



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

# IJIRSET

International Journal of Innovative Research in  
**SCIENCE | ENGINEERING | TECHNOLOGY**

# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 12, Issue 11, November 2023

**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA

Impact Factor: 8.423

9940 572 462

6381 907 438

[ijirset@gmail.com](mailto:ijirset@gmail.com)

[www.ijirset.com](http://www.ijirset.com)

# A Novel Method for Detection of Breast Cancer using Artificial Intelligence Algorithms on CBIS DDSM DATABASE

A Malarvizhi<sup>1</sup>, A Nagappan<sup>2</sup>

Research Scholar and Assistant Professor of Electronics and Communication Engineering, Vinayaka Mission's Kirupananda Variyar Engineering College, Vinayaka Mission's Research Foundation (Deemed to be University), Salem, Tamilnadu, India<sup>1</sup>

Registrar, Vinayaka Mission's Research Foundation (Deemed to be University), Salem, Tamilnadu, India.<sup>2</sup>

**ABSTRACT:** Deep Neural Networks for the Identification of Carcinogenic Tissue in Mammograms the number of cases of breast cancer has been rising over time. Early diagnosis and treatment of lesions is crucial for increasing the likelihood of a favorable outcome. Mammograms and other screening techniques are among the best ways to accomplish this. However, because of their morphological makeup, certain lesions can be difficult to recognize. For example, the diameters of calcifications are usually modest, ranging from 0.1 to 1 mm. Radiologists must carefully and thoroughly analyze the exams in order to find calcifications. This article suggests a method that uses deep Convolutional neural networks to help detect malignant calcifications in mammograms. The deep Convolutional neural network divides the mammography image into patches that indicate whether or not they contain malignant lesions (cancerous tissue) or not (healthy tissue). After that, a summary of the patch-wise cancer probability is obtained in order to make a judgment on the entire mammography. In addition, a heatmap is created that shows areas that are likely to contain malignant tissue. Radiologists can utilize this heatmap to highlight particular locations. The top model assessed in the classification task had an F-Measure of 0.99, according to experiments conducted with the CBIS-DDSM dataset. These findings suggest that implementing the approach in a practical setting may enhance the identification of cancer cases.

**KEYWORDS:-** breast cancer ,CBIS-DDSM, Convolutional neural networks (CNNs), Mammograms

## I.INTRODUCTION

Meta-learning-Based Classification of Breast Abnormalities on Screening Mammograms the diagnostic classification and the breast abnormality classification are the two processes in the detection of general breast cancer. Computational technology can facilitate the process of determining the abnormality, which is crucial for the subsequent phases. To automate the detection, numerous machine learning techniques have been used. The majority of them, meanwhile, concentrate on the diagnostic categorization, with little emphasis placed on the classification of breast abnormalities. Many machine learning algorithms are also hindered by the small size of available mammography datasets. We suggested a meta-learning-based breast abnormality classification approach, taking into account the significance of the categorization of breast abnormalities and the lack of publicly available large-scale medical datasets. Our model was inspired by the most recent meta-learning-based image classifier research and made several modifications. In particular, we retrained a pretrained embedding neural network using the concept of meta-learning to adjust its feature extraction performance for the CBIS-DDSM dataset [1]. The dataset consists of two types of abnormal breast scans: calcification and mass, and each type is composed of two medical imaging categories: ROI and complete mammograms [2]. The inadequate training sample issue was resolved with the use of data augmentation techniques and the concept of meta-learning, which resulted in a final accuracy of 76%, surpassing the accuracy of a neural network baseline model, which was 71%.

## II.CBIS-DDSM

A deep learning method with great specificity that focuses on breast cancer screening the most frequent cancer in women and the second most prevalent cancer in general is breast cancer. With hundreds of breast screening tests conducted every day all around the world, radiologists have a time-consuming work ahead of them when classifying and analyzing all the exams, the majority of which turn out to be normal instances[3]. Reviewing more complicated cases that need more attention could take up the time allotted to negative case reviews. The authors of this work suggest a diagnosis support system that focuses on mammography screening[4]. It is built on an ensemble classifier and deep convolutional neural networks (CNNs), and it is intended to relieve radiologists of routine cases.

To produce an image-level classification, the architecture leverages patient metadata, patch-level features, and macroscopic complete mammogram-level characteristics [5]. The findings indicate that the architectures have a high degree of confidence that the exam is either normal, freeing the radiologist from further examination, or questionable, necessitating the attention of an expert radiologist for a final classification. Using three publicly available datasets (CBIS-DDSM, BCDR, and IN breast), the created system was trained and validated, ultimately achieving an exam-level AUC of 0.98, along with 96% specificity and 87% sensitivity[6]. As a result of this work, the workload of radiologists may be reduced, thereby eliminating the need for a second reader and opening up more time for the evaluation and analysis of more complicated cases.

