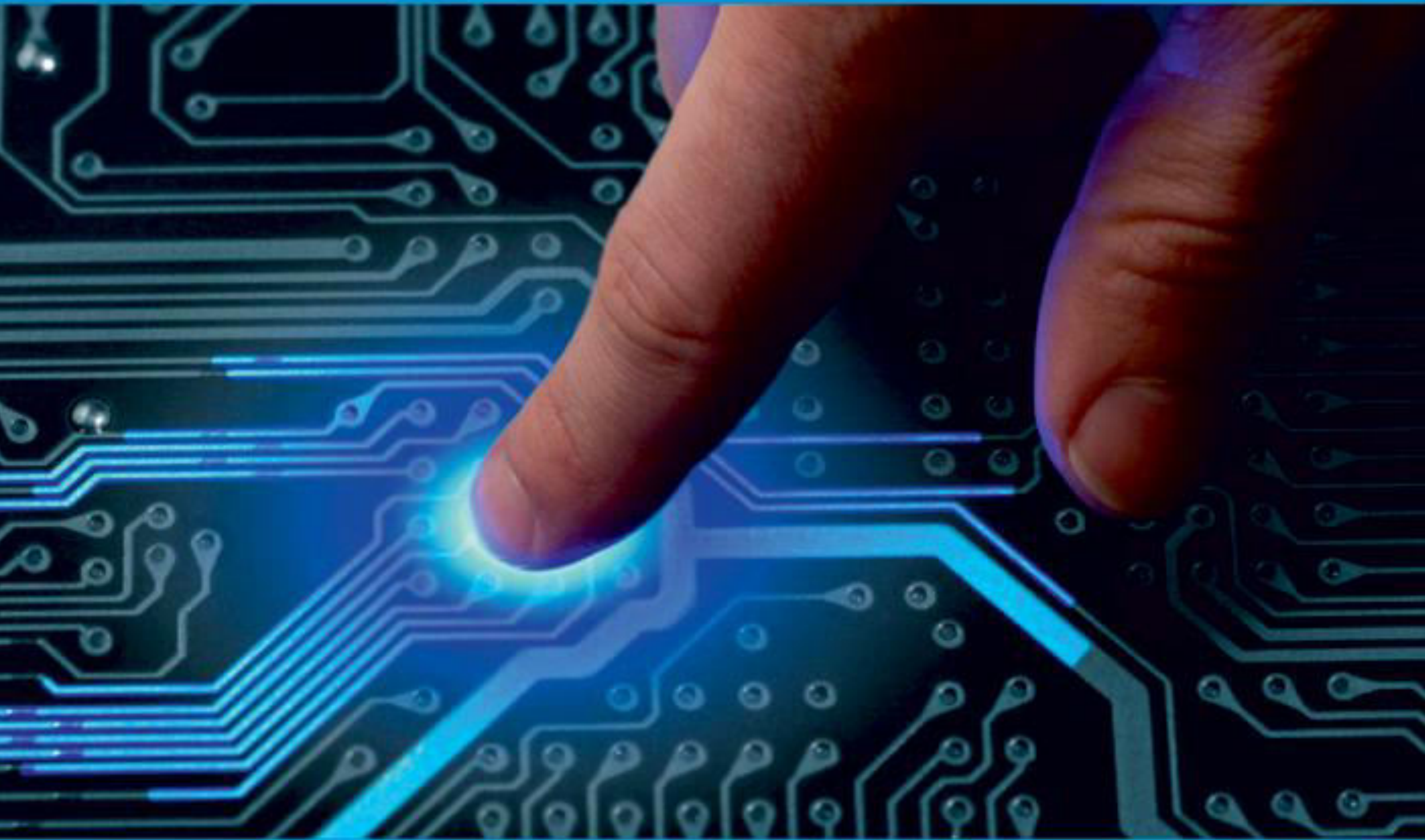
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 2, February 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 2, February 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1202006 |

Breast Cancer Classification Using Transfer Learning with Ensemble

Mrs. Preethi Kolluru Ramanaiah[1], Dr. Ramanaiah[2]

