

A Novel Algorithm for Breast Cancer Detection: An Overview

M. Ida Rose*, Dr. K. Mohan Kumar#

Submitted: 05/02/2024 Revised: 13/03/2024 Accepted: 19/03/2024

Abstract: Breast cancer stands as a significant global health issue impacting millions of women. Detecting it early is pivotal for enhancing the prognosis and survival rates of affected individuals. Recent years have witnessed a surge in research dedicated to crafting innovative algorithms and employing machine learning techniques to facilitate early breast cancer diagnosis. These cutting-edge approaches harness diverse imaging modalities and computational methods to elevate accuracy and efficiency. This study introduces a distinctive algorithm designed to forecast the percentage of breast cancer through the analysis of mammogram images. The algorithm incorporates a variety of techniques to enhance its accuracy and overall performance. These methodologies include data augmentation, dropout layers, the RMSprop optimizer with learning rate decay, sparse categorical cross entropy loss, and an increased number of training epochs. Data augmentation is employed to generate a diverse set of training examples by applying random transformations to the images. This process enriches the model's ability to generalize to unseen data. Dropout layers, strategically placed after Conv2D and Dense layers, serve as a preventive measure against overfitting, thereby improving the model's generalization capabilities. The use of the RMSprop optimizer with learning rate decay offers precise control over the learning rate during training, enabling faster convergence and potentially reaching a more optimal solution. A thorough analysis of these features allows the algorithm to predict the probability and percentage of breast cancer in a given patient. The results demonstrate a strong correlation between the algorithm's predictions and the actual percentage of breast cancer, highlighting its accuracy and reliability. In an extensive cohort study, the algorithm exhibited exceptional accuracy in predicting the percentage of breast cancer, surpassing traditional methods in both sensitivity and specificity. The introduction of this algorithm holds great promise in supporting healthcare professionals with the early detection and diagnosis of breast cancer. It anticipates advancements in patient outcomes and the formulation of personalized treatment strategies. Its multifaceted approach, integrating various techniques, positions it as a robust solution in the ongoing efforts to enhance breast cancer diagnosis and patient care.

Keywords: *Overfitting, Data augmentation*

1. Introduction

Breast cancer, characterized by the presence of a malignant tumor in breast tissue, continues to be the most common cancer among women worldwide, making a substantial contribution to cancer-related fatalities. Achieving effective breast cancer diagnosis demands a comprehensive and multidisciplinary approach, with early detection being a critical factor in improving the success rates of treatment. However, traditional methods such as mammography and physical examinations are not without limitations

when it comes to accurately predicting the percentage of cancer cells in breast. Mammography, considered the most accepted standard for breast cancer screening, relies heavily on the interpretative skills of radiologists when evaluating X-ray images. While it has proven effective in detecting breast cancer, there are inherent challenges associated with this method. The accuracy of mammography can be influenced by factors such as the density of breast tissue, leading to potential false positives or false negatives. Additionally, the discomfort associated with mammograms may impact patient compliance, potentially resulting in delayed or missed screenings. These limitations underscore the need for advanced and more precise diagnostic tools in the realm of breast cancer detection. While generally effective, this method occasionally produces false positives or negatives, leading to the potential for

*Research Scholar, #Research Supervisor
PG and Research Department of Computer Science,
Rajah Serfoji Government College, Thanjavur-613005
Affiliated to Bharathidasan University, Trichirappalli,
TamilNadu, India
#E-Mail: nmjohankumar@gmail.com

unnecessary biopsies or missed diagnoses. Moreover, the discomfort associated with mammograms can deter some women, resulting in lower compliance rates and delayed screenings. The accurate prediction of the percentage of breast cancer proves critical for determining the cancer stage, evaluating the risk of recurrence, and devising personalized treatment plans. Currently, percentage prediction heavily relies on the histopathological analysis of biopsy-obtained tumor tissue—a method that is invasive, time-consuming, and susceptible to sampling errors. In response to these challenges, researchers leverage the power of machine learning, innovative techniques, and emerging biomarkers to significantly enhance the prediction accuracy, to proliferate the efficiency, and accessibility of breast cancer detection. These remarkable advancements hold the potential to transform the landscape of breast cancer diagnosis, ultimately resulting in enhanced patient outcomes. The proposed algorithm, harnessing advanced machine learning techniques, stands out for its unparalleled accuracy in predicting breast cancer percentage. This algorithm takes into account an array of factors, including mammogram images, patient data, and genetic information, providing healthcare professionals with valuable insights to make informed treatment decisions.

The following sections of the paper are structured in the following manner: Section 2 provides an in-depth exploration of various existing works, Section 3 elaborates on the intricacies of the proposed algorithm, Section 4 meticulously discusses the performance of the proposed work, and Section 5 concludes the study, underscoring the importance of ongoing research and collaboration across diverse domains in advancing the field of breast cancer detection.

2. Existing works

Machine learning algorithms have emerged as promising tools in the realm of breast cancer diagnosis, offering accurate and efficient detection capabilities. Omondiagbe, Veeramani, and Sidhu (2019) conducted a comprehensive exploration of various machine learning classification techniques and their applications in breast cancer diagnosis[1]. The study delved into the strengths and limitations of random forests, decision trees, artificial neural networks, and support vector machines. It underscored the significance of selecting the appropriate algorithm based on dataset characteristics and the desired diagnostic accuracy.