## III.BREAST CANCER CATEGORIZATION

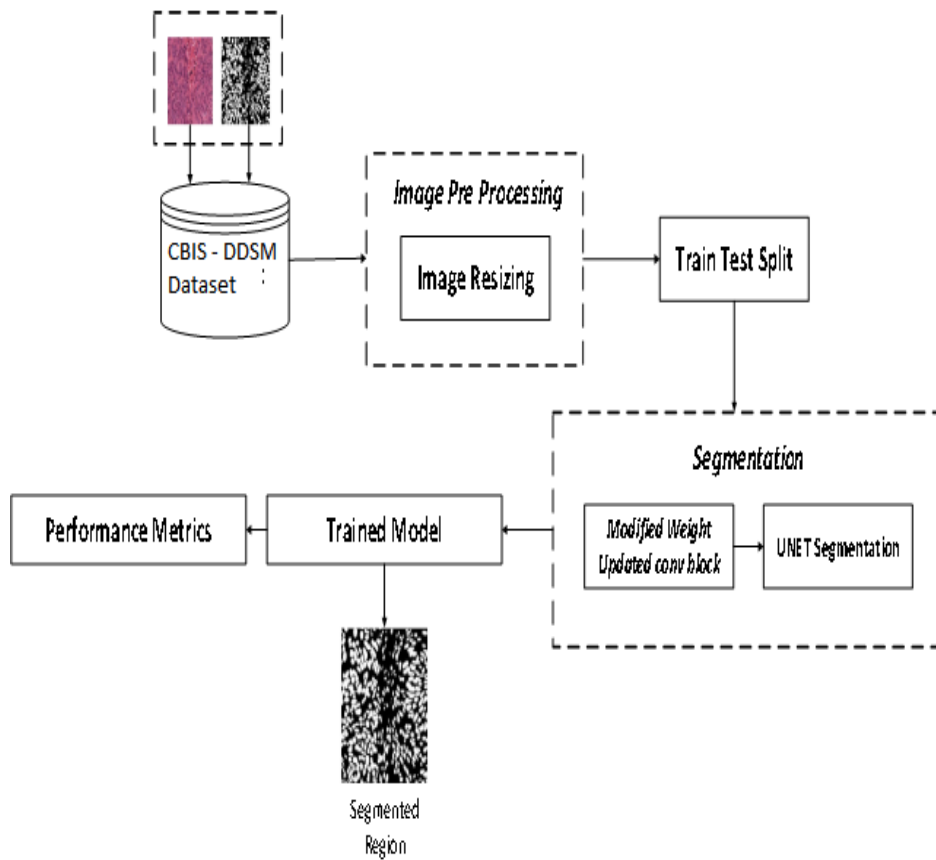
Elemental Elements in More In-Depth Deep Learning Layers for breast cancer categorization treating breast cancer, early detection is more important than cure in subsequent mammography screening procedures. Researchers have long suggested using artificial intelligence to create a computer-aided diagnosis (CAD) that can identify the type of breast tumor lesion, whether benign or malignant. This strategy is important for reducing the number of radiologists who misread false-positive and false-negative diagnosis outcomes. It is now primarily due to a lack of publicly available, well-established medical datasets that the system is not yet extensively adopted in clinical settings[7]. The purpose of this work is to examine how well a convolutional neural network (CNN) performs in identifying different types of malignant lesions[8]. The pre-trained CNN networks are evaluated using CBIS-DDSM and INbreast, two well-known public datasets. To reduce the impact of over fitting, preprocessing techniques such as denoising and contrast limited adaptive histogram equalization (CLAHE) and augmentation were used[9]. The transfer learning technique was used to train the pre-trained CNNs AlexNet and InceptionV3, which stand for shallow and deeper neural networks, respectively. The system's accuracy, losses, sensitivity, specificity, and receiver operating characteristic curve (ROC) are assessed after testing and performance testing[10]. Using the IN breast dataset, the InceptionV3 network outperforms the shallower AlexNet network with the maximum testing and area under the curve (AUC) of 99.93% compared to 98.92%. Testing accuracy is demonstrated to increase with augmented data training, reaching 86.7% from 60.26% with a non-augmented dataset in low-quality input photos. Conversely, transfer learning with a shorter network yields great accuracy outcomes without sacrificing computational expense. By changing the pretrained networks that are employed and obtaining distinct characteristics from each Convolutional layer that will be trained later, this study provides the foundation for future improvements in the system's performance [11].

## IV. APPLICATIONS OF ARTIFICIAL INTELLIGENCE IN BREAST CANCER RESEARCH

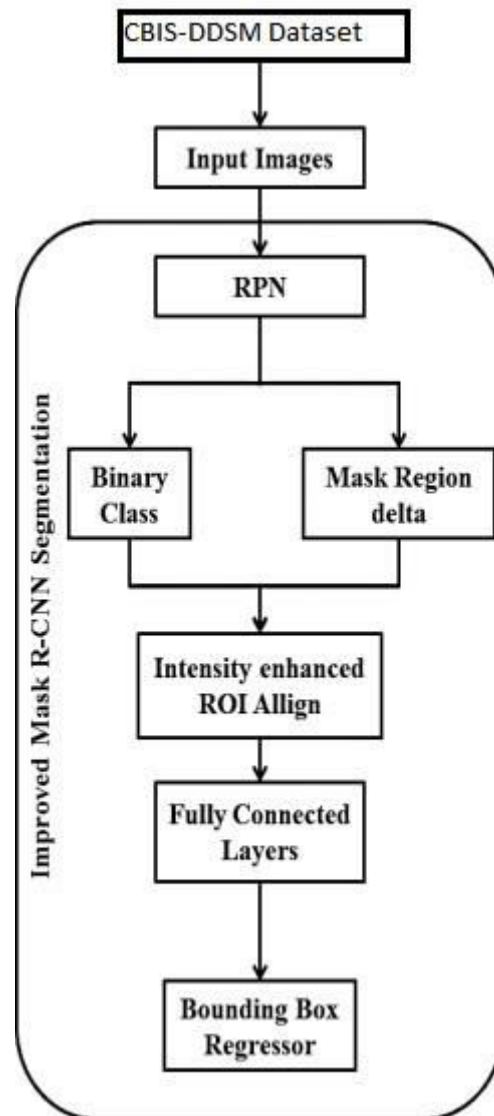
This paper's goal is to develop and put into practice an algorithm that uses a mammography dataset to automatically generate photo acoustic numerical models. The algorithm's steps are displayed in the flowchart in Figure 1 and 2 below. First, various breast tissues are removed using various extraction algorithms. Once this is done, the extracted tissues are combined with new images using a mathematical set operation. Lastly, photo acoustic parameters are assigned to various tissues in new image modes according to their unique optical and acoustic characteristics.

**V. FLOWCHART OF ALGORITHM FOR AUTOMATICALLY GENERATING PHOTO ACOUSTIC NUMERICAL MODELS**

10,239 processed mammograms are included in the CBIS-DDSM (Curated Breast Imaging Subset of DDSM) dataset, which is an improved version of the DDSM (Digital Database for Screening Mammography) [12]. We preprocess the dataset by data augmentation to improve the algorithm's implementation. Specifically, fibroglandular tissue which serves as the input for a Convolutional neural network was manually cropped under the supervision of clinicians[13].