Cloud Architect, Ernst and Young, New York, USA[1]

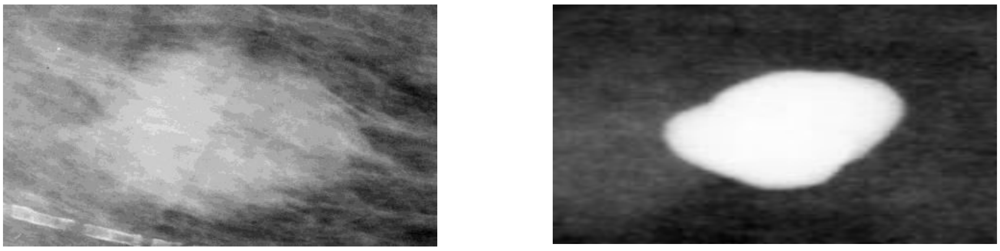
Head of Department of Mathematics (Retd), Jamal Mohammed College, Trichy, India[2]

ABSTRACT*:* Breast cancer (BC) is one of the leading causes of cancer death in women. By allowing patients to receive appropriate care at an early stage, BC can improve their chances of survival. This study develops a new deep learning (DL) model based on the transfer learning (TL) technique to aid in the automatic detection and diagnosis of the suspected BC. The proposed system has two techniques namely CNN with Gabor Channel and VGG16. Problem-specific model is used to create DL architectures. TL applies the skills learned from resolving a specific problem to another that is pertinent. The features of the proposed model are based on the CBIS-DDSM (Curated Breast Imaging Subset of DDSM). Mammographic picture dataset with Structure of Convolutional neural networks (CNNs) and Visual Geometry Group (VGG)-16 networks. Breast Cancer classification whether the tumor is Benign or Malignant is classified using both models. Four assessment metrics—F1-Score, Recall, Support, and Accuracy—have been selected to evaluate how well the suggested model performs. According to experimental results, CNN using the Gabor Filter model provides 96% accuracy, while the TL of the VGG16 model gives 70% accuracy.

KEYWORDS: Breast Cancer, Deep Learning, CNN, VGG16, Gabor Filter

I. INTRODUCTION

In the human body, cancers come in two distinct assortments as shown in Figure 1. Non-harmful cells that develop locally and don't spread across the body make up a harmless cancer. A threatening growth is comprised of dangerous cells that spread to the tissues around them. All through their lives, 12% of Indian ladies will encounter BC. In India, a lady is determined to have BC like clockwork [1], [2]. For ladies, this implies that BC is the most widely recognized sort of disease [3]. The unusual development of bosom cells is the reason for bosom disease. It relies upon which cells form into harmful ones. Anyplace in the bosom BC can shape. The curves, channels, and connective tissue are the three fundamental pieces of the bosom. The conduits or lobules are where most BC creates. Early BC ID is consequently significant to raising patient endurance rates. Since malignant growth is connected to a high bleakness rate and critical medical care costs, specialists have been looking for additional precise models to analyze the sickness. A biopsy and mammography is one of the most famous ways of tracking down bosom disease. Research demonstrates that the utilization of mammography has diminished the quantity of passings inferable from bosom disease. Mammography distinguishes early signs of bosom disease in ladies by utilizing a specific sort of bosom picture. A biopsy can likewise be completed for BC.



Malignant ***Benign***

Fig 1: Breast Cancer Dataset Sample Images

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 2, February 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1202006 |

