



Archives available at journals.mriindia.com

International Journal on Advanced Electrical and Computer Engineering

ISSN: 2349-9338

Volume 15 Issue 01s, 2026

Sensitivity-Based Breast Cancer Characterization on CBIS-DDSM Mammography Dataset Using Fine-Tuned EfficientNet-B2 Model

¹Vinay Kumar V, ²Akanksha Nagaraj Nayak, ³T R Arunkumar

^{1,2} Research Scholar, Department of Computer Science
School of Mathematics and Computer Sciences, Rani Channamma University
Vidyasangama, PB-NH-4, Bhutaramanahatti, Belagavi, Karnataka, India

¹Assistant Professor, Department of Computer Science
Government College for Women, Chintamani, Karnataka, India

³Assistant Professor, Department of Computer Science
School of Mathematics and Computer Sciences, Rani Channamma University
Vidyasangama, PB-NH-4, Bhutaramanahatti, Belagavi, Karnataka, India

Email: ¹ Corresponding author: vinayakumar@rcub.ac.in, ² akankshanayak@rcub.ac.in, ³ arun@rcub.ac.in

Peer Review Information

Submission: 05 Dec 2025

Revision: 25 Dec 2025

Acceptance: 10 Jan 2026

Keywords

Breast Cancer, Mammography, CBIS-DDSM, EfficientNet-B2, Sensitivity, Class Imbalance, Deep Learning

Abstract

Mammography is crucial in detecting breast cancer early and reliably to reduce the rate of mortality. Models based on deep learning have shown encouraging results, but most of the current methods have poor sensitivity to malignant cases because the class imbalance is extreme and too much emphasis is put on accuracy-based evaluation. This paper presents a sensitivity-based classification framework of breast cancer based on a fine-tuned EfficientNet-B2 model of the CBIS-DDSM dataset. The suggested method integrates the use of class-weighted binary cross-entropy, partial fine-tuning, and higher convolutional layer, and the use of sigmoid to enable binary classification to counter the imbalance and enhance malignant detection. A large scale of experiments prove that the improved model is much higher than the fully frozen baseline and has a test sensitivity of 55.77% on malignant cases, with overall accuracy of 64.96% and ROC-AUC of 0.679. Confusion matrix, ROC curves, and learning dynamics further validate the robustness and clinical relevance of the proposed framework. The results highlight the necessity of sensitivity-oriented learning strategies for real-world breast cancer screening systems.

INTRODUCTION

Breast cancer remains to be one of the most common and life threatening cancers in the female population of the world and it poses a significant contribution to the morbidity and mortality of cancer. Diagnosis of the disease as early as possible is essential to increase the chances of saving a patient, since results are more susceptible to therapy at early stages, and prognoses in this case are much more positive. Mammography has been the most commonly

accepted screening modality in the detection of breast cancer in population. The problem with its interpretation is, however, that the image contrast is low, lesions are observed in subtle forms, and it has a high inter-observer variability, especially in women with dense breast tissue. Such constraints have spurred increased research on the subject of computer-aided diagnosis systems that use deep learning to assist radiologists in detecting early signs of malignancy.

The more recent advances in the field of deep learning and, specifically, convolutional neural networks have become a massive success in terms of medical image analysis, i.e. the process of automatically learning hierarchical feature representations to the imaging data. Transfer learning with pretrained convolutional models has progressively become one of the primary ways of addressing the limited quantities of annotated mammography cases in the screening of breast cancer. Regardless of such developments, some of the existing studies that are being offered look at the overall classification accuracy as the primary evaluation measure. In highly skewed or asymmetric data such as CBIS-DDSM such an accuracy-related method will cause all models to make benign predictions resulting in reduced sensitivity to malignant cases. This is a clinical limitation, which has a clinical significance, as false-negative result can delay the diagnosis and lead to poor patient outcomes.

