International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** V **Month of publication:** May 2025

DOI: https://doi.org/10.22214/ijraset.2025.71017

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

Mammary Tumor Screening

Prof. Santushti Betgeri[1], Manali Dubla[2], Manasi Patil[3], Tanishq Pardeshi[4], Ashumal Palde[5]
Department of Computer Engineering, Vishwakarma Institute of Technology, Pune, India

Abstract: Mammary Tumor Screening using deep learning provides an innovative approach for early breast cancer detection. In this work, a model trained on Convolutional Neural Networks (CNNs) on the Kaggle Multi Cancer dataset, consisting of 10,000 high-resolution histopathological images of benign and malignant tumors. To improve model performance and lessen overfitting, preprocessing methods like resizing, normalisation, and data augmentation are used. The CNN model .The CNN model is designed for binary classification, and itsF1-score, recall, accuracy, and precision are used to assess performance. This inquiry seeks to fashion a faultless core, employing an exhaustive dataset used by a cancer detection system. The high-resolution dataset comes from Kaggle, consisting of histopathological images of both benign and malignant growths, specifically malignant breast tumors. It furnishes multiple images to procure thorough diagnostic evaluations quickly. Experimental results show high effectiveness, implying a high level of helpfulness. Every instance of breast cancer being detected early improves patient outcomes

Keywords: Medical Imaging, Tumor Classification, Machine Learning, Deep Learning, Computer-Aided Diagnosis, Convolutional Neural Networks , Feature Extraction.

I. INTRODUCTION

Breast cancer continues to be one of the most common and deadly cancers worldwide, further highlighting the important need for dependable early detection systems[1]. Customary diagnostic methods, such as mammography and ultrasound, are undoubtedly important. However, these methods frequently involve frustrating delays and are quite susceptible to human error[2]. Classifying each of the histopathological images of mammary tumors using Convolutional Neural Networks (CNNs) and other deep learning models provides an opportunity to automate and improve the precision of breast cancer screening[3].

This investigation formulates a deep CNN-based model for Mammary Tumor Screening, leveraging the Kaggle Multi Cancer dataset, which includes a total of 10,000 exceedingly detailed images of both benign and malignant tumors. The system aims to improve diagnostic accuracy and speed through image preprocessing, normalization, data augmentation, and exhaustive model training for binary classification[4]. A cloud-based platform with an exceptionally user-friendly graphical interface facilitates the model's deployment, enabling healthcare experts to acquire prompt, accurate tumor classification results and readily access these outcomes[5].

II. LITERATURE REVIEW

Breast cancer remains a leading cause of mortality among women worldwide, emphasizing the critical need for effective early detection methods[1]. Mammography, widely regarded as the primary screening tool, offers high sensitivity for detecting microcalcifications, contributing significantly to early diagnosis and improved survival rates[2]. However, its diminished effectiveness inside dense mammary tissue has encouraged enhancements such as Digital Breast Tomosynthesis(DBT), which strengthens lesion detectability as well as lessens false positives when contrasted with standard mammography[3]. Notwithstanding these improvements, mammography retains certain limitations, all-embracing of radiation exposure joined with potential overdiagnosis, which might give rise to unnecessary treatments[4].

Wang, Sultana, and Kavanagh's 2023 investigation evaluated the repercussions of mammography screening on both breast cancer deaths and the incidence of overdiagnosis apparent within Australia's BreastScreen program, observing that while screening notably reduced mortality, a 21% overdiagnosis rate was also obvious [5]. These consequences stress the value of considering the benefits and detriments in breast cancer screening programs.