II. RELATED WORK

F. Ting et al. [1], in a profound brain network were carried out for the grouping of BC sores. There was one information layer, 28 secret layers, and one result layer in this organization. The component wise-information expansion (FWDA) method forestalled overfitting for responsiveness, exactness, and explicitness, the relating values were continuously achieved: 89.47%, 90.5%, and 99.71% of their proposed approach. The BreastNet was proposed by To§çar et al. [2] and could extricate the best qualities from bosom pictures. It was made out of convolutional, pooling, leftover, and thick blocks. With 98.80% precision, BreastNet outflanked AlexNet, VGG-16, and VGG-19 models.A multi-facet DL engineering was introduced by Abbas [03] to order harmless and destructive regions in bosom pictures. The invariant elements in this organization were extricated in four phases and afterward changed over into profound invariant highlights and learning highlights for direction. Utilizing the MIAS dataset, [3] got aftereffects of 92%, 84.2%, and 91.5%, with awareness, explicitness, exactness, and AUC all approaching in at 0.91. Sha et al. [4] detailed a strategy for consequently recognizing and sorting the threatening region in bosom pictures utilizing the equivalent dataset. Their proposed approach utilized both a grasshopper streamlining calculation and a calculation in light of CNN. The discoveries showed the way that the recommended approach could accomplish responsiveness, explicitness, and precision of 96%, 93%, and 92%, separately. A CNN was prepared by Charan et al. [5] to identify BC. Three completely associated layers (FCLs), four normal pooling layers, and six convolution layers made up their proposed CNN. The information image had a 224 × 224 element, and the characterisation results were applied using the Softmax (SM) capability. Using the MIAS data set, the organization's average is not set in stone at 65%. Wahab et al. used a pre-prepared CNN to order mitoses and then transferred the resulting boundaries to another CNN in [6]. The accuracy, review, and F-proportion of the recommended approach were, individually, 0.50, 0.80, and 0.621. Additionally, to order BC in many classes, Lotter et al.

III. PROPOSED WORK

As in Figure 2 the proposed technique for BC recognition and grouping consists of two basic aspects. CNN grouping is the subsequent part, and information preprocessing is the utilization of the principal part. Likewise BC arrangement finished with move learning model VGG16, CNN order with Gabor Channel is performing great contrast and VGG16.

Fig 2: CNN with Gabor Filter work flow diagram

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

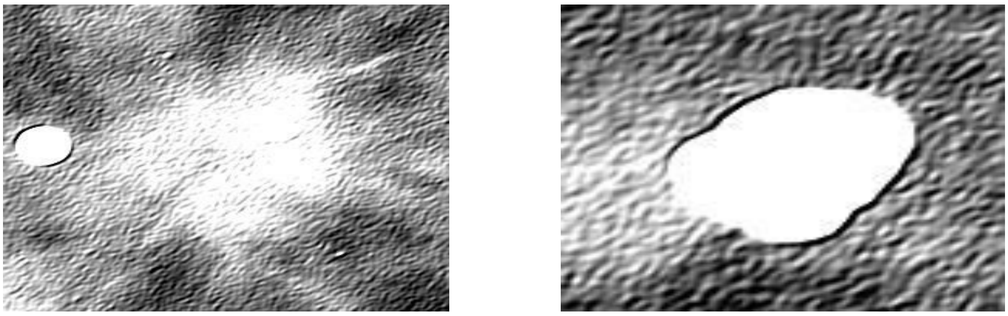
|| Volume 12, Issue 2, February 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1202006 |

This system has three main algorithms which are explained below.

A. Gabor Filter

As in Figure 3, an image control The Gabor channel, so named to pay tribute to Dennis Gabor, is a direct channel utilized for surface examination in pictures. Set forth plainly, it look through a bound district encompassing the examination point for specific recurrence content in the picture specifically headings. Numerous vision researchers guarantee that Gabor channels' portrayals of heading and recurrence are tantamount to those of the human visual framework. They are especially helpful for portrayal and segregation of surfaces. A sinusoidal plane wave balancing a Gaussian part capability in the spatial space makes up a 2-D Gabor channel.



Malignant ***Benign***

Fig3: After Gabor Filter

B. Convolution Neural Network

In 1988, Yann LeCun made a specific sort of fake brain network called the convolutional brain organization (CNN). CNN utilizes specific elements of the visual cerebrum. Classification of pictures is one of the most well known utilizes for this engineering. For example, Facebook involves Google for client photograph look, Amazon for item proposals, and CNN for programmed labeling calculations.

We should investigate how CNN can be utilized to arrange pictures. The basic undertaking of picture characterization is to acknowledge the information picture and afterward characterize its class. People have this capacity from birth, which empowers them to immediately recognize an elephant in a photo. Conversely, the PC sees the pictures in a totally unique manner.