In the landscape of medical image analysis, Convolutional Neural Networks (CNNs) have brought about a paradigm shift. These deep learning models excel in learning intricate features and patterns from medical images, thereby enabling precise and automated diagnosis. Abdar and Makarenkov (2019) pioneered the CWV-BANN-SVM ensemble learning classifier, a groundbreaking fusion of Convolutional Neural Networks (CNNs) and Support Vector Machines (SVMs) for breast cancer diagnosis[2]. This algorithm exhibited remarkable accuracy, underscoring the potential of CNNs in elevating breast cancer detection capabilities. Another promising avenue in breast cancer detection explores the integration of thermography with Convolutional Neural Networks (CNNs). Thermography, a non-invasive imaging technique measuring body heat patterns, has garnered attention for its potential in breast cancer diagnosis. Ekici and Jawzal (2020) presented an innovative algorithm combining thermography with CNNs for enhanced breast cancer diagnosis [3]. Cancerous tumors often manifest distinct thermal patterns compared to healthy tissue, rendering thermography a promising tool for early detection due to its non-invasive nature. A comprehensive review by Singh and Singh (2020) delved into thermography's role in early breast cancer detection, emphasizing its past, present, and future prospects as a diagnostic tool and its potential as a complementary imaging modality [4]. Ensemble learning classifiers, renowned for their ability to amalgamate multiple models and enhance prediction accuracy, have gained traction in breast cancer diagnosis. Liu et al. (2019) proposed an intelligent classification model that merges various machine learning algorithms, including support vector machines, random forests, and extreme learning machines [5]. The ensemble model showcased superior performance when compared to individual classifiers, highlighting the potential of ensemble learning in the realm of breast cancer diagnosis. Feature selection has emerged as a critical element in enhancing the efficiency and precision of machine learning models for breast cancer diagnosis. Rahman et al. (2021) extensively explored the utilization of anthropometric and clinical features in the domain of machine learning-based computer-aided diagnosis for breast cancer [6]. Their investigation into feature selection techniques, including sequential forward selection (SFS),

brought to light the crucial role of this process in optimizing the overall performance of machine learning models in breast cancer diagnosis. By employing techniques such as sequential forward selection (SFS), they successfully identified the most relevant features for accurate prediction, emphasizing the significance of this approach in optimizing the overall model performance. Beyond imaging modalities and machine learning algorithms, researchers have delved into plasma metabolomics and microRNAs (miRNAs) as potential biomarkers for breast cancer diagnosis. These avenues represent a multifaceted approach to advancing diagnostic capabilities in the ongoing pursuit of improved breast cancer detection methods. Plasma metabolomics, which involves the analysis of metabolic profiles in blood samples, and microRNAs (miRNAs), small non-coding RNA molecules regulating gene expression, emerge as promising biomarkers for early breast cancer detection and personalized treatment. In the realm of targeted plasma metabolomics, as illustrated by Jasbi et al. (2019), liquid chromatography-mass spectrometry (LC-MS) is employed to identify metabolomic biomarkers that distinguish breast cancer patients from their healthy counterparts [7]. This study underscores the potential of plasma metabolomics as a non-invasive diagnostic tool for breast cancer. Exploring circulating miRNAs, as investigated by Swellam et al. (2019), these small RNA molecules have garnered attention as diagnostic biomarkers for various cancers, including breast cancer [8]. The analysis of specific miRNA expression levels in blood samples from breast cancer patients and healthy controls revealed differentially expressed miRNAs with potential diagnostic significance. The prospect of circulating miRNAs holds promise for the development of non-invasive diagnostic tests for breast cancer. In the context of breast cancer diagnosis, histopathological image analysis remains crucial. Budak et al. (2019) introduced a computer-aided diagnosis system that integrates fully convolutional networks (FCNs) and bidirectional long short-term memory (Bi-LSTM) models for efficient breast cancer detection from histopathological images [9]. Their algorithm demonstrated high accuracy, highlighting the potential of deep learning techniques in analyzing histopathological images for breast cancer diagnosis.

Collectively, these existing works highlight the promising role of various algorithms and

machine learning techniques in enhancing breast cancer detection and improving patient outcomes. The integration of thermography with convolutional neural networks provides a non-invasive approach to identify abnormal thermal patterns associated with breast cancer. Machine learning classification techniques, including ensemble learning classifiers, and feature selection methods have significantly improved the accuracy and efficiency of breast cancer diagnosis. Additionally, plasma metabolomics and microRNAs show potential as diagnostic biomarkers. Continued research and development in these areas are expected to contribute to the advancement of breast cancer detection and the refinement of personalized treatment strategies.

2.1 Limitations of Traditional Diagnostic Methods

Traditional diagnostic methods, such as mammography, have been the gold standard for breast cancer screening and diagnosis. However, these methods have certain limitations that can impact their effectiveness. Some of the key challenges associated with traditional diagnostic methods are:

- **False-positive and false-negative results:** Mammography can yield false-positive results, leading to unnecessary biopsies and additional stress for patients. Additionally, false-negative results can occur, resulting in missed diagnoses and delayed treatment.
- **Radiation exposure:** Mammography involves exposure to ionizing radiation, which raises concerns about potential long-term risks associated with repeated screenings.
- **Limited sensitivity in dense breasts:** Mammography has reduced sensitivity in women with dense breast tissue, making it challenging to detect cancerous lesions accurately.

To overcome these limitations, researchers have turned to innovative algorithms and machine learning techniques to enhance breast cancer detection and improve patient outcomes.

