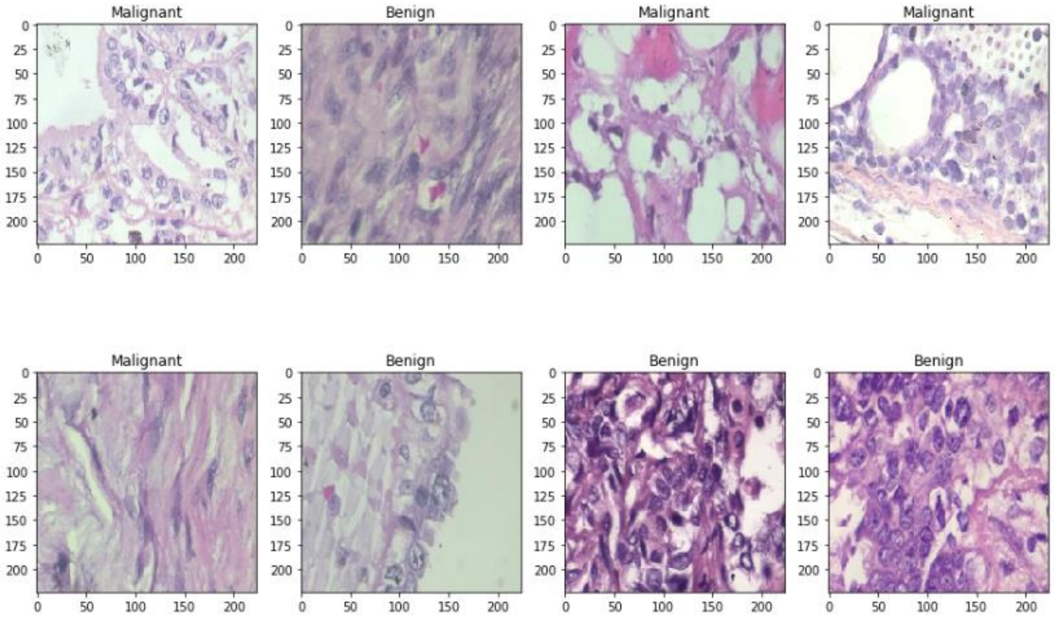


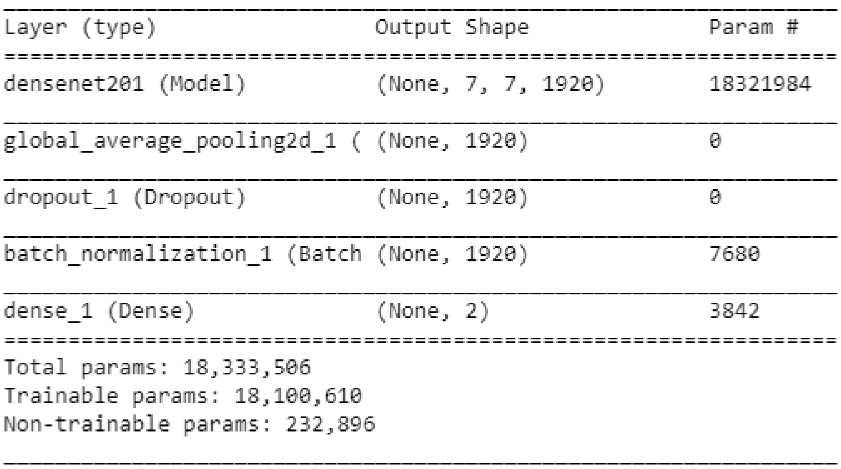
Fig .4.Benign and malignant samples

VI. EXPERIMENTAL RESULTS AND OBSERVATIONS

The Breast cancer classification model is evaluated on architecture.To get accurate results, parameter tuning was performed on the model, since theneural networks are called as parameterized function. Once the pre- process steps, the histology images are applied to the CNN architecture. The input images are flipped and zoomed to reduce the computation.

We used a batch size value of 16. Batch size is one amongst the foremost necessary hyperparameters to tune up deep learning.we value more highly to use a bigger batch size to train the models because it permits computational speedups from the parallelism of GPUs. However, it's accepted that overlarge of a batch size can lead to poor generalization. On the one extreme, using a batch up to the complete dataset guarantees convergence to the worldwide optima of the objective function. However this is at the cost of slower convergence to it optima.