As shown in the Figure 4 the convolution layer is consistently the top layer. It is given the picture, which is a pixel-esteem network. Accept that the upper left corner of the picture is where the info lattice is being perused. Then, a more modest grid called a channel (or neuron, or center) is picked by the product to be embedded there. From that point onward, the channel executes convolution, pushing ahead with the info picture. Duplicating the channel's qualities by the first pixel values is its work. These increases amount to an aggregate. In the long run, only one number remaining parts. The channel continues one unit to the right, doing the indistinguishable capability, since it has just perused the picture in the upper left corner. The all out of these augmentations is added.

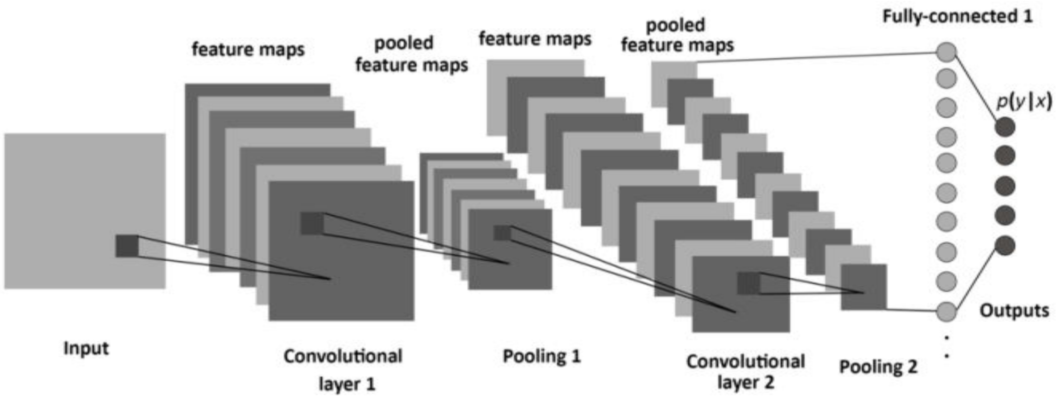


Fig 4: Architecture of CNN

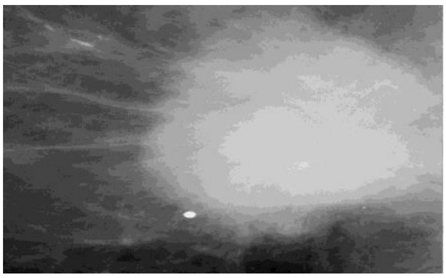
This capability is equivalent to human impression of straightforward varieties and visual lines. In any case, the acknowledgment of more significant level elements, similar to the storage compartment or huge ears, requires the total organization. In the wake of finishing a progression of convolutional, nonlinear, and pooling layers, joining a totally associated layer is significant. This layer gets input from convolutional organizations' result information. An N-layered vector is made when a completely associated layer is connected to the organization's end; N is the number of classes from which the model selects the predicted class.

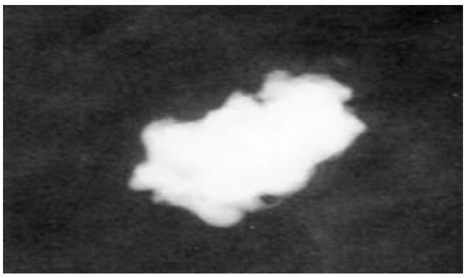
C. VGG16 (VGG - Visual Geometry Group)

In the 2014 ILSVR (Imagenet) rivalry, the convolutional brain net (CNN) design VGG16 won in front of the pack. It is regarded as one of the best designs for vision models created. VGG16 often employs a similar cushioning and max pool layer of a 2x2 channel with a step 2, and on second thought, it focuses on 3x3 channel convolution layers with a step one. Throughout the engineering process, the convolution and max pool layers follow a predictable pattern. Its result is involved two FC (completely associated layers) and a Softmax. The number 16 in VGG16 alludes to the way that it contains sixteen weighted layers. With an expected 138 million boundaries, this organization is genuinely enormous.