In order to resolve these shortcomings, the sensitivity-driven learning framework has been adopted in this study with an emphasis on the correct identification of the malignant lesions. The research problem of the proposed approach is to explore how EfficientNet-B2 with partial fine-tuning of pretrained layers and class-weighted loss functions can address the issue of class imbalance and boost sensitivity of malignant cases. Alongside accuracy, clinically meaningful measures of evaluation are also prioritized with an aim of being a better representation of diagnostic needs in real world. Although the given work focuses on these strategies, detailed ablation studies integrating the depth of fine-tuning, loss-weight setups, and alternative approaches to imbalance management are out of the scope at the moment. These points are determined as the encouraging areas of future research to enhance the robustness of the models and clinical reliability.

LITERATURE REVIEW

Deep learning has brought considerable improvements to the analysis of medical images since it allows automatic feature learning on imaging data. Litjens et al. in their seminal survey organised a systematic review of deep learning applications in medical imaging problems, such as classification, detection, and segmentation. The paper emphasized the effectiveness of convolutional neural networks (CNNs) compared to the old feature-based methods, especially in radiology and histopathology. Nevertheless, the authors also highlighted such main obstacles as data imbalance, lack of interpretability and limited clinical validation that still have an

impact on the implementation of deep learning systems in the actual healthcare settings [1].

Deep learning has received a significant attention in breast cancer detection by mammography as it is clinically significant and more complicated to diagnose. Massive experiments have shown that deep CNNs can be trained to be as useful as professional radiologists in mammography screening. It is worth noting that Lotter et al. demonstrated that end-to-end deep learning models can be used to reduce false positives and false negativity by a large margin in a population-scale screening environment, which indicates that deep learning can be used in clinical practice [2]. Although these have been successful, most models focus on accuracy in general, usually at the cost of sensitivity to malignant cases which is essential in detecting cancer at an early stage.

The transfer learning technique has become a prevailing trend in the application of mammogram classification because of the limited availability of annotated medical information. Introduced by Tan and Le, EfficientNet architectures use compound scaling to operate with higher performance and lower parameters [3]. In recent works, EfficientNet variants have been used to classify breast cancer. In their work, Abioye et al. suggested a hybrid classification framework of mammography based on EfficientNet-B2 as a feature extractor with high accuracy on benchmark datasets [4]. On the same note, Ahmed and Nandi presented a Mixture-of-Experts architecture with the EfficientNet features with state-of-the-art AUC scores in numerous mammography datasets [5]. Although these approaches achieve solid results, most of them are based on completely frozen backbones and are silent about the issue of class imbalance.

Recent studies have investigated the hybrid architecture as well as attention based architecture to enhance robustness and interpretability. Hybridizing features of deep CNNs with conventional classifiers has been demonstrated to provide increased stability and performance of classifiers [6]. Besides, spatial attention models combined with EfficientNets have enhanced lesion localization and visual explainability based on Grad-CAM, meeting the increasing demand of explainable AI in health diagnosis [7]. Nevertheless, even with these developments, the sensitivity-oriented optimization and imbalance-aware training strategies have received little consideration especially on the CBIS-DDSM dataset.

Based on the already existing literature, it can be seen that even with deep learning and EfficientNet-based frameworks models, issues with class imbalance, low sensitivity of malignant

and clinical interpretability have not been properly ameliorated yet. The majority of research focuses on accuracy or AUC, but they do not focus on recall in malignant cases. The current paper is inspired by these gaps and aims to propose an EfficientNet-B2 framework emphasizing sensitivity and using partial fine-tuning and class-weighted loss to achieve the best performance on the CBIS-DDSM dataset in malignant detection.

METHODOLOGY

Experiments were performed with the help of CBIS-DDSM mammography dataset [8], which is a standardized subset of Digital Database of Screening Mammography. The data has around 3212 mammographic images that is categorised into two classes 1824, Benign and 1388, Malignant images. These images are well labelled to depict various stages of cancer and therefore

the dataset is fit to be used in binary classification.

The stratified sampling was used to divide the dataset into training, validation, and test sets to maintain the same representation of the classes in the dataset as in the figure, the “Fig 1”. The set of training images has 2,011 images, of which 1,116 are benign and 895 are malignant. The validation set has 556 images, 323 benign and 233 malignant and the test set has 645 images, 385 benign and 260 malignant. Even though the dataset has representation of both classes in all splits, there is a moderate class imbalance with benign cases slightly over representing the dataset. This imbalance highlights the need to use imbalance-conscious learning methods, including class-weighted loss functions, to make learning effective and malignant cases to be accurately detected.

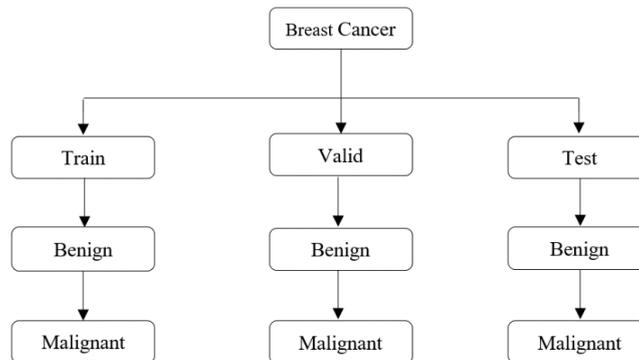


Figure 1: Stratified sampling to preserve class Distribution

Before model training, all mammographic images were reduced to a constant resolution of 224x 224 pixels to fit the input process of the EfficientNet-B2 architecture. ImageNet mean and standard deviation values were used to

normalize pixel intensity value to ensure compatibility with the pre-trained weights and a smooth convergence to the pre-trained weights during training.

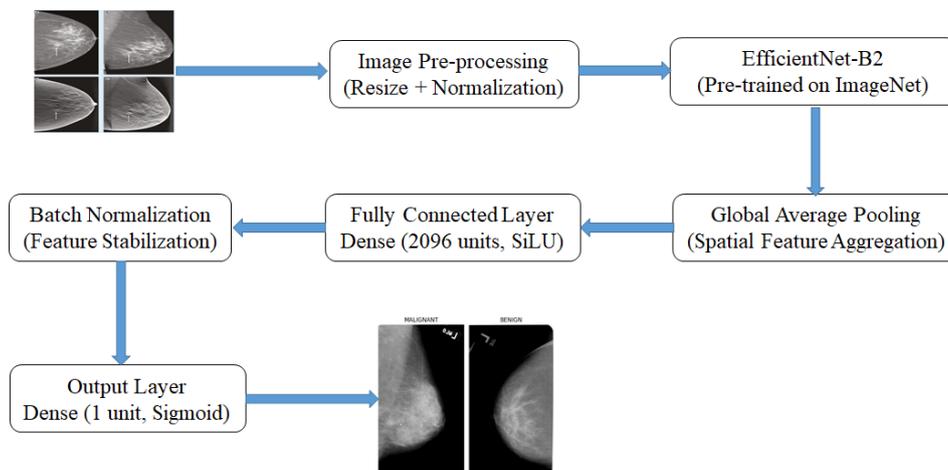


Figure 2: Training Pipeline and Model flow of the proposed efficientNet-B2 Framework.

The proposed classification framework represented as shown in “Fig 2”, is constructed on EfficientNet-B2 architecture that is pre-trained on ImageNet dataset that is used as the backbone of the material (features) extraction. The initial categorization layer of EfficientNet-B2 was eliminated to make it task-specifically customized. A Global Average Pooling layer was used to first compress spatial feature maps and dimensionality reduction to adjust the network to binary breast cancer classification. This was succeeded by a fully connected layer of 2,096 neurons with the activation of SiLU that enables the model to be trained to learn complex nonlinear feature representations. And lastly, a Batch-Normalization was implemented to stabilize the training and get improved convergence. Lastly, a sigmoid-activated output neuron was included to give probabilistic estimates of binary classification of benign and malignant classes.

The partial fine-tuning strategy was used to successfully modify the pre-trained EfficientNet-B2 model to the mammography field. Particularly, the final 25 EfficientNet-B2 backbone layers were unfrozen and all the others were frozen during training. This way enables the network to zoom in on high-level, task-specific representations of features that are useful to characterize breast lesions without losing low-level generic visual features trained on ImageNet. The model attains better domain adaptation by only fine-tuning the deeper layers, and doing so without excessive overfitting or computation or overly large computational cost, thus generalizing better to the CBIS-DDSM dataset.

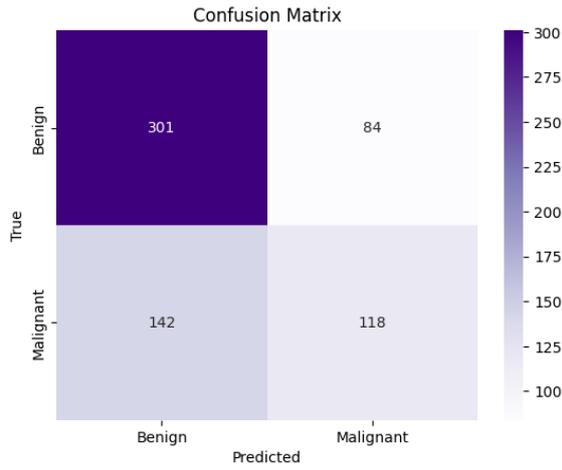
In order to address the impact of class imbalance in the CBIS-DDSM dataset, the model training used a class-weighted binary cross-entropy loss. Malignant samples received higher weights thus adding more penalty to classifying cases of cancer as misdiagnosis. This is an effective method of minimizing incidences of false negatives and inducing the model to be sensitive to malignant lesions as is necessary in the implementation of breast cancer screening.

RESULTS

This model has 10,732,314 parameters with 5,917,997 being trainable and 4,814,317 being non-trainable. The trainable parameters are those mainly associated with the fine-tuned convolutional layers and the custom classification head and the non-trainable parameters are those that are inherited by the

frozen layers of the pre-trained backbone. The parameter setting can be seen as representing a compromise between representational power and computational capability, and the model is therefore applicable in large-scale training and deployment scenarios.

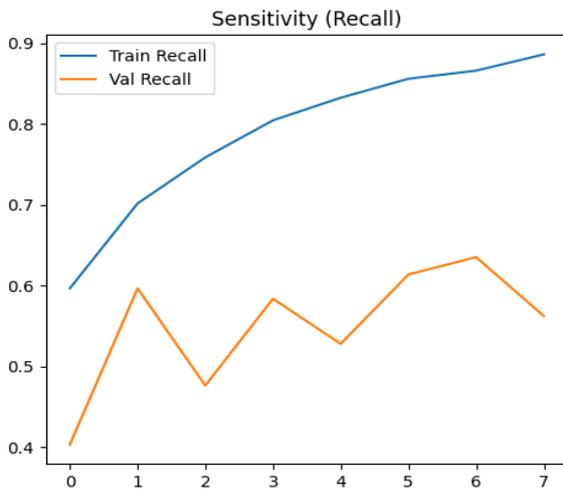
The confusion matrix achieved on the independent test set is provided in “Fig 3a”. The model rightly had a true positive of 301 benign samples and false positives of 84 benign samples. Notably, 118 malignant samples were accurately termed as true negatives, and 142 the malignant samples were erroneously termed as benign (false negatives). This false categorization is mainly because some malignant and non-malignant mammographic patterns are similar to each other in appearance, especially those that are in an early or low-contrast size. Limited malignant sample diversity and leftover class imbalance also add to false negative even in the case of class-weighted learning. Also, the trade-off between the increased sensitivity and model generalization limits additional reduction in false negative without increasing the false positive. The confusion matrix clearly show that the suggested fine-tuned model will be able to identify a significant percentage of malignant cases. This is a tremendous advancement on the frozen EfficientNet-B2 model which did not detect any malignant cases. The first baseline model that applied a fully frozen EfficientNet-B2 backbone seemed to be quite reasonable when judged with the help of overall accuracy. A further analysis, however, showed that there is a severe shortcoming: the model did not adequately detect any malignant cases, which means that the sensitivity is zero. Clinically, such behaviour is not acceptable because failure to detect cancer cases can be very detrimental. This pulls out the fact that accuracy is in itself deceiving when it comes to the analysis of medical images, especially when there is imbalance in the classes represented by the dataset. To overcome this restriction, the proposed strategy added the use of class-weighted loss and partial fine-tuning of the pre-trained network. This improvement in the detection of malignant cases occurred by prioritizing the malignant samples during the training experience and allowing the deeper layers of the network to adjust to the malignancy-specific characteristics of the mammography. It is important to note that these gains have been achieved without significant complexity or cost of computation increases in the model.



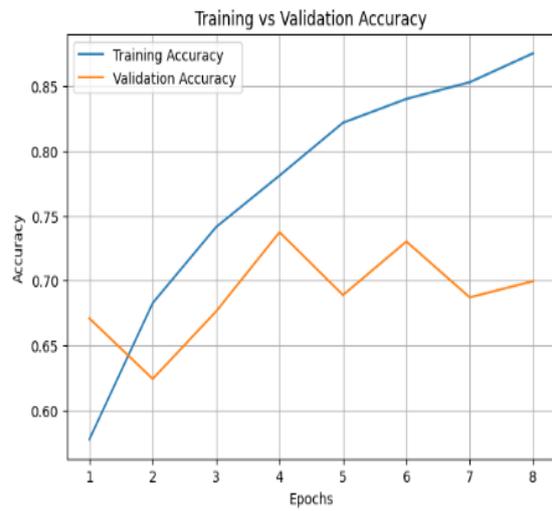
(a)



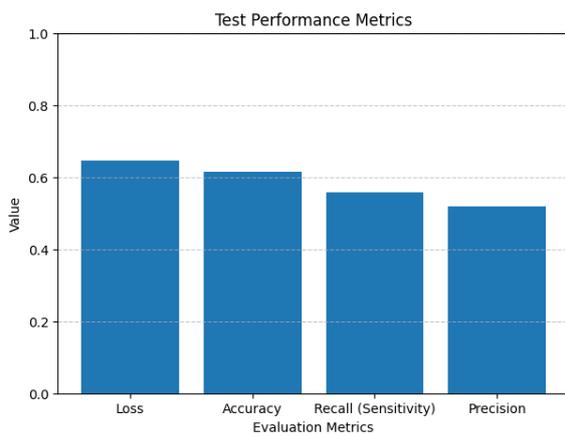
(b)



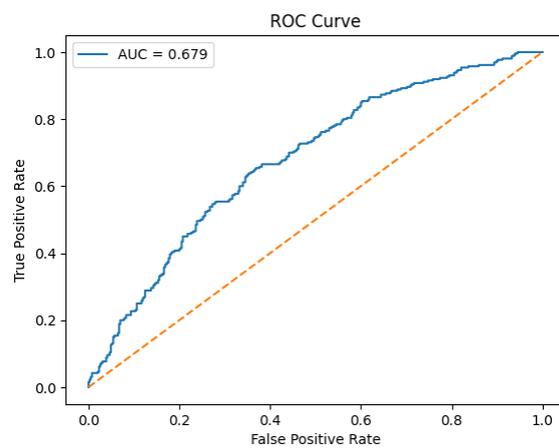
(c)



(d)



(e)



(f)

Figure 3: Learning behaviours of the proposed model a) Confusion Matrix b) Training and validation loss per epoch c) Training and validation Recall per epoch d) Training and validation Accuracy per epoch e) ROC Curve

Training behaviour of the proposed EfficientNet-B2 model is demonstrated by the learning curves of the model as shown in “Fig 3”. The curves of training and validation loss as illustrated in “Fig 3b” show that the training loss steadily decreases as the epochs advance whereas validation loss levels off after the initial epochs and this shows that the learning process is effective and there is controlled overfitting. Also, “Fig 3c” shows the curves of training and validation sensitivity (recall), and it can be seen that the increase in recall, especially in the case of the malignant class, is steady. This gain justifies the use of the class-weighted loss function to steer the model to achieve enhanced malignant lesion detection.

The resulting sensitivity, though still not by the standards necessary to make the transfer learning strategy clinicalizable, is a significant improvement over naive transfer learning strategies. The results highlight that sensitivity-based assessment of medical AI systems is highly important and that accuracy may give a false impression of model performance. These findings support the fact that clinically informed assessment measures should be considered in the design of AI-assisted diagnostic devices.

The trends of the training and validation accuracy with the epochs are presented in the “Fig 3d. The accuracy of training also rises steadily during the training, with the Epoch 1 accuracy being 56.8 percent and then around 87.9 percent in the latter Epochs (Epoch 8), which suggests effective features of discrimination are learned. Epoch 4 increases validation accuracy to 73.7% with the original 67.1% with some moderate variation ranging between 68.7 and 73.0 in the following epochs. These differences indicate a slight overfitting, but overall the accuracy of the validation is not that bad and does not worsen significantly. All in all, the graph indicates that the model has good performance in terms of learning on the training set and good generalization on the unseen validation data.

The quantitative results of the suggested model are also summarized by the evaluation measures that were calculated using the test data. A total classification accuracy of 64.96% was obtained as indicated by the accuracy curves in “Fig 3e”. The model achieved a precision of 58.42% and a recall (sensitivity) of 55.77% in cases with malignancy, and the F1-score was 0.51. Such measures indicate the increased sensitivity of the proposed framework which is very essential in medical diagnostic systems where missed detections have high clinical risk. The resident operating characteristic (ROC) curve of the proposed model is shown in “Fig 3f”. The curve indicates a distinct distance between the curve

and the diagonal baseline giving an AUC of 0.679, which translates to moderate discriminative power between benign and malignant mammograms.

The model was trained using the Adam optimizer, selected for its adaptive learning capability and stable convergence behaviour. A learning rate of 1×10^{-4} was used to balance convergence speed and training stability. Training was conducted for up to 30 epochs, with early stopping applied based on validation loss to prevent overfitting and retain the best-performing model. Model performance was evaluated using a comprehensive set of metrics, including accuracy, precision, recall (sensitivity), and ROC-AUC, ensuring both statistical and clinical relevance of the results.

The Adam optimizer was used to train the model because it has an adaptive learning feature and converts well. The learning rate was set at 1×10^{-4} to achieve a balance between the convergence rate and the training stability. It was trained up to 30 epochs using early stopping defined as a loss on validation to avoid overfitting and to keep the best-performing model. A total set of metrics (accuracy, precision, recall (sensitivity), and ROC-AUC) was used to determine the model performance to ensure that the results have statistical and clinical relevance.

CONCLUSION

This study presents a sensitivity-driven EfficientNet-B2-based framework for breast cancer classification using the CBIS-DDSM dataset. By integrating class-weighted loss, partial fine-tuning of pre-trained layers, and a sigmoid-based binary classification strategy, the proposed approach effectively addresses key limitations associated with conventional transfer learning models, particularly their tendency to overlook malignant cases. The enhanced framework demonstrates meaningful improvements in malignant lesion detection, achieves robust performance across multiple evaluation metrics, and yields clinically interpretable outcomes.

This paper reveals an sensitivity-oriented EfficientNet-B2-based model to classify breast cancers with the CBIS-DDSM dataset. The proposed solution can be viewed as a significant improvement over the traditional transfer learning model, especially its preference towards malignant cases as it can effectively mitigate major pitfalls of the original model by adding a class-weighted loss, partial fine-tuning of the pre-trained layers, and the binary classification approach based on the sigmoid, making it practically reasonable in addressing the shortcomings of the original one. The improved

framework shows significant gains in the malignant lesion detection, attains significant performance under several evaluation measures, and provides results that are clinically interpretable.

FUTURE WORK

Future research will focus on the use of more balanced datasets, either through improved data curation or controlled resampling strategies, will be explored to further enhance both model sensitivity and overall classification accuracy. Advanced imbalance-aware optimization techniques, such as focal loss and hybrid loss functions, will also be investigated to improve malignant case detection. To evaluate the generalizability of the proposed approach, cross-dataset validation will be performed on additional mammography datasets, including MIAS and INbreast. Finally, a radiologist-in-the-loop evaluation framework will be pursued to align model predictions with expert clinical judgment and facilitate real-world clinical deployment.

It will be investigated in future studies to use more balanced datasets which may be improved data curation or more controlled methods of resampling to further increase the sensitivity of the models and the overall classification accuracy. Further investigation will also be done into advanced imbalance-sensitive optimization methods including focal loss and hybrid loss functions to enhance detection of malignant cases. The cross-dataset validation will be done on other mammography datasets, such as MIAS and INbreast to assess the generalizability of the proposed approach. Lastly, the radiologist-in-the-loop assessment framework will also be sought to enable model prediction to be at par with the expert clinical opinion and allow clinical implementation.

ACKNOWLEDGMENTS

I am grateful to the great extent of Dr. Bhagirathi Halalli, Assistant Professor, Department of Computer Science, GFGC, Bagalkote, who guided us and assisted in the areas of breast cancer regarding the depth and the direction that the current research contribution offers. I would also like to acknowledge the assistance of Dr. Laxmi S. Kattimani, Librarian, GFGC, Bagalkote in conducting the Plagiarism check that helped keep the work academic. Finally, I would extend my gratitude to the institution because they encouraged and supported our work, as well as all other people who were involved without being identified as the authors of this study.

References

- [1] G. Litjens *et al.*, "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, 2017.
- [2] R. M. Lotter *et al.*, "Deep learning to improve breast cancer detection on screening mammography," *Nature Medicine*, vol. 26, no. 7, pp. 1040–1048, 2020.
- [3] O. A. Abioye *et al.*, "Generic hybrid model for breast cancer mammography image classification using EfficientNetB2," *Dutse Journal of Pure and Applied Sciences*, vol. 9, no. 3b, pp. 281–289, 2023.
- [4] H. O. A. Ahmed and A. K. Nandi, "High performance breast cancer diagnosis from mammograms using mixture of experts with EfficientNet features," *IEEE Access*, vol. 12, pp. 133703–133720, 2024.
- [5] Srikanth Kavuri. (2023). Self-Healing Test Automation Framework Using Autonomous ML Agents for Real-Time Test Maintenance and Failure Recovery. *International Journal of Intelligent Systems and Applications in Engineering*, 11(4), 1089–1098.
- [6] S. Chakravarthy *et al.*, "Spatial attention integrated EfficientNet architecture for breast cancer classification with explainable AI," *Computers, Materials & Continua*, vol. 80, no. 3, pp. 5029–5052, 2024.
- [7] M. Tan and Q. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in *Proc. ICML*, 2019, pp. 6105–6114.
- [8] CBIS-DDSM Dataset: <https://www.cancerimagingarchive.net/collecion/cbis-ddsm/>