On the other hand, using smaller batch sizes are shown to own quicker convergence to sensible results. This is often intuitively explained by the actual fact that smaller batch sizes enables the model to start out learning before having to envision all the data. The downside of using a smaller batch size is that the model isn't sure to converge to the global optima. There's model outline given below.



```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
densenet201 (Model)          (None, 7, 7, 1920)        18321984
_________________________________________________________________
global_average_pooling2d_1 ( (None, 1920)              0
_________________________________________________________________
dropout_1 (Dropout)          (None, 1920)              0
_________________________________________________________________
batch_normalization_1 (Batch (None, 1920)              7680
_________________________________________________________________
dense_1 (Dense)              (None, 2)                 3842
=================================================================
Total params: 18,333,506
Trainable params: 18,100,610
Non-trainable params: 232,896

_________________________________________________________________
```

Fig .5.Model summary

Precision, Recall and F1-Score

We use the following metric to urge a much better plan of true positives (TP), true negatives (TN), false positive (FP) and false negative (FN).

Precision - The ratio of correctly predicted positive observations to the overall predicted positive observations.

Recall - The ratio of correctly predicted positive observations to all or any observations in actual class.

F1-Score - The weighted average of Precision and Recall.

$$F1 = \frac{2 * (Recall * Precision)}{(Recall + Precision)}$$

The higher the F1-Score, the higher the model. For all three metric, 0 is that the worst whereas 1 is that the best.

ROC

An ROC curve is a graph showing the performance of a classification model the least bit classification thresholds. ROC can be defined as the curve between true positive rate and false positive rate. True positive refers to the positive samples that were properly labeled by the classifier. And false positive refers to the negative samples that were incorrectly labeled as positive by the classifier.

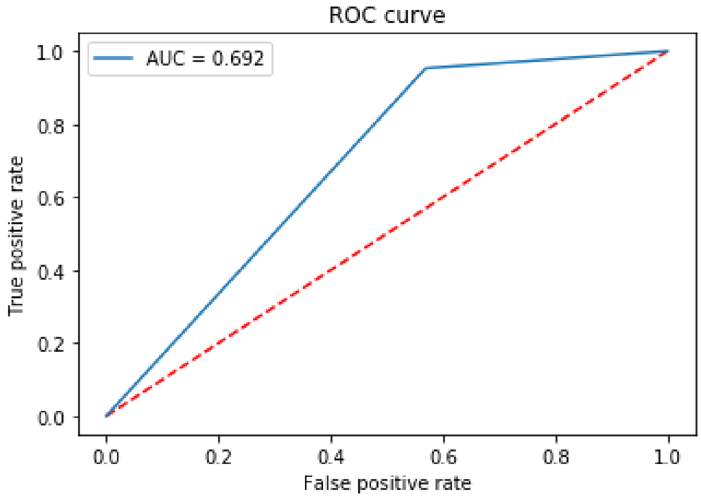


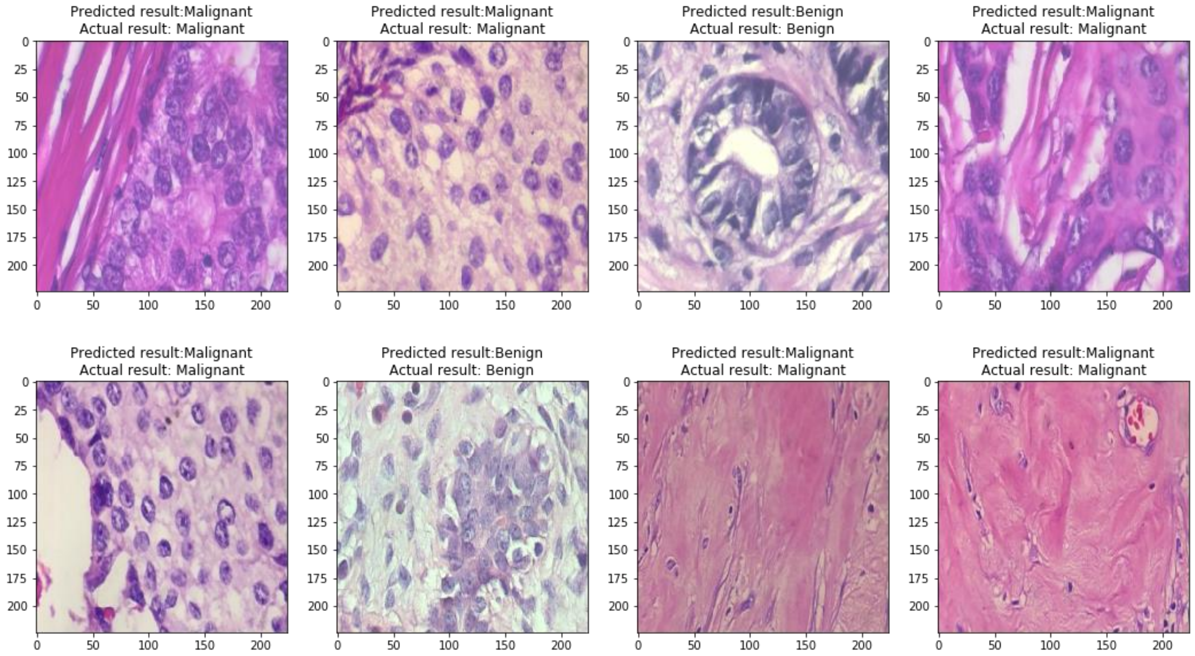
Fig. 6. Receiver operating characteristics (ROC)