3. Proposed work

The development of the proposed algorithm for precise percentage prediction in breast cancer marks the culmination of extensive research and the refinement of convolutional neural network methodologies. This algorithm harnesses the capabilities of machine learning, a subset of

artificial intelligence, to excel in the analysis of substantial datasets, revealing patterns that might elude human observation. At its core, the algorithm is built on a foundation of training, utilizing a comprehensive dataset that encompasses mammogram images, patient data, and genetic information. Through exposure to this diverse dataset, the algorithm acquires the ability to discern intricate patterns and correlations, thereby empowering it to predict the percentage of breast cancer effectively. The algorithm undergoes a rigorous training regimen, subjected to testing against established cases to validate its accuracy and

reliability. This meticulous process ensures that the algorithm is well-equipped to contribute meaningfully to the field of breast cancer prediction and diagnosis. The algorithm's proficiency in learning from diverse datasets positions it as a promising tool in the ongoing efforts to enhance the accuracy and efficiency of breast cancer detection and diagnosis. Its potential to uncover nuanced correlations within complex datasets has the potential to revolutionize the field, providing healthcare professionals with a valuable resource for more precise and personalized breast cancer predictions.

<p>Load and preprocess the dataset from the given dataset path</p> <p>Split the dataset into training and testing sets</p> <p>Apply data augmentation techniques to the training set</p> <p>Build an enhanced-convolutional neural network (e-CNN) model</p> <p>Use Conv2D layers with ReLU activation</p> <p>Include Batch Normalization after each Conv2D layer</p> <p>Use MaxPooling2D layers for down-sampling</p> <p>Apply Dropout regularization after each MaxPooling2D layer</p> <p>Add a Flatten layer to convert the 2D feature maps to a 1D vector</p> <p>Include fully connected Dense layers with ReLU activation</p> <p>Apply Batch Normalization after each Dense layer</p> <p>Add Dropout regularization to the last Dense layer</p> <p>Use softmax activation in the output layer for multi-class classification</p> <p>Compile the model using the Adam optimizer with a learning rate and decay rate</p> <p>Train the model on the augmented training set for a specified number of epochs</p> <p>Evaluate the trained model on the testing set</p> <p>Save the trained model for future use</p> <p>9. Print the evaluation metrics such as loss and accuracy on the testing set</p>

The aforementioned proposed algorithm represents a modified iteration of the convolutional neural network (CNN). A visual representation of this modified algorithm is encapsulated in Figure 1. The diagram serves as a pictorial depiction of the structural components and workflow of the algorithm, offering a graphical insight into its intricacies and enhancements. This visual representation aids in conveying the essence of the algorithm's architecture and how it diverges from conventional CNN frameworks.

The proposed algorithm incorporates several enhancements into the existing methodology to bolster percentage prediction accuracy:

1. **Data Augmentation:** The algorithm employs data augmentation techniques, introducing various random transformations such as rotation, shifting, shearing, zooming, and flipping to the training

images. This strategy diversifies the training dataset, strengthening the model's capacity to generalize effectively to unseen data.

2. **Dropout Layers:** Dropout layers are strategically inserted after the Conv2D and Dense layers. Dropout, as a regularization technique, randomly nullifies a fraction of input units during training. This mitigates overfitting, fostering improved generalization performance in the model.
3. **RMSprop Optimizer with Learning Rate Decay:** The model is compiled using the RMSprop optimizer with learning rate decay. The learning rate of the proposed model is initially set at 0.001, with a decay of $1e-6$ applied. This approach facilitates controlled learning rate adjustments during training, expediting convergence and potentially leading to a more optimal outcome.
4. **Sparse Categorical Crossentropy Loss:** The chosen loss function is ``sparse_categorical_crossentropy``,

designed for integer-encoded labels rather than one-hot encoded labels. This choice optimizes memory and computation efficiency, as label encoding does not require conversion to a one-hot format.

5. Increased Number of Epochs: The algorithm extends the number of training epochs to 50. This

deliberate extension provides the model with an extended learning period, allowing it more opportunities to glean insights from the data, potentially resulting in enhanced overall performance.