Researchers have examined the incorporation of thermal infrared (IR) imagery with deep learning (DL) methodologies[6]. Thermal imaging constitutes a non-intrusive and radiation-free technique of great utility. It determines those thermal variations upon the epidermis which might denote malignant expansion[7]. Tsietso and Yahya (2023) detailed a Computer-Assisted Diagnosis (CADx) system incorporating multiple thermograms and clinical information, achieving 90.48% precision, thereby suggesting the potential of uniting thermal imaging and advanced diagnostic instruments[6]. This strategy aligns with studies that suggest AI's role in

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

increasing diagnostic accuracy and lessening radiologist workload[8]. The integration of Artificial intelligence (AI) and mechanized erudition (ML) during breast cancer diagnosis has finished in Prominent advances occurred in recognition and. These were in regards to approval classification accuracy[9]. Multiple ML algorithms, including Support Vector Machines (SVM), together with k-Nearest Neighbors (k-NN). Convolutional Neural Networks (CNNs) do, in fact, exist. Formerly used to classify mammary carcinoma as benign.or malignant, showing superior Efficiency versus prior approaches.[10].

Advanced CNN structures ,such as AlexNet,GoogleNet, YOLO, and RetinaNet, do subsist.increasingly used for breast cancer diagnoses[11]. These specific models have displayed capacity inside tasks such as identifying traumas and classifying tumors types[11]. Specific models offer certain benefits. YOLO (You Only Look Once) is prominent due to due to its pace, RetinaNet provides a fairer result. Employ an efficient technique. Reference it in source[12]. Recent Refinements feature an integrated rendition that Integrates a customized CNN with EfficientNetV2B3. Framework, achieving precision of 96.3% through the Kaggle BC-IDC data repository is at[13]. This model's Accomplishment accentuates the skill for consolidation.Advanced architectures with large preprocessing techniques, like data enhancement, to elevate diagnostic performance [13].

Common utilization of many ML models in breast cancer frequently, diagnosis includes preprocessing methods. Including procedures like feature scaling toward standardizing independent variables elements such as preserving the twofold attribute from dependent variables[7]. Studies using the Dataset for diagnostic breast cancer from Wisconsin used constructs like Random Forest (RF).Decision Tree, along with K-Nearest Neighbors k-Nearest Neighbors (KNN), Logistic Regression (LR), and Support Vector Machines stand as examples.Vector Classifier (SVC)[7]. Random Forest, in singular, manifested superior performance. accentuating the importance in feature prominence

explication in, and scrutiny of, ML models clinical settings[7]. Such strategies in preprocessing combined with model choice stay important. devising durable classification structures for breast cancer detection.

Examination of histopathological images represents still another important aspect to breast cancer diagnosis, supplying understandings into tissue configurations and cell characteristics such as nuclear size and shape[8]. The procedure is generally, manually, and lengthy and susceptible to variation from aspects like variable staining inconsistencies, coupled alongside the pathologist's experience[9]. Automated Computer-Helped CAD diagnostic systems, created through ML techniques, with the goal of elevating the aforementioned process through mechanization, stages like preprocessing, segmentation, and feature extraction, and classification [10]. These systems make it quite consistent and systematic avenues exist for detailed scrutinizing of histopathological images. Multiple methods are available for careful examination of such images. Reducing errors made by people and improving diagnoses.consistency [11].

Explainable AI (XAI) methodologies are increasingly common integrated in breast cancer diagnosis strengthen exactness and clarity [12]. These Techniques can be useful when handling the enigmatic attribute of DL models, presenting clinicians as perspicuous. Many understandings exist regarding model determinations[12]. For example, ResNet50-like models have frequently exhibited.

Heightened precision through enhancement and transfer learning techniques[12]. Additionally, preprocessing techniques like Contrast Limited Adaptive Histogram Equalization (CLAHE) has exhibited effectiveness regarding amplification mammographic images [12]. XAI methods like Grad-CAM, LIME, and SHAP lend themselves toward contribution assurance clinically through clarification of how AI models determine their predictions thus sanctioned medicinal decision-making [12].