The 45 degree line is that the random line, wherever the Area below the Curve or AUC is 0.5 . The more the curve from this line, the higher the AUC and better the model. The highest a model will get is an AUC of 1, wherever the curve forms a right angulate triangle. The ROC curve may also facilitate debug a model. For example, if the bottom left corner of the curve is adjacent to the random line, it implies that the model is misclassifying at Y=0. Whereas, if it's random on top right, it implies the errors occurring at Y=1.

Results

Accuracy	**Precision**	**Recall**	**F1 Score**	**ROC-AUC**
98.3%	0.65	0.95	0.77	0.692

Output Snapshots:-



VI. CONCLUSION

The present work is relies on the classification of breast cancer using CNN architecture.It is outstanding to see the success of deep learning in such varied globe issues. From this work, we've demonstrated a way to classify benign and malignant breast cancer from a collection of histological images using convolutional neural networks and transfer learning. It's clear that the performance of the conventional architectures is improved by data pre-processing and parameter tuning. The results show that this technique is used as an automated tool to help doctors in disease diagnosis, which can cause higher concentration within the treatment at early stages instead oft diagnosis and might increase the cancer survival rate.

VII. REFERENCES

[1] Siegel R. L., Miller K. D., Jemal A., "Cancer statistics", CA Cancer J Clin. 68:7- 30. W.-K. Chen, *Linear Networks and Systems*, Belmont,
CA: Wadsworth, pp. 123–135, 2018.

[2] Sabour Sara, Frosst Nicholas, E. Hinton Geoffrey, "Dynamic Routing Between Capsules", 31st Conference on Neural Information Processing Systems NIPS, 2017.

[3] Iesmantas T., Alzbutas R., "Convolutional Capsule Network for Classification of Breast Cancer Histology Images", In: Campilho A., Karray F., ter Haar Romeny B. (eds), Image Analysis and Recognition, Lecture Notes in Computer Science, vol 10882. Springer, Cham, ICIAR 2018.

[4] Ferreira C. A. et al. Classification of Breast Cancer Histology Images Through Transfer Learning Using a Pre-trained Inception Resnet V2. In: Campilho A., Karray F., ter Haar Romeny B. (eds)., Lecture Notes in Computer Science, vol 10882. Springer, Cham., Image Analysis and Recognition, ICIAR 2018.

[5] Teresa Arajo, Guilherme Aresta, Eduardo Castro, Jos Rouco, Paulo Aguiar, Catarina Eloy, Antnio Polnia, Aurlio Campilho, "Classification of breast cancer histology images using Convolutional Neural Networks", vol. 12, 2017

[6] Dr. Mohan Kumar S, Dr. Balakrishnan, Classification Of Breast Microcalcification- CAD System And Performance Evaluation Using SSNE, International Journal of Advanced Research in Computer Science and Software Engineering, Volume 5 , Issue 9, 824-830, ISSN: 2277 128X, Sep- 2015

[7] Reinhard E., Ashikigmin M., Gooch B., Shirley P., "Color Transfer between Images", IEEE Computer Graphics and Applications, pp.34-40, 2001.