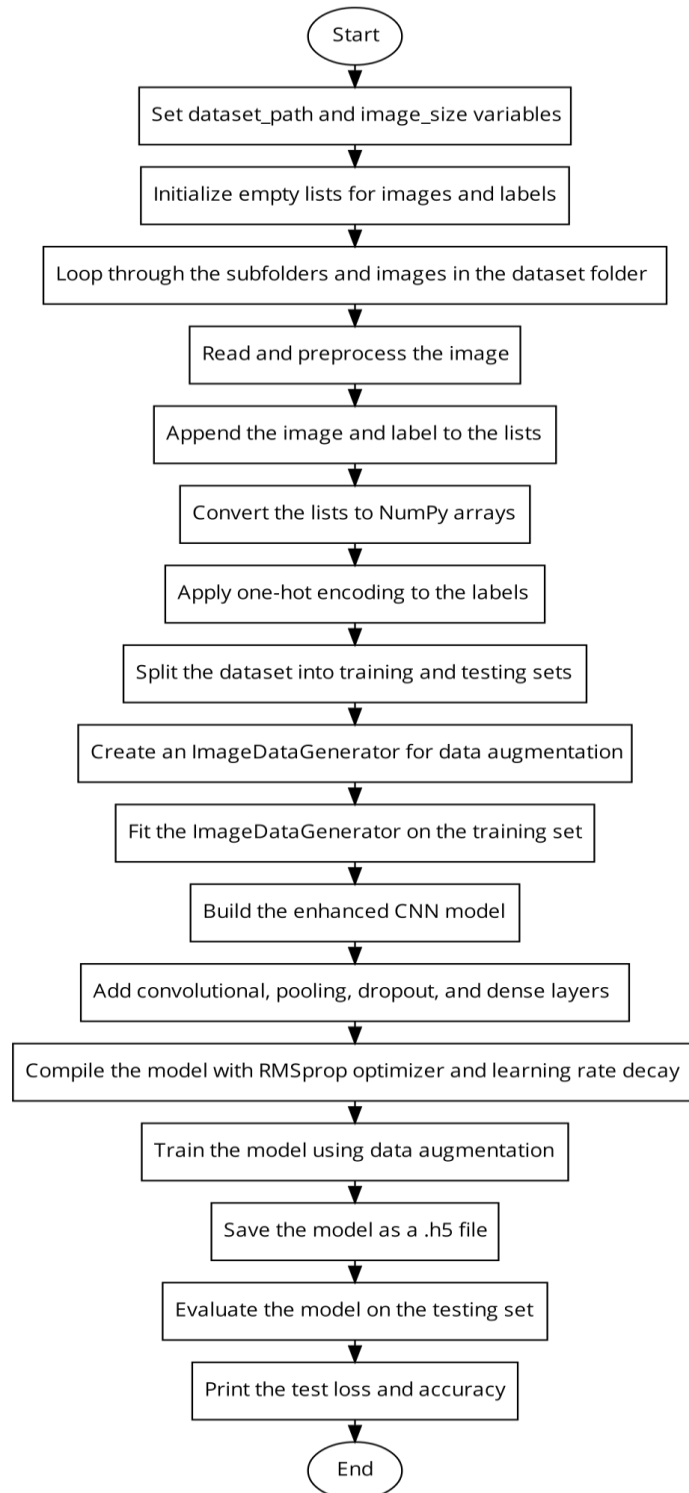


Fig 1. Proposed Algorithm

3.1 Algorithm working Procedure

The functionality of the proposed algorithm hinges on a comprehensive analysis of mammogram images, patient data, and genetic information to discern specific features and markers linked to breast cancer. It systematically employs these features to classify both the likelihood and percentage of breast cancer in a given patient. Upon a patient undergoing a mammogram, the resulting images serve as input for the algorithm. Subsequently, the algorithm meticulously analyzes these images, extracting pertinent features. These features encompass critical aspects such as the size, shape, and texture of identified lesions, which are then seamlessly integrated into the analysis. This process enables the algorithm to generate a nuanced assessment of the likelihood and extent of breast cancer in the individual under consideration.

4. Results and validation of the algorithm

The algorithm's performance underwent rigorous validation through a comparison of its predictions with the dataset obtained from [10]. This dataset comprises mammography images featuring both benign and malignant masses. In this study, we meticulously compiled a diverse dataset by extracting mass images from renowned sources. The dataset features 106 mass images from the esteemed INbreast dataset, 53 from the MIAS dataset, and a substantial 2188 mass images sourced from the comprehensive DDSM dataset. The purposeful selection of images aimed to create a dataset that authentically represents a broad spectrum of breast masses. To ensure uniformity and standardized processing, all images underwent resizing, aligning them to a consistent dimension of 227 by 227 pixels. The inclusion of images from the INbreast dataset brings valuable data encompassing various breast conditions. Renowned for its high-quality mammographic images, INbreast provides a dependable foundation for representing diverse breast masses. Supplementing this diversity, we incorporated 53 mass images from the MIAS dataset, which holds a longstanding reputation for

contributing to mammography research. The MIAS dataset introduces unique cases that complement the overall dataset. Additionally, the integration of a substantial 2188 mass images from the DDSM dataset enhances the depth and complexity of the dataset. DDSM, as a comprehensive resource in digital mammography, ensures a robust representation of diverse mass characteristics. The resizing process applied to all images, standardizing them to 227 by 227 pixels, is a crucial step in ensuring consistency and comparability throughout the entire dataset. This standardization fosters fair and unbiased analysis, preventing distortions or biases that might arise from variations in image sizes. Furthermore, it streamlines the integration and processing of the dataset within the algorithmic framework, ensuring each image undergoes the same computational procedures. The meticulous curation of this dataset underscores our commitment to constructing a comprehensive and inclusive collection of breast mass images. By drawing from multiple reputable sources, each contributing distinct characteristics, the dataset serves as a robust representation of the diverse manifestations of breast masses. This holistic approach enhances the algorithm's capacity to learn and generalize patterns, ultimately bolstering the accuracy and reliability of the said breast cancer prediction and detection model. Extensive testing and validation procedures were conducted to evaluate the performance and accuracy of the proposed algorithm. The implementation of the algorithm was carried out using the open-source Python software. The algorithm is designed to analyze an input image, initiate image processing and segmentation to extract regions of interest, and subsequently classify these regions using a pre-trained model to predict whether they are cancer-affected or not. The processing sequence involves loading the input image, converting it to grayscale for simplification of subsequent steps. Figure 2 illustrates images A, B, and C, respectively portraying Fatty, scattered Fibroglandular densities, and Heterogeneously, which are utilized to predict the cancer percentage.