D. Dataset

An improved and normalized adaptation of the Computerized Data set for Screening Mammography is this CBIS-DDSM (Arranged Bosom Imaging Subset of DDSM) version (DDSM). With 2,620 investigations, the DDSM is a data set of checked film mammography pictures. It contains pathology information along with instances of harmless, dangerous, and typical patients. The DDSM ends up being an important instrument in the creation and assessment of choice emotionally supportive networks because of its broad degree and capacity to approve ground truth. A certified mammographer picked and assessed a piece of the DDSM information for the CBIS-DDSM gathering. The pictures were de-pressurizeed and switched over completely to the DICOM design.





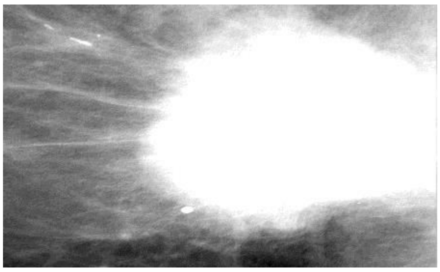
Malignant ***Benign***

Fig 6: Original Dataset Image

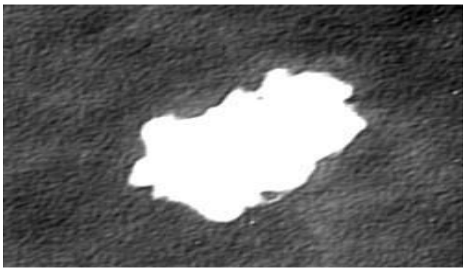
| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 2, February 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1202006 |



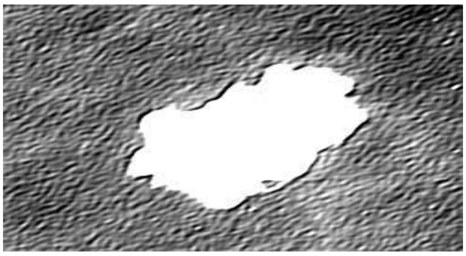
Malignant



Benign

Fig 7: Gabor Filter1 Dataset Image

Malignant



Benign

Fig 8: Gabor Filter2 Dataset Image

IV. RESULT ANALYSIS

This study compares the performance of CNN with Gabor Channel and VGG16 in breast cancer classification, revealing superior accuracy, recall, and F1-Score for CNN with Gabor Channel.

Table 1: comparison between CNN and VGG16

Model	**Accuracy**	**Recall**	**F1-Score**	**Support**
CNN with Gabor Filter	96%	71%	82%	419
VGG16	70%	65%	65%	290

Comparison

As shown in Table 1 the CNN calculation with Gabor Channel had the option to characterize bosom disease more accurately than the other two AI calculations that were surveyed, the exactness is 96% and VGG16 bosom malignant growth order has 70% accuracy.CNN with Gabor Channel has less misclassification blunders than the VGG16. Contrasting CNN model and VGG16, CNN model give more exactness with Gabor Channel dataset. CNN model Review score is 71% where as VGG16 is 65%, F1 Score of CNN is 82% where as VGG16 is 65% and Backing score of CNN is 419 where as VGG16 is 290.From this examination we can plainly comprehend in every one of the measurements CNN is performing great contrast with VGG16.

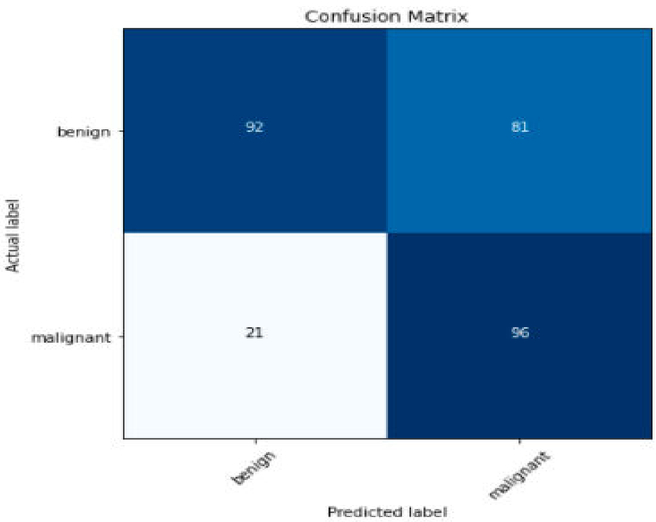
| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