The structured assessment by Chakravarthy, in conjunction with Rajaguru (2021) accentuates the revolutionary capacity to integrate clever technologies, regarding mammary carcinoma, machine learning goes hand in hand with robotics screening and classification[1]. Their findings align with diverse propensities across medical diagnostics where innovating innovations such as robotic-assisted biopsy procedures coupled alongside AI-propelled diagnostic platforms ease preliminary improvement and customization treatment plans [1]. These technologies not only improve diagnostic accuracy and additionally lessen lowered invasiveness coupled with increased patient outcomes[1]. The National Health Service throughout the UK as well Recognized the potential of AI by commencing the AI's primary trial concerning mammary cancer diagnosis, Including 700,000 mammograms under AI's assessment exactitude versus radiologists [4].

Despite these promising improvements,cdifficulties remain. AI models must be Popular across multiple factions and unrestricted stemming from predispositions inside instruction information, so as to obstruct inequities from start to finish diagnosis [9]. Furthermore, the clinical deployment. For artificial intelligence constructs, thorough verification is required to assure safety and efficacy[10]. Researchers are dealing with these quandaries through employment of Approaches like SMOTE (Synthetic Minority A Data-management Method involving Oversampling imbalance and increase model effectiveness[7]. Future examination must concentrate on wide-ranging. Execute multi-location research, and contemplate assimilation. Additional imaging methods for evaluating the real result of AI-backed diagnoses on patients. outcomes [8]

In summary, while discernible symbolism specific techniques like mammography, when joined with MRI, remain subject to thorough assessment The inclusion of retains large worth within breast cancer detection. Advanced ML models along with transparent AI methodologies is transforming diagnostics[6][11]. These ameliorations transcend a simple augmentation accuracy and effectiveness are also important during evaluations strengthen additional specialized routes targeted toward treatment[12]. Continued research and clinical Confirmations are vital for complete understanding. Undoubted expertises in AI-guided diagnostic equipment, certainly expediting prompt diagnoses, considerably lessening mortality. Qualities along with expansion within the mutual effectiveness of the mammary gland cancer screening programs [13].

III. METHODOLOGY

This study attempts to construct a strong breast cancer detection system, capitalizing on a repository acquired from Kaggle, featuring high-resolution histopathological images of both benign and malignant mammary tumors[6]. The "Multi Cancer" dataset includes images of multiple cancer types, especially pointing out breast cancer [7]. Each image in the dataset was derived from many clinical reports, which confirms that the data adequately depicts large tumor characteristics [8].

The complete dataset for this study includes 10,000 images, distributed among a pair of separate classifications, specifically benign as well as malignant tumors. Benign images depict completely noncancerous proliferations[6], while malignant images depict purely cancerous proliferations. Images are procured at many magnifications (40×, 100×, 200×, and 400×) to document exhaustive variations in tumor composition[7].

Rendered in RGB mode, these images exhibit a 700x460-pixel resolution. The dataset is divided into
two special classifications, specifically designated for both model instruction and assessment.

Benign images show tumors. These tumors are non-cancerous[8].

Malignant images are representative of cancerous tumors[9].

To guarantee the model operates skillfully and dependably, every image is uniformly downscaled to 224x224 pixels and carefully normalized to a pixel intensity range of [0, 1]. Furthermore, to considerably improve the dataset's diversity and substantially reduce the likelihood of overfitting, a variety of data augmentation techniques including rotation, mirroring, scaling, and cropping are implemented[11].

For classification, a Convolutional Neural Network (CNN) is used[2]. We employ the identical network for all feature extractions. Employing transfer learning with pretrained models like ResNet50 and VGG16 expedites training and elevates accuracy[7]. CNNs are purpose-built for binary classification, integrating dropout layers to preclude overfitting. The Adam optimizer quickens training greatly[10].

The dataset includes many training segments (70%, 7,000 images), alongside sizable validation (15%, 1,500 images) and testing (15%, 1,500 images) partitions[12]. Accuracy, precision, recall, in addition to the F1 score, constitute a suite of metrics used to thoroughly assess the model's performance, thereby guaranteeing consistently reliable as well as constant forecasts[13].

Once the model is trained, it is integrated into a graphical user interface (GUI) decision support system, making it easy for healthcare professionals to upload images and receive diagnostic results. To make the system more accessible, a cloud-based platform is also implemented for remote diagnostics. This allows the system to be used anywhere, anytime. To ensure the model's effectiveness in real-world scenarios, it is rigorously tested with unseen data, and regular updates are planned to continuously improve the system's accuracy and reliability[5].

IV. METHODOLOGY

The process starts with Data Collection, where histopathology images of breast cancer tumors (benign and malignant) are sourced from Kaggle's Multi Cancer dataset. Moving to Data Preprocessing, pictures are rescaled to 64x64 pixels, normalized to the range[0, 1], and augmented using techniques like flipping, rotation, and zooming to increase variability and reduce overfitting. The data is then split into Training (70%), Validation (15%), and Testing (15%) sets. A CNN Model is designed with Conv2D levels for attribute removal, MaxPooling2D layers for dimensionality reduction, and completely linked layer with a sigmoid output for binary classification[7]. During Model Training, early stopping is implemented to avoid overfitting by monitoring validation accuracy across epochs. Once trained, the model's performance is evaluated using metrics like accuracy, precision, recall, F1-score, and a confusion matrix. Finally, the model is deployed as a Web Tool, allowing healthcare professionals to upload images and receive instant classifications of tumors as benign or malignant.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

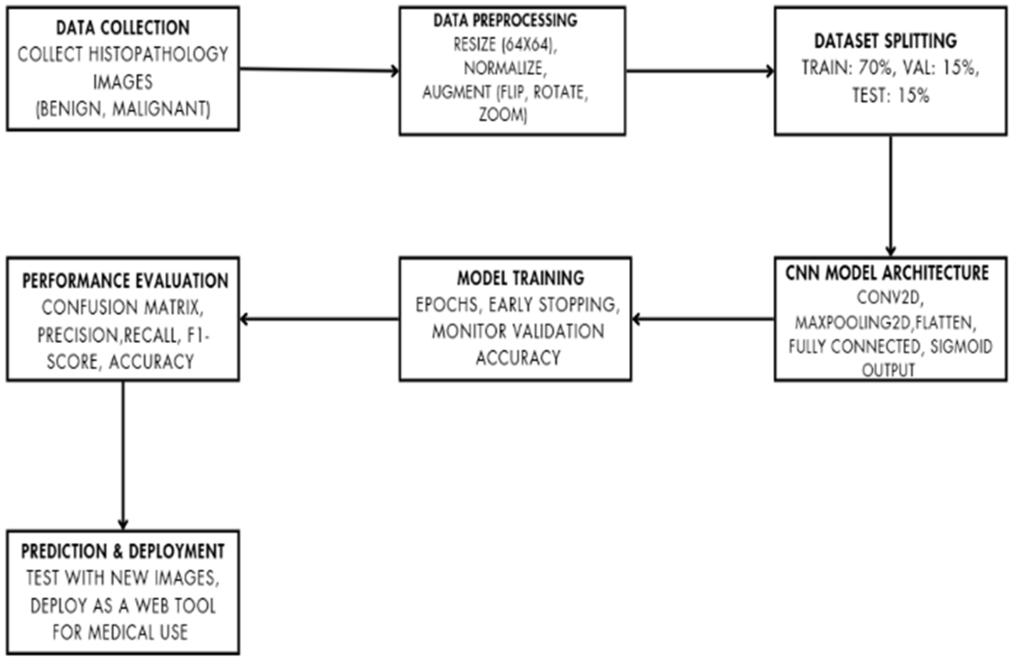
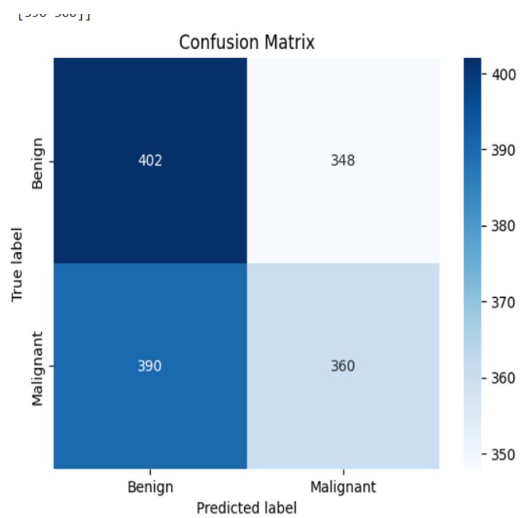
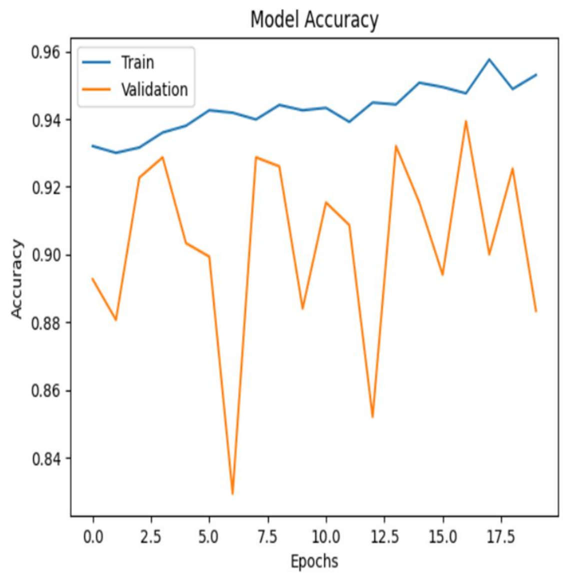
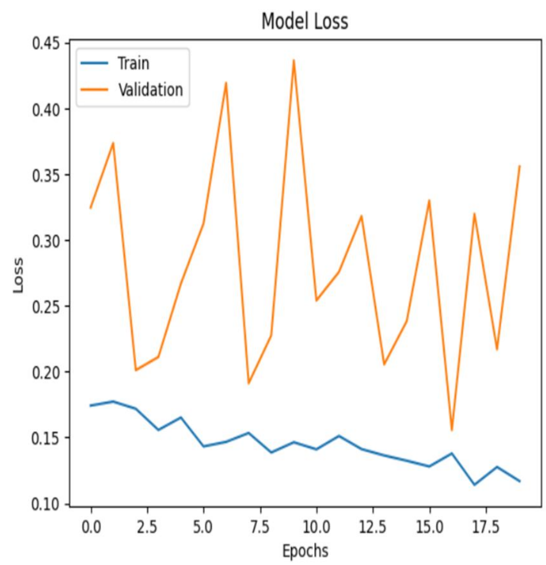