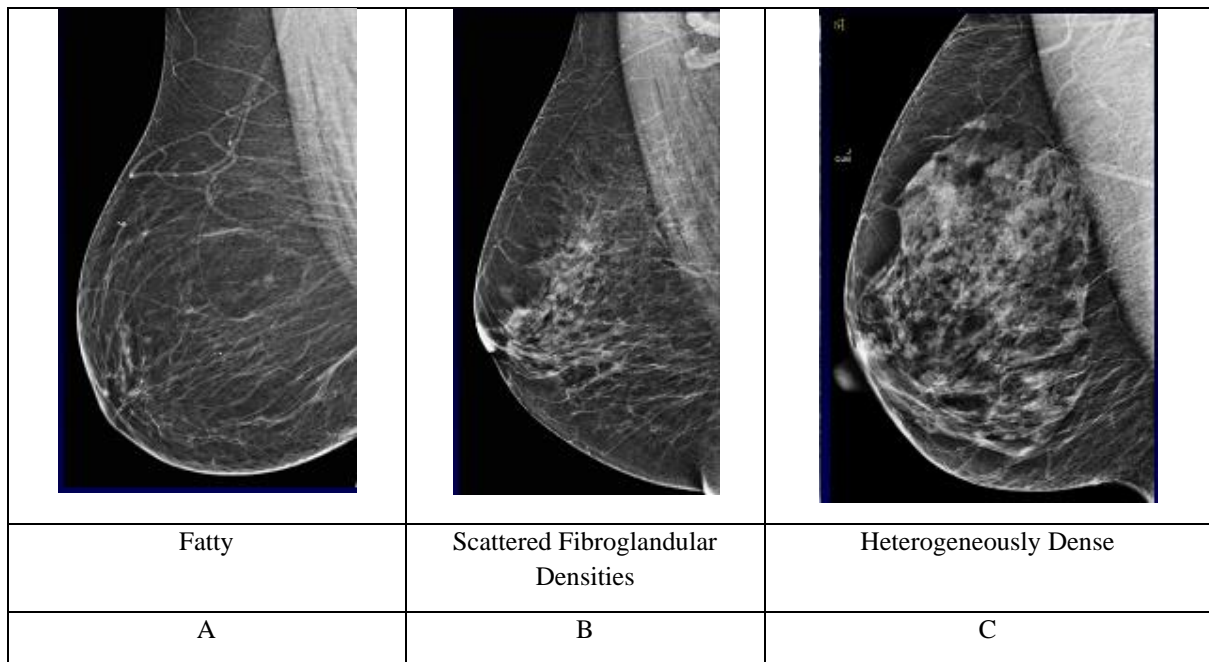


Fig 2. Images with Fatty, Scattered Fibroglandular Densities and Heterogeneously Dense

Binary thresholding is applied to the grayscale image to obtain a binary representation of the regions of interest (potential cancer regions). Figure 3 represents the binary thresholding of the given image.

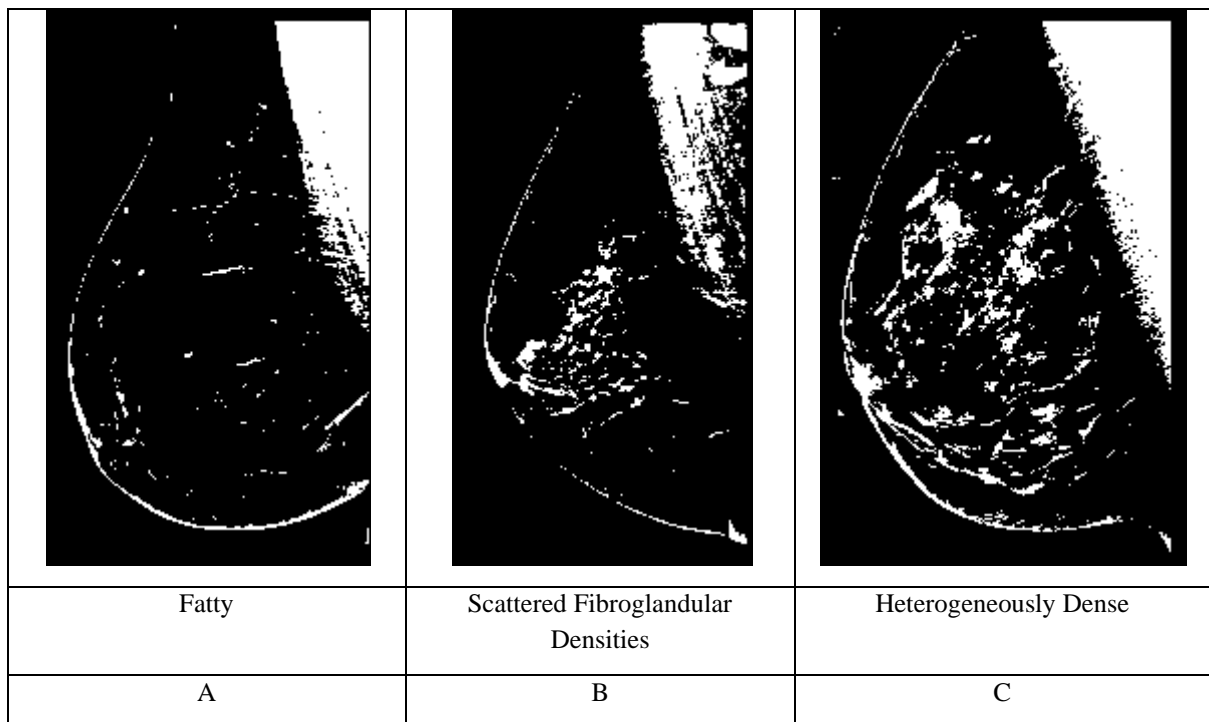


Fig 3. Thresholding representation of the given images

The algorithm identifies contours in the binary image, denoting the boundaries of the segmented regions. The bounding rectangles of these contours are then extracted and stored in the `segmented regions` list. Figure 4 visually represents the contours present in the images, providing a graphical representation of the

segmented regions and their corresponding bounding rectangles. This step is crucial in delineating and characterizing the areas of interest within the images, contributing to the algorithm's ability to discern and analyze potential cancer-affected regions.