One of the most prevalent malignancies and a major contributor to cancer-related mortality among women globally is breast cancer[14]. A good prognosis depends on early diagnosis and treatment. Artificial intelligence technology is being used in the medical profession in more and more ways, such as image analysis, automated diagnosis, intelligent pharmaceutical systems, customized treatment plans, and more[15]. In addition to lessening the burden of physicians, AI-based technologies for adjuvant therapy, pathology, and breast cancer imaging continuously increase the sensitivity and accuracy of breast cancer diagnosis and treatment[16]. In order to generate ideas for future study, this paper analyzes the use of AI in breast cancer treatment and looks ahead, posing difficulties to the development of AI in breast cancer detection and therapy.

## VI. ABRIDGED TERMS

AUC is for area under the receiver operating characteristic curve. Figure 3 and 4 AI stands for artificial intelligence. Breast Imaging Reporting and Data System, or BI-RADS Digital breast tomosynthesis, or DBT.Lower recall rates and fewer exams in the simulated workflow were the results of an artificial intelligence system that weeded out digital breast tomosynthesis examinations that did not show malignancy.Key Findings N An artificial intelligence (AI) model maintained non inferior sensitivity (90.0% vs. 90.8%; P,.001) and reduced screening effort by 39.6% in a retrospective

analysis of 5182 digital breast tomosynthesis screening exams. The AI model eliminated cancer-free exams in a simulation, which also resulted in a 25% drop in the number of women who would have been called back (6.9% vs. 9.2%; P,.001). Classification of Breast Cancer Using Deep Learning Features Enhanced with Specialized Features one of the main causes of death for women worldwide is breast cancer. Advanced stages are tough to treat, yet millions of women's lives are improved and survival rates are greatly increased with early identification. It is critical that the scientific community develop a paradigm for early detection, classification, and diagnosis of breast cancer given the disease's high prevalence. These kinds of frameworks are being developed by the artificial intelligence research community in collaboration with medical practitioners to automate the detection task. Large datasets, increased processing power, and a spike in research efforts all point to the possibility that the outcomes of AI frameworks may assist even more clinicians in generating accurate predictions. This paper proposes a unique framework for mammography-based breast cancer classification. The suggested framework blends manually created features like HOG (Histogram of Oriented Gradients) and LBP (Local Binary Pattern) with robust features taken from innovative Convolutional Neural Network (CNN) features. The CBIS-DDSM dataset yielded results that surpass the current state of the art.

## VII. ARTIFICIAL INTELLIGENCE

It is undeniable that the field of artificial intelligence (AI) has had an impact on numerous fields, including medicine. Particularly, there is a boom in artificial intelligence (AI) systems that examine medical photos to identify a variety of illnesses, including breast cancer. This article introduces a revolutionary system designed to analyze mammograms in order to diagnose breast cancer. First, we used transfer learning on the most advanced CNN architectures to obtain baseline findings. We developed a unique CNN architecture, called cVGG, to address the shortcomings of the transfer learning approach. The goal of training this innovative CNN architecture on mammograms was to categorize them as benign or malignant.

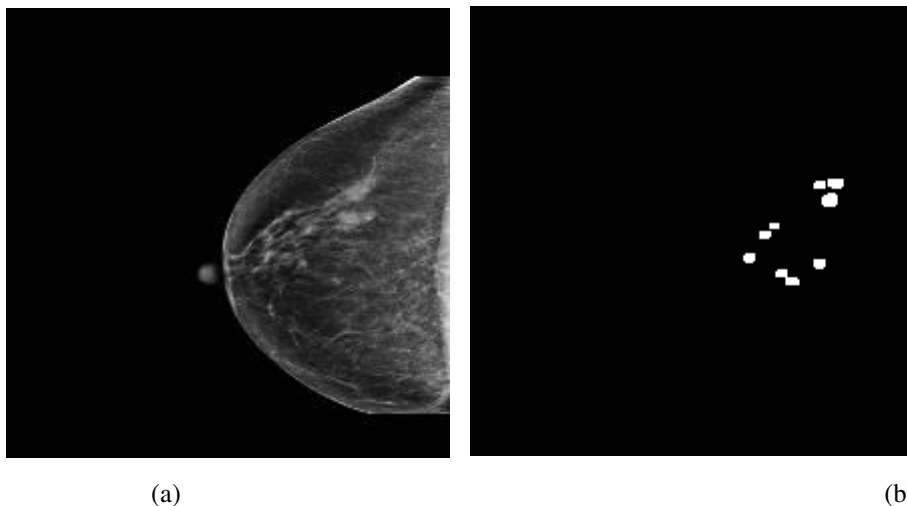


Figure 3 a & b : Improved mask rcnn segmentation input & output image

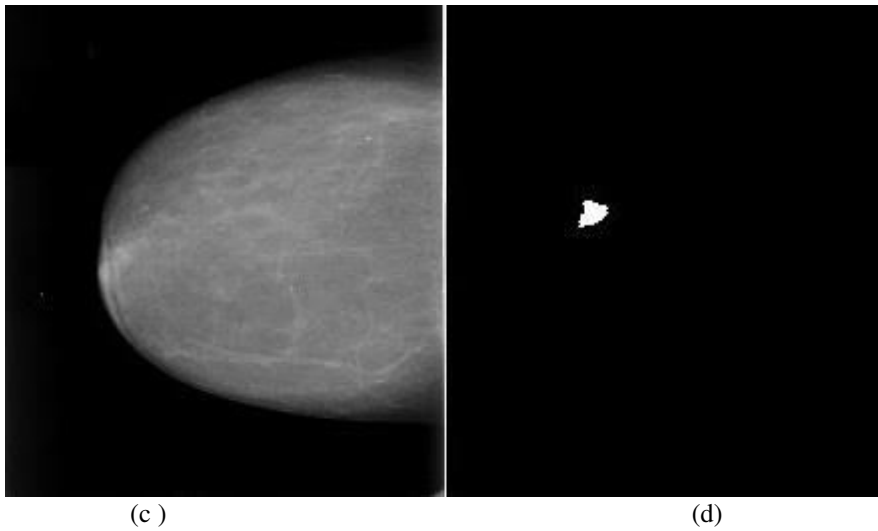


Figure 4 c & d : MWu-Conv-UNET Segmentation & Output image

	precision	recall	f1-score	support
0	0.98	1.00	0.99	61
1	1.00	0.99	0.99	93
accuracy			0.99	154
macro avg	0.99	0.99	0.99	154
weighted avg	0.99	0.99	0.99	154

Fig 5 Classification Report of .modified weight updated UNET segmentation

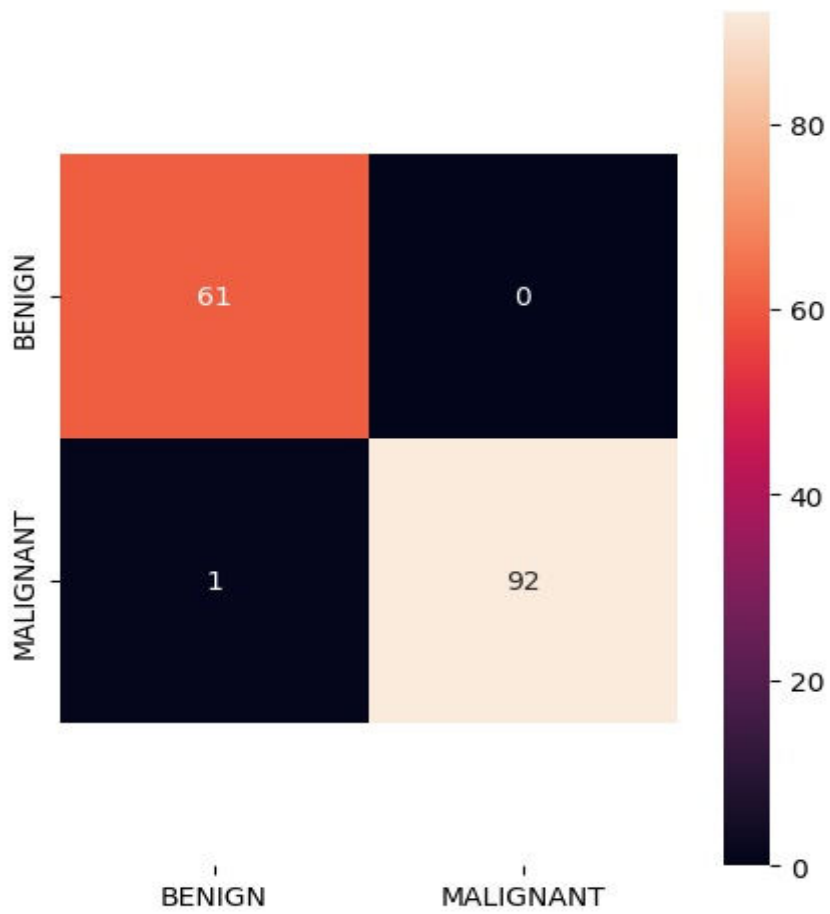


Fig 6 Confusion Matrix

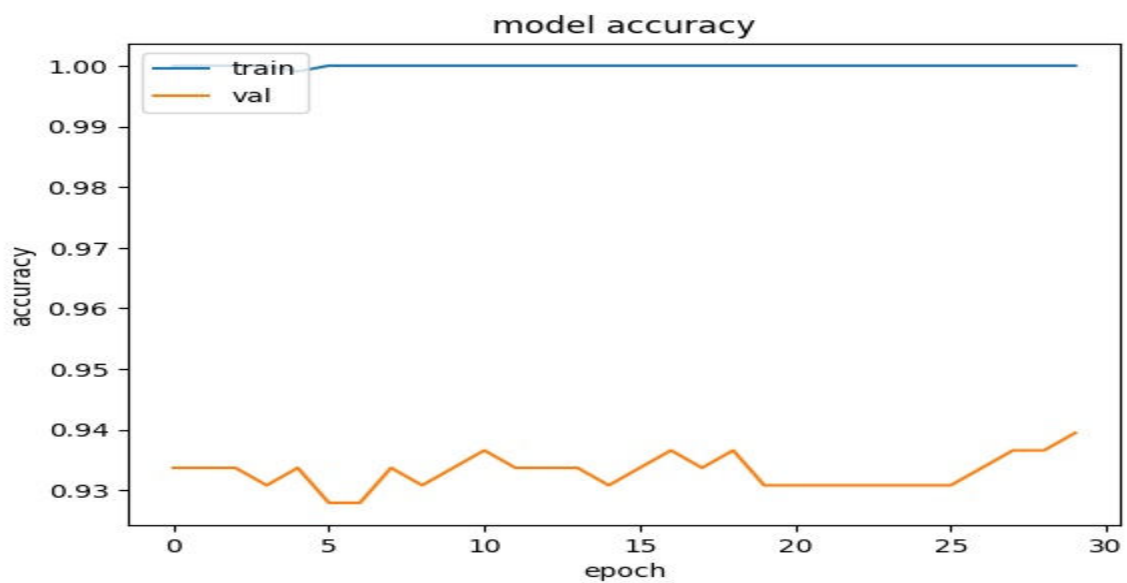


Fig 7 Model Accuracy plot of modified weight updated UNET segmentation



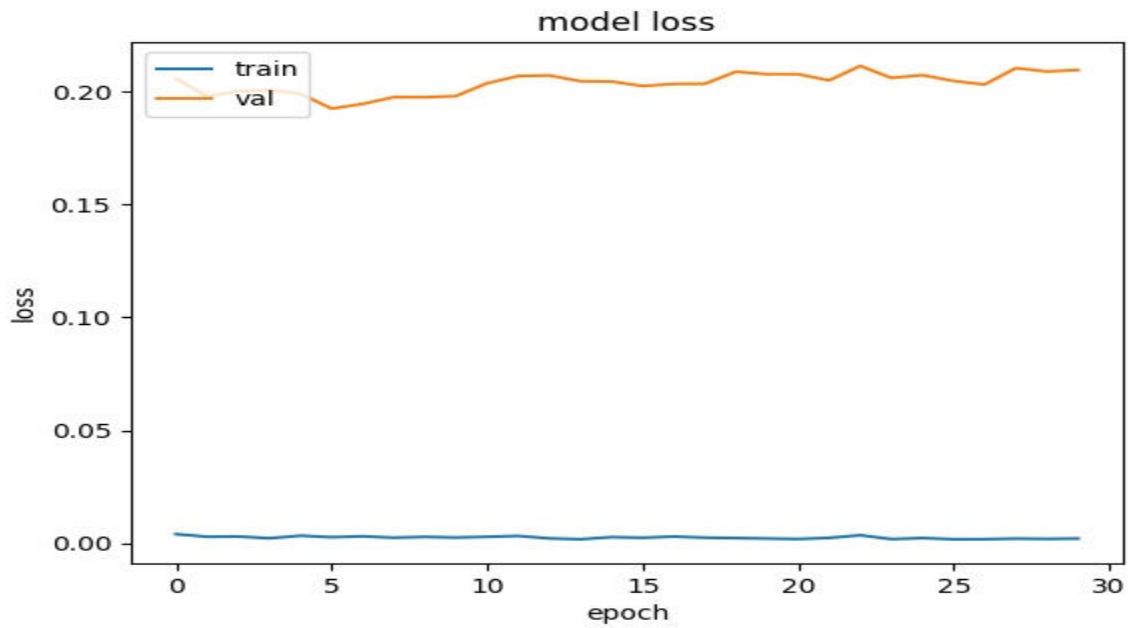


Fig 8 Model Loss plot of modified weight updated UNET segmentation

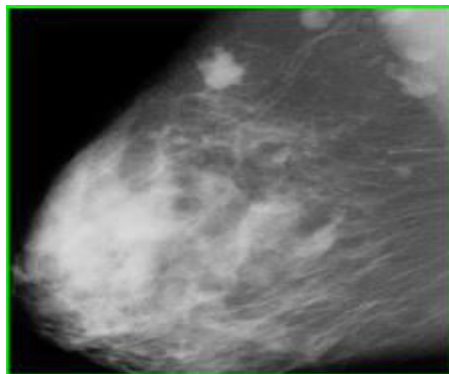


Fig 9 a Input

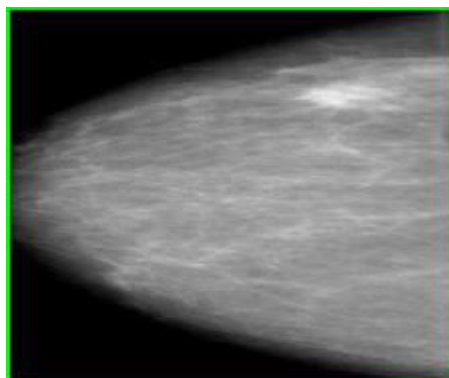


Fig 9 b Input

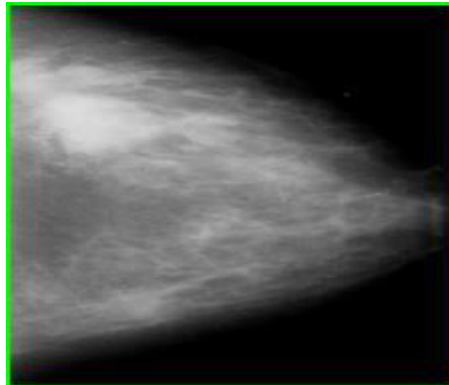


Fig 9 C Input

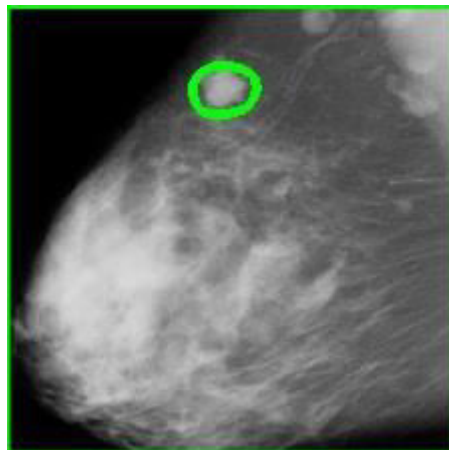


Fig 10 a Cancer Detection Output

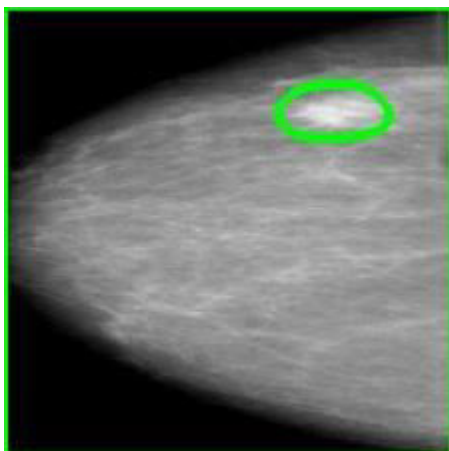


Fig 10 b Cancer Detection Output

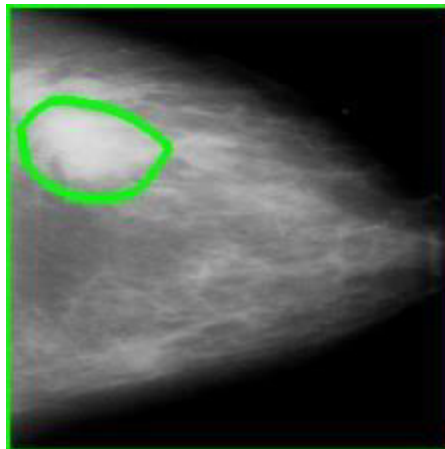


Fig 10 c Cancer Detection Output

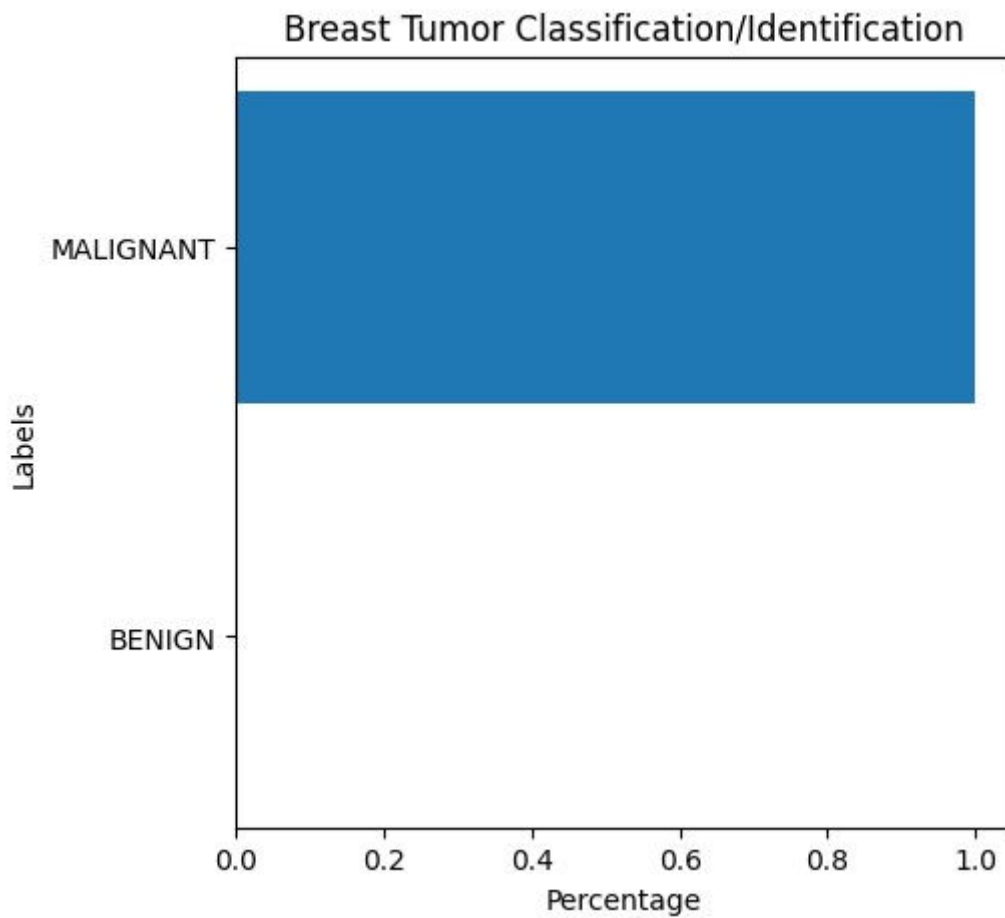


Fig 11 a Classification output

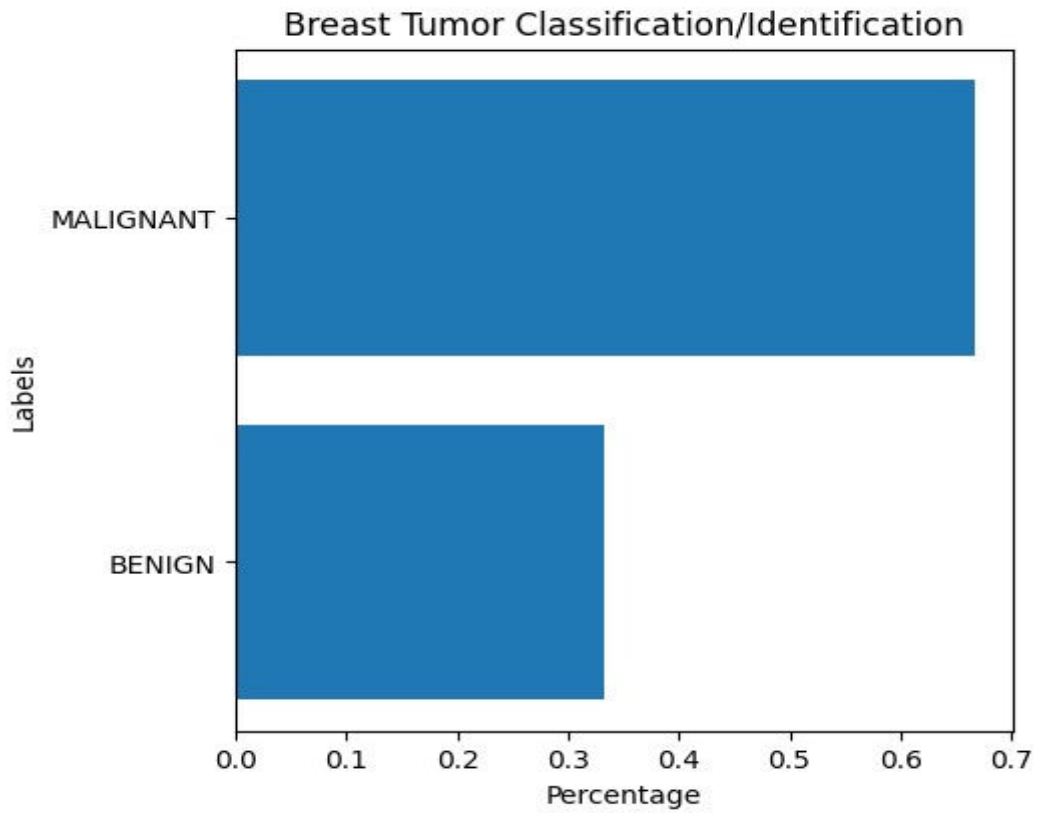


Fig 11 b Classification output

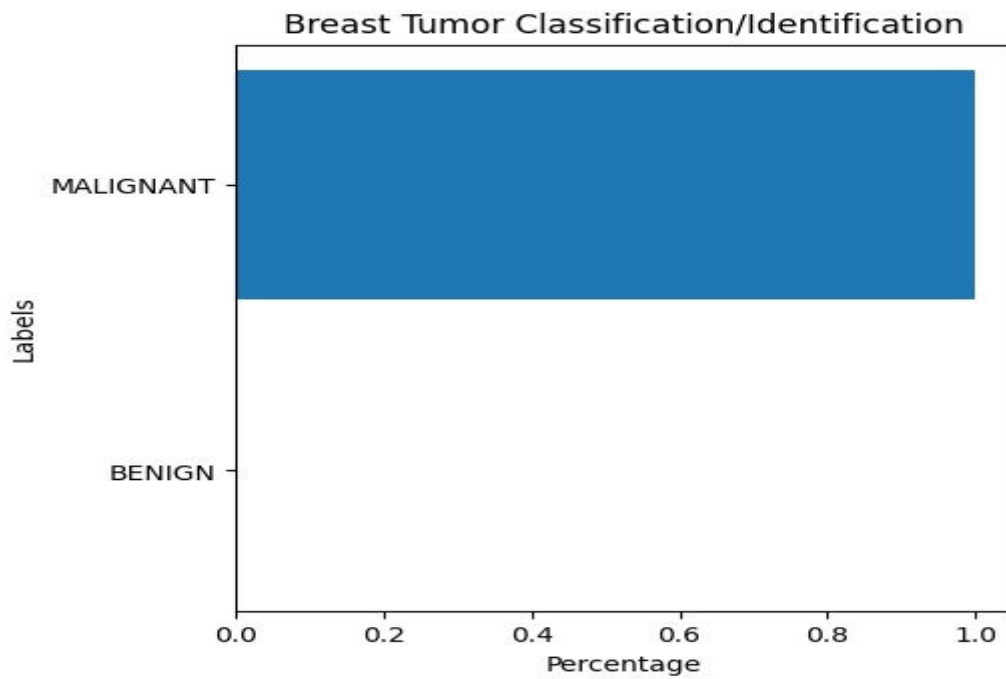


Fig 11 C Classification output

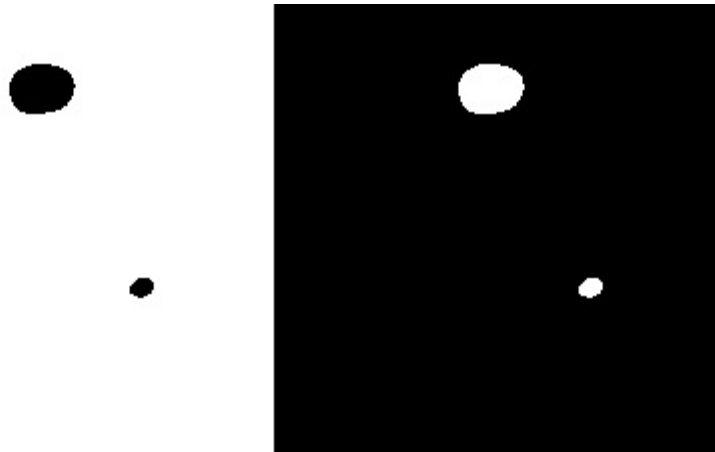


Fig 12 a modified weight updated UNET segmentation Output

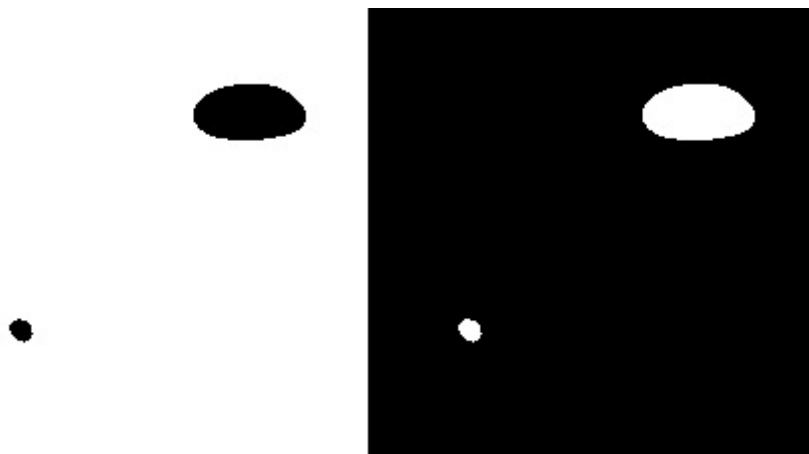


Fig 12 b modified weight updated UNET segmentation Output



Fig 12 C modified weight updated UNET segmentation Output

Handcrafted elements were included with the intention of enhancing the results produced by the innovative CNN architecture by providing further information that incorporates domain experts' experience. The study's conclusions included the following: 1. The transfer learning strategy is useful for achieving benchmarks. 2. The outcomes of transfer learning may be enhanced. 3. Robust outcomes are not always ensured by CNN architectural complexity. 4. By combining low-level and high-level features, recognition accuracy can be increased.

### VIII. CONCLUSION

Although several frameworks and solutions have been presented by AI researchers for the diagnosis and categorization of breast cancer, doctors and clinicians are not actively implementing these solutions because of the "black-box" nature of the frameworks. CNN and deep learning architectures are inherently opaque, making it challenging to understand the logic behind an AI algorithm's prediction. Nonetheless, it is critical for doctors to comprehend the rationale behind any given decision. Therefore, a key factor in the adoption of technology is the model's explain ability. A further obstacle to the development of AI-based systems is the lack of complete, fully labeled and annotated datasets that adhere to international distribution guidelines and ethical standards. To be trained robustly, deep learning models require large amounts of data, which makes them data-hungry.

### REFERENCES

1. X. Jia, W. Meng, S. Li, Z. Tong and Y. Jia, "A rare case of intracystic Her-2 positive young breast cancer," *2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, Houston, TX, USA, 2021, pp. 2598-2602, doi: 10.1109/BIBM52615.2021.9669897.
2. Y. Shao *et al.*, "Breast Cancer Detection Using Multimodal Time Series Features From Ultrasound Shear Wave Absolute Vibro-Elastography," in *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 2, pp. 704-714, Feb. 2022, doi: 10.1109/JBHI.2021.3103676.
3. A.P. A. Sharma, S. R. Kawale, S. P. Diwan and D. G. V, "Intelligent Breast Abnormality Framework for Detection and Evaluation of Breast Abnormal Parameters," *2022 International Conference on Edge Computing and Applications (ICECAA)*, Tamilnadu, India, 2022, pp. 1503-1508, doi: 10.1109/ICECAA55415.2022.9936206.
4. M. Kavitha, P. Venkata Krishna, V. Rama Krishna, S. Digavinti, M. Kalyani and T. V. S. Naga Divya, "Android based Recommender System (ARS) to Detect Breast Abnormalities," *2022 International Conference on Automation, Computing and Renewable Systems (ICACRS)*, Pudukkottai, India, 2022, pp. 973-978, doi: 10.1109/ICACRS55517.2022.10029264.
5. A.K. S, J. P. Dharmarajan, V. D. K and N. Subbanna, "Atlas based breast registration and segmentation in the Mediolateral Oblique and Craniocaudal Views," *2023 Second International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT)*, Trichirappalli, India, 2023, pp. 1-5, doi: 10.1109/ICEEICT56924.2023.10157170.
6. M. E. H. Rasheed, M. Youseffi, M. M. A. Jamil and N. A. A. Rahman, "Medical Imaging and Analysis of Dense Breast Tissue: A Case Study," *2021 IEEE 17th International Colloquium on Signal Processing & Its Applications (CSPA)*, Langkawi, Malaysia, 2021, pp. 1-5, doi: 10.1109/CSPA52141.2021.9377288.
7. J. Li *et al.*, "A Preclinical System Prototype for Focused Microwave Breast Hyperthermia Guided by Compressive Thermoacoustic Tomography," in *IEEE Transactions on Biomedical Engineering*, vol. 68, no. 7, pp. 2289-2300, July 2021, doi: 10.1109/TBME.2021.3059869.
8. Z. KHOMSI, A. ELOUERGI, A. AFYF and L. BELLARBI, "Contribution for the Early Detection of Breast Cancer by a Superficial Thermography Solution," *2020 International Conference on Electrical and Information Technologies (ICEIT)*, Rabat, Morocco, 2020, pp. 1-4, doi: 10.1109/ICEIT48248.2020.9113225.
9. U. Khasana, R. Sigit and H. Yuniarti, "Segmentation of Breast Using Ultrasound Image for Detection Breast Cancer," *2020 International Electronics Symposium (IES)*, Surabaya, Indonesia, 2020, pp. 584-587, doi: 10.1109/IES50839.2020.9231629.
10. B. Barufaldi *et al.*, "Computational Breast Anatomy Simulation Using Multi-Scale Perlin Noise," in *IEEE Transactions on Medical Imaging*, vol. 40, no. 12, pp. 3436-3445, Dec. 2021, doi: 10.1109/TMI.2021.3087958.





11. P. Karatza, K. Dalakleidi, M. Athanasiou and K. S. Nikita, "Interpretability methods of machine learning algorithms with applications in breast cancer diagnosis," *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, Mexico, 2021, pp. 2310-2313, doi: 10.1109/EMBC46164.2021.9630556.
12. M. Akhil and P. V. S. Kumar, "Breast Cancer Prognosis using Machine Learning Applications," *2022 4th International Conference on Advances in Computing, Communication Control and Networking (ICAC3N)*, Greater Noida, India, 2022, pp. 488-493, doi: 10.1109/ICAC3N56670.2022.10074517.
13. G. Fontaine and S. Pistorius, "Initial characterization of a portable breast microwave detection device," *2021 IEEE 19th International Symposium on Antenna Technology and Applied Electromagnetics (ANTEM)*, Winnipeg, MB, Canada, 2021, pp. 1-2, doi: 10.1109/ANTEM51107.2021.9518535.
14. H. Q. Vo, P. Yuan, T. He, S. T. C. Wong and H. V. Nguyen, "Multimodal Breast Lesion Classification Using Cross-Attention Deep Networks," *2021 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI)*, Athens, Greece, 2021, pp. 1-4, doi: 10.1109/BHI50953.2021.9508604.
15. J. -Y. Yeh, Y. -C. Hou and Y. -L. Wang, "Breast Cancer Classification on Mammograms Using Extreme Learning Machines," *2023 International Conference on Consumer Electronics - Taiwan (ICCE-Taiwan)*, PingTung, Taiwan, 2023, pp. 509-510, doi: 10.1109/ICCE-Taiwan58799.2023.10226965.
16. P. S. Shekar Varma, S. Kumar and K. Sri Vasuki Reddy, "Machine Learning Based Breast Cancer Visualization and Classification," *2021 International Conference on Innovative Trends in Information Technology (ICITIIT)*, Kottayam, India, 2021, pp. 1-6, doi: 10.1109/ICITIIT51526.2021.9399603.



**INNO SPACE**  
SJIF Scientific Journal Impact Factor  
**Impact Factor: 8.423**



**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 **9940 572 462**  **6381 907 438**  **ijirset@gmail.com**



[www.ijirset.com](http://www.ijirset.com)

Scan to save the contact details