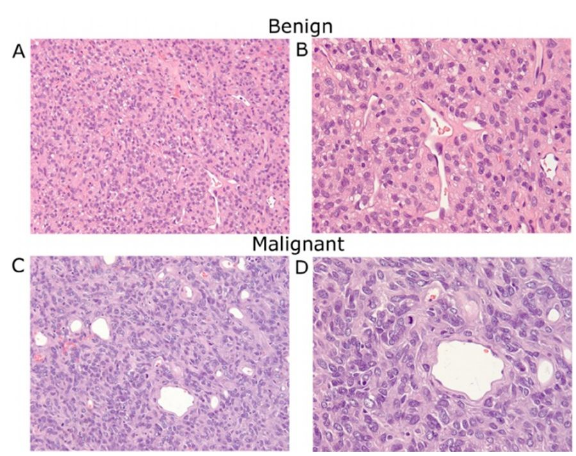
Figure 2: Methodology

V. RESULTS




Model Accuracy
Train
Validation
Accuracy
Epochs
0.96
0.94
0.92
0.90
0.88
0.86
0.84
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5


Model Loss
Train
Validation
Loss
Epochs
0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5


Benign
A
B
Malignant
C
D
,

VI. CONCLUSION

In conclusion, this innovative approach holds great promise and analysis of image is very powerful procedure to recognize the existence of breast cancer. Continued research, for optimal results, the integration of new modern techniques on diverse datasets to enhance detection of accuracy, and for better outcomes. For advanced imaging technologies such as ultrasound, histopathology images, mammograms, and many more for high accuracy. The accurate categorizing of tumors as benign or malignant remarkably upgrade early diagnosis and treatment planning.

VII. FUTURE SCOPE

Mammary Tumor Screening have bright future in data collection and analysis. The data can be analyzed by the images at early stage and that can reduce false positives and negatives. Here are some probable advancements and uses for Mammary Tumor Screening in the future:AI Integration and early detection: The technology in AI sector is increasing rapidly that can improve early stage detection through image processing and reducing the false negatives and positives. Care Screening: Low cost screening and portable tools can be access in rural areas and not so developed area, that can improve early treatment outcomes at early diagnosis. Telemedicine Integration: In this the application can be created for the remote screening, cloud based diagnostic tool and AI based consultations can be provided that everyone has access to expertise and can get proper opinions and this will reduce diagnosis time.

VIII. ACKNOWLEDGMENT

We like to thank Vishwakarma institute of Engineering for providing laboratory facilities for the execution of this work.
We are so grateful to our respected Prof. Santushti Betgeri. we would like to express our deepest gratitude and appreciation to our guide who review our work and pointed out the mistakes in our project.
We interviewed most of the hospitalized older patients and discussed their issues. We also like to thank them. They listened to us and talked so enthusiastically about the problems that they suffer during their hospitalized days, which encouraged us to work on this project and help them to have a better quality of life.

REFERENCES

[1] Benefits and harms of breast cancer screening: cohort study of breast cancer mortality and overdiagnosis. Wang S, Sultana F, Kavanagh A, 2023.
[2] Multi-Input Deep Learning Approach for Breast Cancer Screening Using Thermal Infrared Imaging and Clinical Data , Dennies Tsietso, Abid Yahya , 2023.
[3] Breast Cancer Detection and Classification, 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE), Poonam Kathale, Snehal Thorat, 2020.
[4] A Systematic Review on Screening, Examining and Classification of Breast Cancer, 2021 Smart Technologies, Communication and Robotics (STCR), Sannasi Chakravarthy S R, Harikumar Rajaguru, 2021
[5] Deep Learning to Improve Breast Cancer Detection on Screening Mammography , Li Shen, Laurie R. Margolies, Joseph H. Rothstein, Eugene Fluder, Russell McBride & Weiva Sieh, 2019
[6] Mammography with deep learning for breast cancer detection, Lulu Wang, 2024.
[7] Breast Cancer Detection Using Machine Learning Algorithms , Kumar Shubham, Dr. R. Kamalraj,2022.
[8] Breast histopathological image analysis using image processing techniques for diagnostic purposes:A methodological review ,Rashmi R , Keerthana Prasad,Chethana Babu K Udupa, 2021.
[9] Computer-aided breast cancer detection and classification in mammography: A comprehensive review, Kosmia Loizidou , Rafaella Elia, Costas Pitris,2021.
[10] Breast Cancer Detection and Prevention Using Machine Learning , Arslan Khalid , Arif Mehmood , Amerah Alabrah , Bader Fahad Alkhamees , Farhan Amin , Hussain AlSalman and Gyu Sang Choi,2023.
[11] Deep Learning in Breast Cancer Detection and Classification, Ghada Hamed(B) , Mohammed Abd El-Rahman Marey, Safaa El-Sayed Amin, and Mohamed Fahmy Tolba,2020.
[12] Enhancing Breast Cancer Diagnosis in Mammography: Evaluation and Integration of Convolutional Neural Networks and Explainable AI,Maryam Ahmed, Tooba Bibi, Rizwan Ahmed Khan , and Sidra Nasir,2024.
[13] Using hybrid pre-trained models for breast cancer detection , Sameh ZarifID1, Hatem Abdulkader , Ibrahim Elaraby , Abdullah Alharbi, Wail S. Elkilani , Paweł PławiakI,2023.