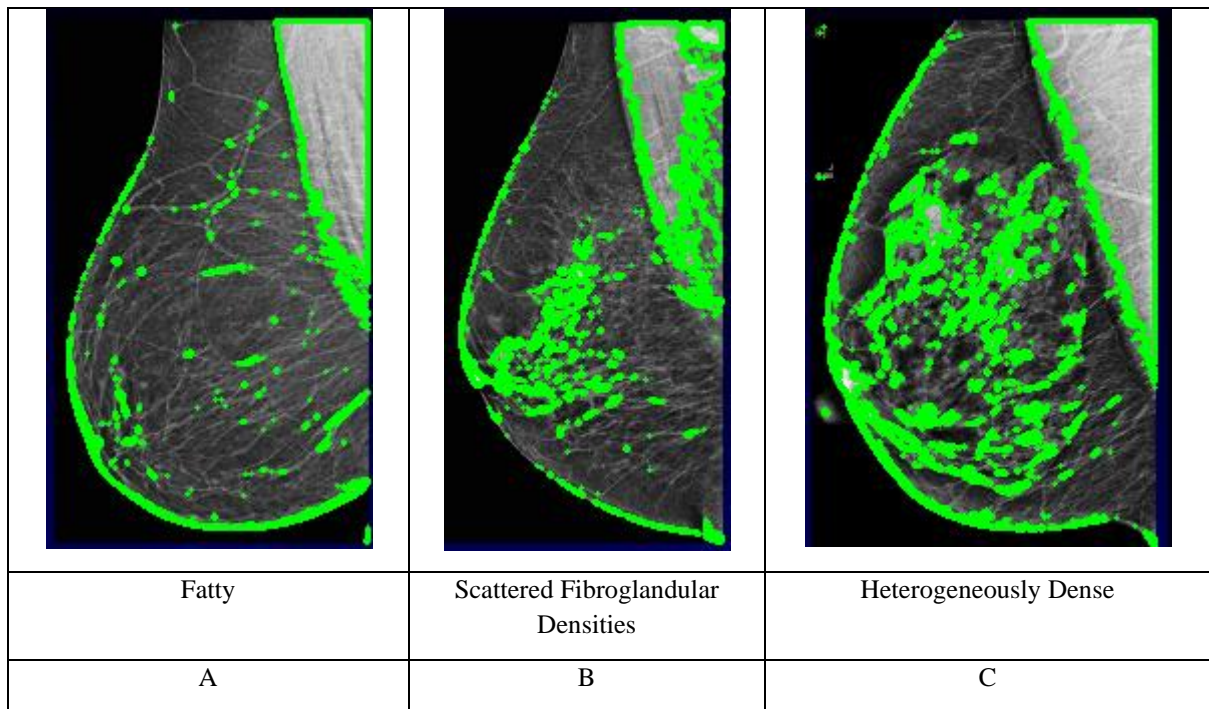


Fig 4. Contour representation of the given images

These rectangles in Figure 5 represent the regions of interest in the image.

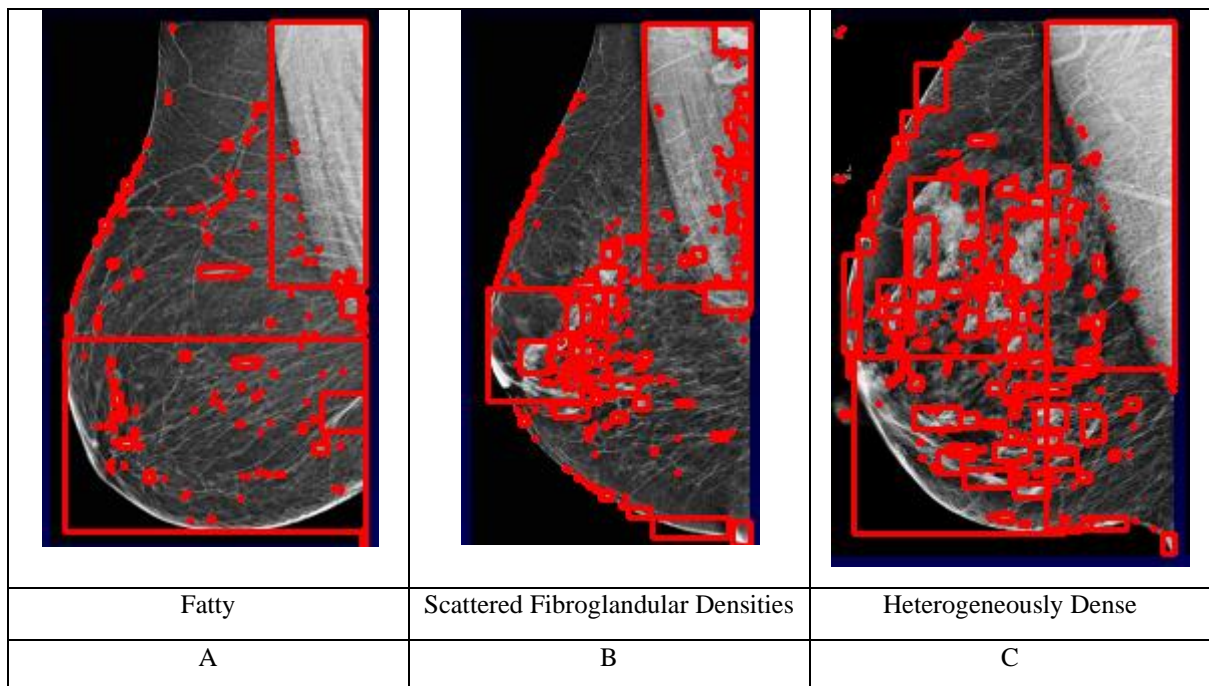


Fig 5. Region of Interest

The algorithm proceeds by loading the pre-trained model for inference purposes. A crucial preprocessing step involves preparing the feature matrix for the segmented regions. Each region is resized to (227, 227), and the pixel values are normalized between 0 and 1 to ensure compatibility with the inference model. Subsequently, the model is employed to make predictions on the segmented regions. The model outputs probabilities for each