[8] Filipczuk P., Fevens T., Krzyzak A., Monczak R., "Computer-aided breast cancer diagnosis based on the analysis of cytological images of fine needle biopsies", IEEE Transactions on Medical Imaging. pp.2169– 2178, 2013.

[9] Brook A., El-Yaniv R., Issler E., Kimmel R., Meir R., Peleg D., "Breast Cancer Diagnosis From Biopsy Images Using Generic Features and SVMs.", pp. 1–16, 2007.

[10] Esteva A., Kuprel B., Novoa R. A., Ko J., Swetter S. M., Blau H. M., Thrun S. "Dermatologist-level classification of skin cancer with deep neural networks", Nature, vol.542, pp.115–118, 2017.

[11] Swapna G, Soman, K. P., and Vinayakumar R., "Automated detection of diabetes using CNN and CNN-LSTM network and heart rate signals", Procedia Computer Science, vol. 132, pp. 1253-1262, 2018.

[12] Abhinav:2019, Author- Abhinav Sagar , Title- Breast-cancer-classification, Year -2019, Publisher -GitHub,Journal - GitHub repository.

[13] Dr. Mohan Kumar S & Dr. Balakrishnan, Classification Of Breast Mass Classification – CAD System And Performance Evaluation Using SSNE, IJISET – International Journal of Innovative Science, Engineering & Technology, Vol. 2, Issue 9, 417-425, ISSN 2348 – 7968

[14] Dr.S. Mohan Kumar and Dr G. Balakrishnan, Wavelet And Symmetric Stochastic Neighbor Embedding Based Computer Aided Analysis For Breast Cancer, Indian Journal of Science and Technology ISSN 0974-6846 and 0974-5645(Print&Online), Volume 9, Issue 47, 12-16

[15] Dr. S. Mohan Kumar & Anisha Rebinth, Automated detection of Retinal Defects using image mining, A review, European Journal of Biomedical and Pharmatical Sciences, European ISSN : 2349 – 8870, Volume 5 , Issue : 01 year : 2018, pp No.: 189 – 194

[16] S Mohan kumar, Analysis of different wavelets for brain image classification using support vector machine, International Journal of Advances in Signal and Image Sciences Volume 2, Issue 1, 1-4

[17] S Mohan Kumar & Dr. Balakrishnan, The Performance Evaluation of the Breast Mass classification CAD System Based on DWT, SNE AND SVM , International Journal of Emerging Technology and Advanced Engineering, 2013, ISSN 2250–2459, Volume 3, Issue

10, October 2013, Page Numbers: 581-587

[18] S Mohan Kumar & Dr. Balakrishnan, Classification of Micro Calcification And Categorization Of Breast Abnormalities - Benign and Malignant In Digital Mammograms Using SNE And DWT, Karpagam Journal of Computer Science 2013, ISSN No: 0973-2926, Volume-07, Issue-05, July-Aug, 2013. Page Numbers: 253 to 259

[19] S Mohan Kumar & Dr. Balakrishnan, Categorization of Benign And Malignant Digital Mammograms Using Mass Classification – SNE and DWT, Karpagam Journal of Computer Science, 2013, ISSN No: 0973-2926, Volume-07, Issue-04, June-July-2013, Numbers: 237-243.

[20] S Mohan Kumar & Dr. Balakrishnan, Classification of Microclacification in digital mammogram using SNE and KNN classifier, International Journal of Computer Applications - Conference Proceedings published in IJCA, 2013 ISBN: 973-93-80872-00-6, ICETT proceedings with IJCA on January 03,2013, Page Numbers: 05-09

[21] S Mohan Kumar & Dr. Balakrishnan, Mutiresolution analysis for mass classification in Digital Mammogram using SNE, IEEE international Conference- ICCSP-13 organized by Athiparasakthi Engineering College, Chennai , 2013, ISBN:978-1-4673-4864-5, Page Numbers: 2041-2045.

[22] Spanhol, F., Oliveira, L. S., Petitjean, C., Heutte, L., A Dataset for Breast Cancer Histopathological Image Classification, IEEE Transactions on Biomedical Engineering (TBME), 63(7):1455-1462, 2016.