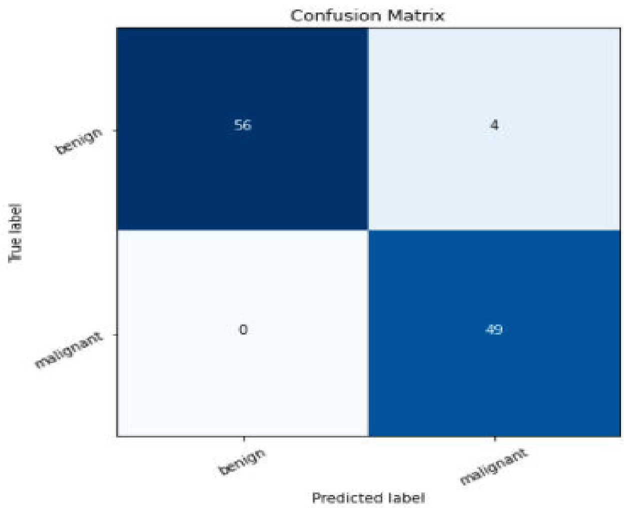
|| Volume 12, Issue 2, February 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1202006 |

Confusion Matrix

The confusion matrix, illustrated in Figures 9 and 10, provides a snapshot of a classification model's performance, illustrating the counts of true positives, false positives, true negatives, and false negatives. It serves as a basis for calculating metrics such as accuracy, precision, recall, and F1 score. A disorder framework is a commonly used table that assesses how well an order model (also known as a "classifier") performs on a collection of test data for which the true attributes are known.





ROC Curve

The ROC curves in Figures 11 and 12 depict the performance of the classification models across different classification thresholds. The ROC curve is a visual representation of the trade-off between sensitivity and specificity. These curves help in understanding how well the models perform at various levels of classification. The collector working trademark bend, or ROC bend, is a diagram that shows a characterization model's presentation over all grouping levels. Two boundaries are plotted on this bend.

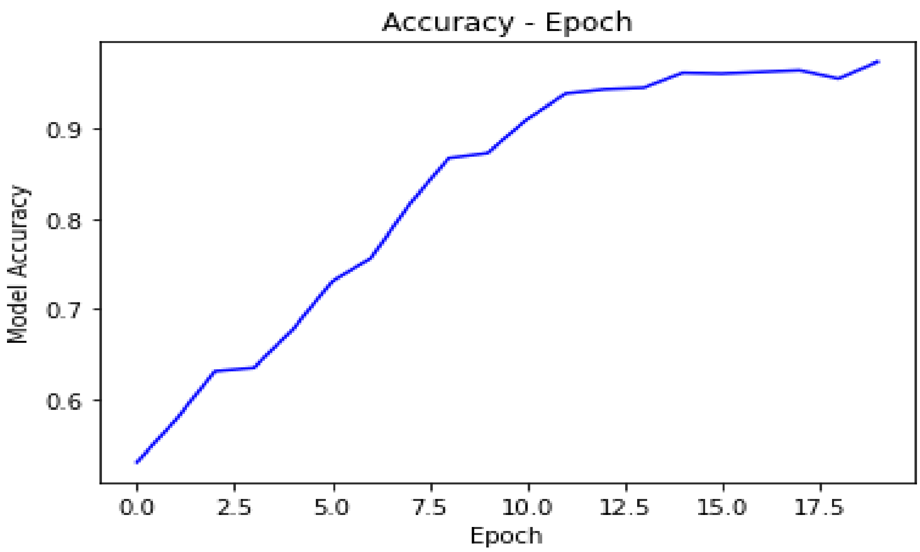


Fig11: ROC Accuracy Curve

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 2, February 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1202006 |