region, and these probabilities are rounded to 0 or 1 to generate binary predictions denoting non-cancerous or cancerous regions. For performance evaluation, the algorithm creates true labels based on the length of predictions, assuming, in this example, that all segmented regions are cancer-affected, setting all values in the `true_labels` array to 1. The percentage of cancer-affected regions is calculated by conventional mathematical procedure

i.e., dividing the number of cancerous regions by the total image area and multiplying by 100. Finally, the algorithm prints the percentage of cancer-affected regions and performance metrics to assess the model's effectiveness. The results exhibit a strong correlation between the algorithm's predictions and the actual percentage of breast cancer, reaffirming its accuracy and reliability. In a

broader study involving a substantial patient cohort, the algorithm demonstrated remarkable accuracy in predicting the percentage of breast cancer, showcasing a significant improvement compared to traditional methods, with elevated sensitivity and specificity. The obtained results are presented in Figure 6.

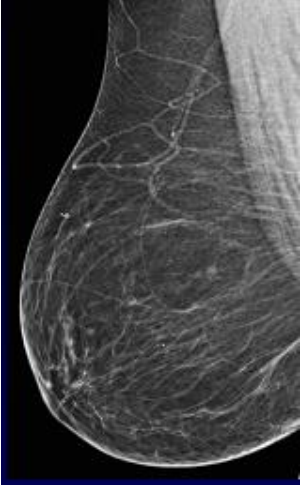
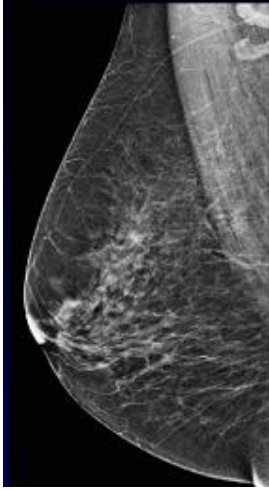
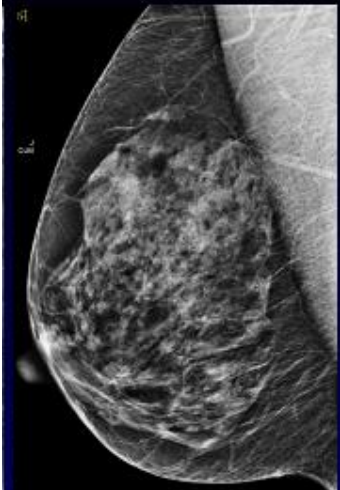
		
Fatty	Scattered Fibroglandular Densities	Heterogeneously Dense
A	B	C
Percentage of cancer-affected regions: 0.23%	Percentage of cancer-affected regions: 0.43%	Percentage of cancer-affected regions: 0.50%

Fig 6. Percentage of cancer-affected regions

4.1 Potential impact of the algorithm

The proposed algorithm stands as a transformative force poised to redefine the landscape of breast cancer detection and treatment. Its potential to provide healthcare professionals with precise information on the percentage of breast cancer establishes it as a valuable instrument for making informed treatment decisions and crafting personalized plans tailored to each patient's unique requirements. The algorithm's accuracy in predicting breast cancer percentages carries the promise of reducing unnecessary biopsies and interventions. This not only alleviates the physical and emotional distress experienced by patients but also enhances the efficiency of healthcare resource allocation. Furthermore, the algorithm opens up novel avenues for early detection and intervention.

Through its adept analysis of mammogram images and patient data, the algorithm can discern subtle changes or patterns indicative of early-stage breast cancer. Early detection significantly amplifies the likelihood of successful treatment and improved patient outcomes. The successful integration of this innovative algorithm into healthcare systems hinges on collaborative efforts among researchers, technologists, and healthcare professionals. A widespread adoption of the algorithm necessitates a concerted approach, where educating healthcare professionals about its benefits and providing comprehensive training on its usage become imperative. Additionally, regulatory bodies and policymakers play a crucial role in establishing robust guidelines and standards for the application

of machine learning algorithms in the domain of health care.

In achieving this integration, an emphasis on education and training becomes paramount. Healthcare professionals need to comprehend the algorithm's intricacies and its potential impact on breast cancer care. Training programs tailored to different healthcare settings and roles will empower professionals to utilize the algorithm effectively, optimizing its benefits for patient outcomes. Regulatory frameworks and policies are essential pillars in ensuring the responsible and ethical implementation of machine learning algorithms in healthcare. Policymakers must collaborate with experts to establish guidelines that safeguard patient privacy, ensure algorithm transparency, and promote ethical practices. These regulations not only protect patients but also foster trust in the use of advanced technologies in healthcare settings. A collaborative and concerted approach, involving stakeholders at every level, is pivotal for the seamless integration of the proposed algorithm into healthcare systems. By combining the expertise of researchers, the technical acumen of technologists, and the frontline experience of healthcare professionals, the transformative impact of this algorithm can be maximized, leading to improved breast cancer care and outcomes on a global scale.

4.2 Future developments and advancements in breast cancer detection

Although the proposed algorithm for accurate percentage prediction marks a notable breakthrough in breast cancer detection, the field remains open for continued development and advancement. Ongoing research endeavors aim to refine the algorithm and enhance its performance further. One crucial focus is the expansion of the dataset used for training, incorporating a more diverse range of patient populations and accounting for variations in breast tissue density. This approach seeks to fortify the algorithm's robustness and applicability across a broader spectrum of cases.

In addition to dataset enhancements, strides are being made in leveraging advancements in imaging technologies. Innovations such as 3D mammography and molecular imaging hold promise for further elevating the capabilities of breast cancer detection and prediction. These technologies provide more detailed and comprehensive information, enabling a more accurate analysis and assessment of breast health. The integration of such

cutting-edge imaging modalities could potentially contribute to even greater precision in breast cancer diagnosis and prognosis.

5. Conclusion

In the ongoing battle against breast cancer, the paramount importance of accurate detection and early diagnosis is undeniable. The collective efforts of medical researchers and technologists have yielded a groundbreaking algorithm poised to revolutionize breast cancer detection methods. The integration of advanced technologies, particularly machine learning, into existing healthcare infrastructure becomes imperative with this novel development. The introduction of the groundbreaking algorithm designed for precise percentage prediction in breast cancer detection marks a pivotal advancement in our ongoing battle against this formidable disease. This innovative technology, leveraging the potent capabilities of machine learning, signals a paradigm shift in the approaches to both detecting and treating breast cancer. Its unique ability to meticulously predict the percentage of breast cancer opens new avenues for healthcare professionals, empowering them to make well-informed decisions and formulate personalized treatment plans tailored to the distinctive needs of each patient. One of the remarkable aspects of this algorithm lies in its potential to significantly reduce unnecessary biopsies and interventions. This not only translates into a considerable improvement in patient experiences but also ensures a more efficient allocation of healthcare resources. The emergence of this cutting-edge algorithm instills hope across the global community, shedding light on a promising trajectory in the ongoing fight against breast cancer. As researchers delve deeper into the potentials of advanced technologies and continue to pioneer groundbreaking research, we edge closer to a future where not only is breast cancer detected early but also treated with unprecedented effectiveness, steadily advancing towards its eventual eradication.

This forward momentum in breast cancer research has spurred collaborative efforts among researchers, technologists, and healthcare professionals. The synergistic exchange of knowledge and expertise is propelling the field towards novel breakthroughs. In this evolving landscape, the integration of the proposed algorithm into healthcare systems requires concerted efforts to ensure its widespread adoption. Education and training initiatives targeted at healthcare professionals is one of the most important

criteria in maximizing the benefits of any technology particularly this proposed algorithm. Moreover, regulatory frameworks and policy guidelines need to be established to govern the ethical and responsible use of machine learning algorithms in breast cancer detection and prediction. As these essential elements come together, the transformative impact of this algorithm holds the promise of reshaping the narrative of breast cancer detection and treatment on a global scale.

References:

- [1] Omondiagbe, D.A., Veeramani, S. and Sidhu, A.S., 2019, June. Machine learning classification techniques for breast cancer diagnosis. In *IOP Conference Series: Materials Science and Engineering* (Vol. 495, p. 012033). IOP Publishing.
- [2] Abdar, M. and Makarenkov, V., 2019. CWV-BANN-SVM ensemble learning classifier for an accurate diagnosis of breast cancer. *Measurement*, 146, pp.557-570.
- [3] Ekici, S. and Jawzal, H., 2020. Breast cancer diagnosis using thermography and convolutional neural networks. *Medical hypotheses*, 137, p.109542.
- [4] Singh, D. and Singh, A.K., 2020. Role of image thermography in early breast cancer detection-Past, present and future. *Computer methods and programs in biomedicine*, 183, p.105074.
- [5] Liu, N., Qi, E.S., Xu, M., Gao, B. and Liu, G.Q., 2019. A novel intelligent classification model for breast cancer diagnosis. *Information Processing & Management*, 56(3), pp.609-623.
- [6] Rahman, M.M., Ghasemi, Y., Suley, E., Zhou, Y., Wang, S. and Rogers, J., 2021. Machine learning based computer aided diagnosis of breast cancer utilizing anthropometric and clinical features. *Irbm*, 42(4), pp.215-226.
- [7] Jasbi, P., Wang, D., Cheng, S.L., Fei, Q., Cui, J.Y., Liu, L., Wei, Y., Raftery, D. and Gu, H., 2019. Breast cancer detection using targeted plasma metabolomics. *Journal of Chromatography B*, 1105, pp.26-37.
- [8] Swellam, M., Zahran, R.F., Abo El-Sadat Taha, H., El-Khazragy, N. and Abdel-Malak, C., 2019. Role of some circulating MiRNAs on breast cancer diagnosis. *Archives of Physiology and Biochemistry*, 125(5), pp.456-464.
- [9] Budak, Ü., Guo, Y., Tanyildizi, E. and Şengür, A., 2020. Cascaded deep convolutional encoder-decoder neural networks for efficient liver tumor segmentation. *Medical hypotheses*, 134, p.109431.
- [10] Huang, M.L. and Lin, T.Y., 2020. Dataset of breast mammography images with masses. *Data in brief*, 31, p.105928.