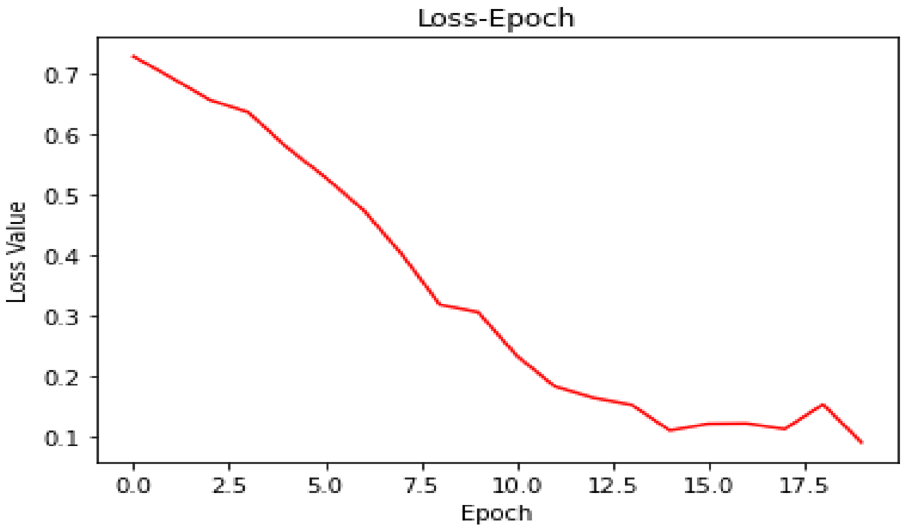


Fig12: ROC Loss Curve

V. CONCLUSION

The forecast discoveries show the viability of ML calculations in characterizing bosom disease into triple negative and non-triple negative subtypes. We applied our two essential calculations, CNN and VGG16, to the CBIS-DDSM (Organized Bosom Imaging Subset of DDSM) Bosom Disease datasets in light of the fact that the objective of bosom malignant growth characterization is to foster precise and reliable classifiers. Subsequent to looking at our calculations precisely, we found that CNN performed all the more proficiently, accomplishing 96%; notwithstanding, even VGG16 performs well, with 70% exactness; if the dataset is bigger, the CNNs might perform significantly more precisely.

REFERENCES

1. F. Ting, Y. J. Tan, and K. S. Sim, "Convolutional neural network improvement for breast cancer classification," Expert Syst. Appl.,vol. 120, pp. 103_115, Apr. 2019.
2. M. To§açar, K. B. Özkurt, B. Ergen, and Z. Cömert, "BreastNet: A novel convolutional neural network model through histopathological images forthe diagnosis of breast cancer," Phys. A, Stat. Mech. Appl., vol. 545,May 2020, Art. no. 123592, doi: 10.1016/j.physa.2019.123592.
3. Q. Abbas, "DeepCAD: A computer-aided diagnosis system for mammographicmasses using deep invariant features," Computers, vol. 5, no. 4,p. 28, Oct. 2016, doi: 10.3390/computers5040028.
4. Z. Sha, L. Hu, and B. D. Rouyendegh, "Deep learning and optimization algorithms for automatic breast cancer detection," Int. J. Imag. Syst.Technol., vol. 30, no. 2, pp. 495_506, Jun. 2020, doi: 10.1002/ima.22400.
5. S. Charan, M. J. Khan, and K. Khurshid, "Breast cancer detectionin mammograms using convolutional neural network," in Proc. Int.Conf. Comput., Math. Eng. Technol. (iCoMET), Mar. 2018, pp. 1_5, doi:10.1109/icomet.2018.8346384.
6. N. Wahab, A. Khan, and Y. S. Lee, "Transfer learning based deep CNN for segmentation and detection of mitoses in breast cancer histopathologicalimages," Microscopy, vol. 68, no. 3, pp. 216_233, Jun. 2019, doi:10.1093/jmicro/dfz002.[07] W. Zhu, C. Ma, X. Zhao, M. Wang, A. A. Heidari, H. Chen, and C. Li,``Evaluation of sino foreign cooperative education project using orthogonalsinecosine optimized kernel extreme learning machine," IEEEAccess, vol. 8, pp. 61107_61123, 2020.

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING