© November 2024 | IJIRT | Volume 11 Issue 6 | ISSN: 2349-6002

Analyzing of Mammography Image Breast Cancer Detection Using CNN and Feature Selection

Sreeji K B[1], Dr.R.Khanchana[2]
[1]*Research Scholar, Sri Ramakrishna Arts and Science for Women, Coimbatore*
[2]*Associate Professor, Sri Ramakrishna Arts and Science for Women, Coimbatore*

Abstract: The prognosis of breast cancer is greatly influenced by the timely and precise identification of breast lesions, including the differentiation between suspicious, non-cancerous, and cancerous cancer. In this work, we present a brand-new feature extraction and reduction technique for the detection of breast cancer in pictures from mammography. First, we extract features from multiple pre-trained convolutional neural network (CNN) models, and then concatenate them. The most informative features are selected based on their mutual information with the target variable. Subsequently, the selected features can be classified using a machine learning algorithm. We evaluate our approach using four different machine learning algorithms: neural network (NN), k-nearest neighbor (kNN),random forest (RF), and support vector machine (SVM).The dataset is newly introduced and includes two views as well as additional features like age, which contributed to the improved performance. We compare our proposed algorithm with state-of-the-art methods and demonstrate its superiority, particularly in terms of accuracy and sensitivity.

Keywords: breast cancer; convolutional neural network (CNN);feature selection; feature classification; mammography images

1. INTRODUCTION

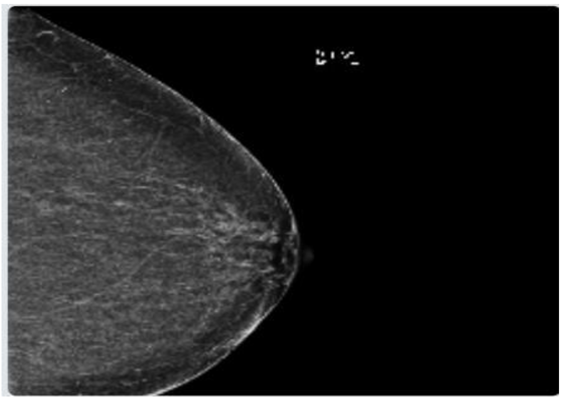
Every year, millions of people are diagnosed with breast cancer (BC), a common type of disease that also claims millions of lives [1]. 2.3 million new cases of breast cancer were diagnosed in 2020 alone, and 685,000 people died from the disease [2]. Even if death rates have decreased as a result of the adoption reducing cancer deaths still requires early discovery, treatment, and routine mammography screening [3]. At the moment, highly skilled radiologists are needed for the early diagnosis of BC using radiological imaging. A looming shortage of radiologists in several countries will likely worsen this problem [4].False positive results from mammography screening are also common. This may lead to needless worry, difficult follow-up care, more imaging tests, and occasionally the requirement for tissue biopsy (usually a biopsy with a needle) [5,6]. Furthermore, the evaluation of multiple-view radiological pictures based on graph-based clustering algorithms may be enhanced by machine learning techniques [7–10]. The interpretation of diagnostic imaging investigations has been completely transformed in recent years by deep learning, a subset of machine learning [11]. One of the most important networks in the field of deep learning is a convolutional neural network (CNN) [12].In contrast to conventional methods of screening, computer-aided diagnostic (CAD) systems Convolutional neural networks (CNN) provide screening that is quicker, more dependable, and more resilient. In image analysis, CNNs have become a popular technique for pattern recognition [13]. CNN has been widely utilized to diagnose breast cancer in a variety of breast cancer picture modalities, including X-ray, magnetic resonance imaging, and ultrasound (US) as follows:

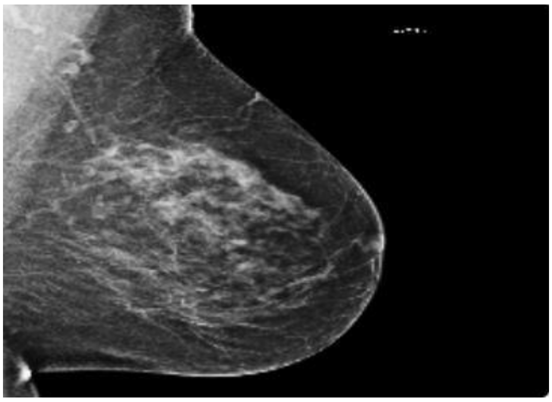
US Photos: By extracting features from Alexnet, MobilenetV2, and Resnet50, then concatenating them, Ero Æglu Y [14] developed a hybrid-based CNN system based on ultrasonography images for diagnosing BC. The best features were then chosen using the mRMR features selection approach. This system supported k-nearest neighbors (k-NN) and vector machine (SVM) as classifiers using machine learning methods. In [15], the breast US pictures were divided into sub-regions using an image segmentation technique, and then an object identification technique that uses methods for feature extraction, selection, and classification to automatically identify BC-related subregions. A technique for segmenting BC images using patch merging and semantic classification was proposed in [16]. A region of interest is cropped, enhanced using filters and clustering techniques, features are extracted, and classification is carried out using a neural network and a k-NN classifier.

MRI Images: Using a poorly supervised method, Zhou J et al. [17] developed a 3D deep CNN for the detection and localization of BC in dynamic contrast-enhanced MRI data, A multi-layer CNN was created in [18] to Use pixel information and online data augmentation to categorize MRI pictures as either benign or malignant tumors. To distinguish between benign and malignant mammography tumors, the authors in [19] employed pre-trained CNN models, InceptionV3 and ResNet50, on the DDSM dataset.Techniques like data augmentation, pre-processing, and transfer learning were employed because restricted information. A CNN model that integrates features from numerous mediolateral oblique (MLO) and craniocaudal (CC) views was employed by the authors in [20]. In [21], Ridhi Hela et al. presented a technique for BC detection based on the CBIS-DDSM image dataset. After pre-processing the images, several CNN models (AlexNet, VGG16, ResNet, GoogLeNet, and InceptionResNet) were used to extract features. Reducing false negatives is essential in the field of BC detection in order to guarantee correct diagnosis and avoid the possible harm that could result from missing positive instances. In order to improve the precision of BC detection, we specifically suggest a new CNN-based method in this study. Concentrating on databases of X-ray images. Our approach seeks to enhance overall detection accuracy and drastically lower false negatives by addressing the shortcomings of earlier studies.The creation of a sophisticated and trustworthy BC detection system has enormous potential to enhance patient outcomes and progress the field of medical imaging diagnostics two important additions to the body of literature are made by this study. First of all, it gathers a wide range of features from many pre-trained CNNs for various viewpoints. It also adds other characteristics, such as age, to produce a feature vector. Second, it uses a mechanism to eliminate weak features based on their mutual information with the ground truth, hence reducing the dimensionality of the feature vector.Alexnet, Resnet50, MobileNetS-mall, ConvNeXtSmall, and EfficienNet are the five basis models used in the suggested system. Their features are combined and extracted for using a neural network (NN) model to optimal categorization. This method showed that it could improve the BC classification's accuracy.

2. MATERIALS AND METHODS

The Radiological Society of North America (RSNA) dataset from a recent Kaggle competition serves as the primary dataset for this project [22]. Because it includes instances of both normal and abnormal mammograms, as well as a variety of breast densities and lesion types, the 54,713-image dataset is very useful for researchers working on machine learning methods for BC detection. Experts have identified and extracted the region of interest from a subset of this dataset known as CBIS-DDSM, which is for positive cases. In this study, we classify photos from cancer patients and normal participants using the original DDSM dataset rather than the CBIS-DDSM. This renders any classification system skewed. We use only 2320 photos from negative cases and all positive cases to make up for this. Two sample photos for cancer and normal cases from this collection are shown in Figure 1.





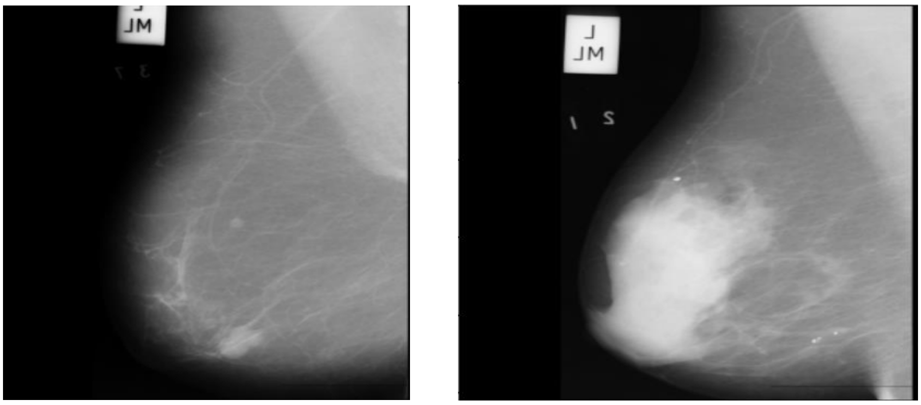
(a)

(b)

Figure 1. These figures show two sample images from the RSNA dataset for (a) a cancerous, and (b) a normal subject.

B. A popular and extensively used dataset for the creation and assessment of CAD systems for BC detection is the Mammographic Image Analysis Society (MIAS) [23] dataset. Each of the 322 mammograms in the set has a corresponding ground truth classification of either benign or malignant tumors. Because the dataset contains samples of both normal and abnormal mammograms, as well as a variety of breast densities and lesion kinds, it is especially useful for researchers working on machine learning methods for BC identification. Two sample photos for cancer and normal cases are shown in Figure 2 from this collection.

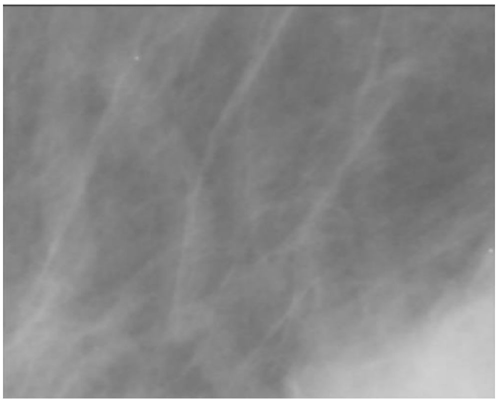
C. The Mammographic Image Analysis Society (MIAS) [23] dataset is a widely used and well-liked dataset for developing and evaluating CAD algorithms for BC identification. There is a corresponding ground truth classification of either benign or malignant tumors for each of the 322 mammograms in the set. The dataset is particularly helpful for researchers working on machine learning techniques for BC diagnosis since it includes examples of both normal and abnormal mammograms, as well as a range of breast densities and lesion types. Figure 2 displays two sample images from this collection, one for cancer and one for normal instances.



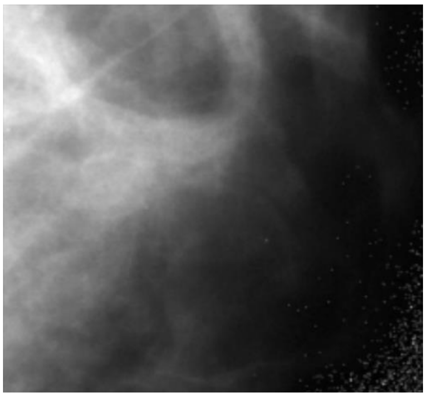
(a) (b)

Figure 2. These figures show two sample images from the MIAS dataset for (**a**) cancerous, and

(b) normal subjects.



(a)



(b)

Models

A. A deep learning architecture called MobileNet is appropriate for reliable and efficient medical picture processing, particularly when it comes to BC diagnosis. MobileNet's focus on computational efficiency allows it to efficiently extract features from mammography images, making it possible to identify minute patterns or anomalies linked to breast cancer. MobileNet optimizes memory usage and computational load by employing depthwise separable convolutions, which makes it perfect for environments with limited resources. Efficiency and compatibility with medical imaging devices are further improved by the use of the ReLU6 activation feature. All things considered, MobileNet is a useful tool for BC analysis that produces precise findings while using little processing power.

B. ConvNeXt is an architecture that improves performance on difficult visual identification tasks by utilizing parallel branches to collect a variety of complementary information, hence increasing the representational capacity of CNNs. It has proven to perform exceptionally well on a range of computer vision tasks, such as semantic segmentation, object identification, and image classification. It is a well-liked option for activities demanding a high level of comprehension of visual data because of its capacity to record intricate correlations between components.

Review of Literature:

Numerous techniques have been put forth for the diagnosis and detection of early-stage breast cancer, which is an important topic in the medical sector. The authors have covered many segmentation techniques, characteristics, and classifiers in this part on progressive review, along with the advantages and disadvantages of each for the identification of breast cancer in its early stages. Deep learning techniques for early-stage breast cancer detection have been studied, as has their importance in breast cancer diagnosis and detection, therapy, and recovery, as well as lowering the likelihood of critical conditions.

A median filtering approach was employed by Shrivastava et al. [1] to identify and eliminate noise from the original mammography image. The mammography image's cancerous areas were divided using the Seeded Region Growing method. The mammography pictures that were accessible in order to test this work MIAS data set. The researchers achieved a 94% detection rate and 96.1% accuracy.

Suresh et al.[2] classified and segmented abnormal cancer areas in mammography images using a deep learning algorithm. Mammography images were subjected to the Laplacian filter in order to identify and eliminate any noise. Next, energy features and the Histogram of Oriented Gradients features were calculated from the image of a denoised mammography. Deep neural networks are then used to classify these combined data, classifying the original mammography pictures as either normal or abnormal. Finally, the cancer regions on abnormal mammography images were segmented using a modified Adaptively Regularized Kernel-based Fuzzy-C-Means segmentation technique.

By combining machine learning and deep learning techniques, Ragab et al. [3]presented a method for identifying and dividing the cancerous regions in mammography pictures. In this study, the region of interest in abnormal mammography pictures was chosen using a threshold technique. Next, deep Neural networks with convolutions To extract the internal features, AlexNet was applied to the chosen region of interest. The Support Vector Machine (SVM) classification technique was then used to train and classify these internal features, dividing the mammography pictures into normal and pathological cases.

Using features that were retrieved using the Hough transform, He et al. [4]provided a classification of mammograms. A two-dimensional transformation is the Hough transform. It is employed to separate an image's elements related to a specific form. The two most important threat indicators are masses and characterisation at the miniature scale and the early detection of breast cancer greatly depends on their robotic identification. Computerized mass location and organization is also quite challenging since masses are frequently not distinct from the surrounding parenchyma. The methods for feature extraction and categorization were discussed in this paper. Here, characteristics of the mammography picture have been identified using the Hough transform, and SVM is utilized for classification. The SVM classifier has a higher classification accuracy.

Vijayarajeswari et al. [5] used deep learning and computer image processing technologies to automatically diagnose breast cancer on Whole Slide Images (WSI). This paper's primary tasks were (1) obtaining training and testing sets by extracting patches from WSIs. Train CNN-based classification models for cancer region classification. Use transfer learning and convolutional neural networks (CNNs) to train feature extraction network models. The SVM classifier is then trained using a combination of texture and CNN features. Cancer detection has made use of the categorization model. Additionally, the accuracy of this categorization system is better. (4) Using sliding windows and vote scoring, select the optimal classification strategy to automatically identify cancerous areas on large-scale pathological images.

Dhungel et al.[6] identified the suspicious mass locations in mammography pictures using a deep learning algorithm. To increase the quantity of mammography pictures and attain a high degree of classification accuracy, the scientists used the data augmentation procedure. Next, linear characteristics that are probabilistic were calculated from every picture during the data augmentation procedure, and a deep learning algorithm was used to train and classify them.

Chu et al.[7]used a straightforward linear iterative clustering technique to present a fully automatic breast cancer detection and segmentation methodology. To increase the average segmentation accuracy, the clutters found in the aberrant mammography pictures were eliminated.

Alsheikhy et al.[8] introduced a method that uses three classifiers and AlexNet to identify and categorize breast cancers. AlexNet was used to carry out the algorithm's deep learning phase, and the authors used mammography images. The system was trained on three Kaggle datasets. This approach's accuracy on the test dataset was 99.147%. The tumors it identified were categorized as either benign or malignant. To that end, a variety of image preparation techniques were used. To assess the deployed model, a number of performance measures were calculated.

Akkur et al.[9] introduced a unique classification model in that uses machine learning methods to identify breast tumors early on. Bayesian optimization methods and feature selection were employed in this model. Two datasets were subjected to several machine learning techniques: a Decision Tree, Naive Bayes, Support Vector Machine (SVM), KNN, and Ensemble Learning. The Mammographic Breast Cancer Dataset (MBCD) and the Wisconsin Breast Cancer Dataset (WBCD) were these databases. To identify the most pertinent traits, the authors used the Least Absolute Shrinkage and Selection Operator (LASSO).

Umer et al.[10] developed a framework in that detects breast tumors by using deep learning and selection based on histology images. Approximately 277,000 images were used. The procedure involved three steps: A novel features selection method that extracted features using deep learning separating the data into two groups—invasive ductal carcinoma and normal—using machine learning techniques and a deep learning tool.

Wang et al.[11] used four datasets to train their algorithm in order to develop a method for diagnosing breast cancers. Thirteen deep learning technologies were used to compare the authors' method. The created method achieved 94.64%, 93.54%, and 87.77% correct diagnoses on less than

140 photos in the first, second, and third datasets, respectively. Given the short sample size, the poor accuracy rate is to be expected. Four performance indicators were used by the authors to gauge the algorithm's effectiveness: accuracy, precision, F-score, and accuracy.

Afolayan et al.[12] who proposed utilizing a PSO algorithm in conjunction with a Decision Tree tool to maximize the classification efficiency of breast cancers using WBCD.

A thorough analysis of the use of Convolutional Neural Networks (CNNs) in the identification of breast cancer from mammography pictures was presented by Abdel rahman et al.[13] It talks about different CNN architectures and how well they classify breast cancer. It does not, however, perform experiments or suggest a novel approach because it is a survey report. One drawback is that deep learning methods are always changing; therefore the research might not cover the most recent developments in CNNs for breast cancer detection.

Abunasser et al.[14]developed a CNN based method for breast cancer detection. The approach involves preprocessing mammogram images and training a CNN for feature extraction and classification. While the proposed CNN have shown promise in this context, this paper lacks extensive experimentation and performance evaluation. Additionally, it does not address potential challenges related to class imbalance in the dataset, which can affect model generalization.

Shewtha K. and others [15] additionally, machine learning algorithms can be utilized to detect and diagnose breast cancer, especially when there is a lack of data or the computer resources needed for deep learning techniques. Furthermore, hybrid approaches have been proposed that combine deep learning with machine learning techniques, utilizing the advantages of both methods to increase speed and accuracy.

V Sansya Vijayam and associates [16]. Finding regions of interest in histopathology images with the aid of clustering methods, such as Lloyd's algorithm, can increase the precision and effectiveness of ensuing classification tasks. CNNs can automatically learn pertinent features from histopathological images and attain high accuracy with the right training, therefore using them for classification is also appropriate.

Deepa B. G. et al[17].The proposed method for breast cancer classification is based on classifier augmentations. By integrating several classifiers and utilizing their unique advantages, augmentation classifiers are a strategy that can enhance the performance of machine learning models. Another helpful machine learning technique is the use of feature selection methods based on information and correlation. These methods can assist minimize the dimensionality of the dataset and find the most significant features for a given task. The comparison of the model's performance with and without feature selection strategies shows a thorough examination of how various methods affect the classifiers' accuracy. Finding the best classifier for a given dataset requires comparing various algorithms, which is what the usage of five classifiers offers. It's crucial to remember that the accuracy attained on the dataset for breast cancer might not transfer to other datasets or therapeutic contexts. Furthermore, variables like data quality, variations in imaging methods, and variations in patient demographics may have an impact on the model's effectiveness. Consequently, additional testing on separate datasets is required to determine the model's generalizability and clinical applicability.

V. M. Patel et al.[18] The proposed method used a convolutional neural network (CNN) architecture with multiple convolutional and pooling layers. The Digital Database for Screening Mammography (DDSM) dataset was utilized by the authors.

Li, W et al.[19], Three components made up the framework: a multi-scale extraction of features module, an area network proposal (RPN). There were two sections to the dataset: a training set of 220 instances and a 56-person testing set instances.

H. K. Verma et al. [20] They put forth the Convolutional Neural Network model based on networks Feature Extractor (CNFE) that takes out. They made use of the 9,109-person BreakHis dataset breast pathological pictures of harmless and cancerous growths. They demonstrated the efficacy of their suggested CNFE model in classifying mild and breast cancerous tumors.

3. CONVENTIONAL METHODS FOR EARLIEST BREAST CANCER

3.1 Identification of Cancer

The conventional methods for the early detection of breast cancer include mammography, clinical breast examination, and breast self-examination.

3.2 Breast self-examination

This straightforward method is one that ladies can do on their own at home. It entails feeling the breasts for any abnormalities, such as lumps or changes in size or texture. Once a month, usually a few days after their menstrual period, when the breasts are less likely to be swollen or sensitive, women should examine their breasts. Breast self-examination can help women get to know their breasts and identify any changes that might need more assessment, but it cannot replace mammograms or other imaging tests.

3.3 Clinical Breast Examination

A clinical breast examination is a physical evaluation of the breast performed by a healthcare professional, such as a physician or nurse. The medical professional will check the surrounding lymph nodes for swelling or soreness and feel the breasts for any lumps or anomalies. For women over 20, a clinical breast examination is usually advised as part of a regular check-up and should be carried out every one to three years.

3.4 Mammography

The most used imaging test for detecting breast cancer is mammography. Although some organizations advise beginning screening earlier for women with specific risk factors, mammograms are advised for people over 50. By identifying malignancies early when they are more curable, mammograms, which are regarded as the gold standard for breast cancer screening, have been demonstrated to lower the death rate from breast cancer. The current standard method for identifying breast cancer at the earliest stage is to combine clinical breast examination, mammography, and breast self-examination.

4. THE EARLIEST TECHNIQUES FOR IDENTIFYING, DETECTING, AND SEGMENTING BREAST CANCER

The topic of breast cancer detection and segmentation has seen a number of developments in recent years. These include some of the first detections of breast cancer, techniques for segmentation and detection.

4.1 Digital breast tomosynthesis : DBT, is a sophisticated type of mammography that takes several pictures of the breast from various perspectives to Make a three-dimensional picture.

4.2 Magnetic Resonance Imaging (MRI): This technique uses radio waves and a magnetic field to create accurate images of the breast tissue. Ladies with a breast cancer risk that is high, or when mammography and because ultrasonography is inaccurate, MRI scans are often used instead.

4.3 Automated Breast Ultrasound (ABUS): ABUS is a newer technology that uses ultrasound to generate images of the breast tissue.

4.4 Artificial Intelligence (AI) and Machine Learning (ML): AI and ML have shown promise in improving breast cancer detection and segmentation. These technologies can analyze mammography and other imaging data to identify patterns and anomalies that may be indicative of cancer.patterns and anomalies that may be indicative of cancer.

4.5 Thermography: This technique uses infrared imaging to identify temperature variations that might be a sign of malignant tissue. Although it is currently regarded as a very experimental Thermography has demonstrated potential as a tool for detecting early-stage breast cancer.

5. INVESTIGATIONS PERFORMED TO FIND BREAST CANCER IN ITS EARLY STAGES

Numerous investigations have been carried out in recent years to identify breast cancer in its early stages.

5.1 AI for Breast Cancer Screening:

In 2019, a research study published in Nature investigated the use of artificial intelligence to get better breast cancer screening. The study found that an AI algorithm was able to identify breast cancer on mammograms with greater accuracy than human radiologists.

5.2 Early Detection Blood Test:

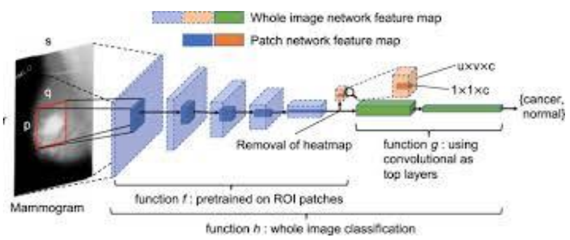
The test detects antibodies in the blood that are generated in reaction to cancer cells.

5.3 Liquid Biopsy:

The potential of liquid biopsy to identify early-stage cancer was examined in a 2020 study that was published in Nature Communications. Breast cancer As part of a liquid biopsy, a patient's check for cancer cells in blood or other body fluids or pieces of DNA. According to the study, liquid biopsy was capable of highly accurately detecting breast cancer in its early stages of precision.

6 DEEP LEARNING

Deep learning models are very useful in a range of applications, including computer programs, because they can automatically learn to extract features from data. Speech recognition, NLP, and vision. The perceptron nevertheless, is a basic neural network model that has a one layer that can only recognise patterns that are linearly separable.



Deep learning has shown great promise in breast cancer detection, particularly in the analysis of mammography images. These algorithms can learn to analyse mammography images and identify features associated with breast cancer, such as the presence of micro calcifications or masses. New mammography pictures can then be categorised as benign or malignant using these algorithms. Numerous investigations have shown that efficacy of deep learning in detecting breast cancer. For instance, in 2019, a group of Google researchers Health released a study wherein they employed a thorough algorithm for learning to examine mammography pictures from more than 76,000 women in the US and the US Kingdom. In terms of lowering false positives and false negatives, the algorithm outperformed radiologists. But it's crucial to keep in mind that deep learning algorithms are still in their infancy and need human supervision. When used in healthcare settings, deep learning Radiologists can utilise algorithms as a tool to help them with diagnosing patients more precisely, but they shouldn't be utilised as a stand-alone diagnostic instrument.

7 Research Methodology

The process of creating and honing a CNN model for breast cancer detection entails mammography pictures as either benign or malignant. The procedures for developing a CNN model are as follows: detection of breast cancer.

7.1 Data collection:

Compile a sizable dataset of mammography pictures and the labels that indicate whether or not they show a malignant tumour. The population being diagnosed should be represented in the dataset, which should be diverse.

7.2 Data pre-processing:

Construct variations of the original photographs to expand the dataset size and ensure that the images are all the same size and quality by performing pre-processing operations on the dataset, such as image resizing, normalisation, and enhancement.

7.2 Model design:

Construct a CNN model architecture with fully connected layers for classification, pooling layers for dimensionality reduction, and convolutional layers for feature extraction. The architecture may be modified and enhanced based on the model's performance on the validation set.

7.3 Model training:

Use an appropriate optimisation technique, such as Adam or stochastic gradient descent, to train the CNN model on the pre-processed dataset. For quicker training, the model should be used on a high-performance computing platform that has access to GPUs or TPUs.

7.4 Model deployment:

To give patients and healthcare professionals real-time breast cancer screening, implement the trained model in a production setting, such as a web or mobile application.7.6 Class imbalance results from the fact that there are usually far fewer malignant lesions than non-cancerous lesions in mammography pictures. This problem can be solved by employing strategies like class weighting, under sampling, and oversampling.

7.5 Class imbalance:
Results from the fact that there are usually far fewer malignant lesions than non-cancerous lesions in mammography pictures. This problem can be solved by employing strategies like class weighting, under sampling, and oversampling.

8. RESEARCH METHODOLOGY

After being reduced, the retrieved traits are divided into two categories: cancer and no cancer. The following are the specifics for every block:

A. Pre-processing: The photos used in this study were taken from a variety of datasets and show differences in resolution and size.

1. Normalization: Images in the RSNA dataset come in a variety of formats, such as 12 and 16 bits per pixel. It also has MONOCHROME1 and MONOCHROME2, two distinct photometric interpretations. The former displays grayscale pictures with pixel values rising from bright to to dark, and the latter depicts grayscale pictures with pixels that rise in values ranging from dark to bright. In order to guarantee uniformity throughout the RSNA dataset, Every MONOCHROME1 image is converted to MONOCHROME2. In order to standardize the intensity normalization and pixel values throughout the RSNA dataset takes place. This entails adjusting the pixel values to fall between 0 and 255. This corresponds to eight bits per pixel. Through the use of this normalization procedure, The dataset's pixel values become more comparable and consistent. However, there is no need for extra normalization because the pixel values in the DDSM and MIAS datasets already fall between 0 and 255. As a result, these datasets' pixel values are judged appropriate, and no more modification is needed.

2. Region of Interest Selection: We first apply a global thresholding technique to the image in order to choose the region of interest. The shape of the largest object in the picture, which represents the breast region, is then extracted. We created a mask using this shape, which allows us to crop the picture. Then separate the particular area of interest for additional examination.

B. Feature extraction: We use the features that the pretrained CNN models that were discussed in Section 2.2 computed to extract features. Since the output of the final fully connected (FC) layer has been trained for 1000 classes of the ImageNet dataset, we skip this layer and extract the features from the last layer before the final FC layer for each model.

C. Feature selection: Most features are superfluous, don't convey any valuable information, and merely make the system more complicated. When it comes to weak features, the distribution of the feature for both healthy and malignant people is comparable, indicating that the feature and the computed.

9. CONCLUSION:

In the medical field, diagnosis techniques are costly and time-consuming. To improve survival rates, machine-learning techniques can be employed as a therapeutic tool to identify breast cancer, particularly in its early stages. In the absence of highly experienced individuals, misdiagnosis is rather common; therefore this can be especially helpful for novice doctors and medical practitioners. This study's primary objective was to investigate feature-based Deep learning methods for breast cancer early detection, particularly in its early stages.

REFERENCES

[1] Ferlay, J.; Colombet, M.; Soerjomataram, I.; Parkin, D.M.; Piñeros, M.; Znaor, A.; Bray, F. Cancer statistics for the year 2020: An overview. Int. J. Cancer 2021, 149, 778–789.

[2] Lei, S.; Zheng, R.; Zhang, S.; Wang, S.; Chen, R.; Sun, K.; Zeng, H.; Zhou, J.; Wei, W. Global patterns of breast cancer incidence and mortality: A population-based cancer registry data analysis from 2000 to 2020. Cancer Commun. 2021, 41, 1183–1194.

[3] Marks, J.S.; Lee, N.C.; Lawson, H.W.; Henson, R.; Bobo, J.K.; Kaeser, M.K. Implementing recommendations for the early detection of breast and cervical cancer among low-income women. Morb. Mortal. Wkly. Rep. Recomm. Rep. 2000, 49, 35–55.

[4] Du-Crow, E. Computer-Aided Detection in Mammography; The University of Manchester: Manchester, UK, 2022.

[5] Evans, A.; Trimboli, R.M.; Athanasiou, A.; Balleyguier, C.; Baltzer, P.A.; Bick, U.; Herrero, J.C.; Clauser, P.; Colin, C.; Cornford, E.; et al. Breast ultrasound: Recommendations for information to women and referring physicians by the

European Society of Breast Imaging. Insights Imaging 2020, 9, 449–461.

[6] Schueller, G.; Schueller-Weidekamm, C.; Helbich, T.H. Accuracy of ultrasound-guided, large-core needle breast biopsy. Eur.Radiol. 2022, 18, 1761–1773.

[7] Shi, X.; Liang, C.; Wang, H. Multiview robust graph-based clustering for cancer subtype identification. IEEE/ACM Trans. Comput. Biol. Bioinform. 2022, 20, 544–556.

[8] Wang, H.; Jiang, G.; Peng, J.; Deng, R.; Fu, X. Towards Adaptive Consensus Graph: Multi-view Clustering via Graph Collaboration. IEEE Trans. Multimed. 2022, 1–13.

[9] Wang, H.; Wang, Y.; Zhang, Z.; Fu, X.; Zhuo, L.; Xu, M.; Wang, M. Kernelized multiview subspace analysis by self-weighted learning. IEEE Trans. Multimed. 2020, 23, 3828–3840.

[10] Wang, H.; Yao, M.; Jiang, G.; Mi, Z.; Fu, X. Graph-Collaborated Auto-Encoder Hashing for Multi-view Binary Clustering. arXiv 2023, arXiv:2301.02484.

[11] Bai, J.; Posner, R.; Wang, T.; Yang, C.; Nabavi, S. Applying deep learning in digital breast tomosynthesis for automatic breast cancer detection: A review. Med. Image Anal. 2021, 71, 102049.

[12] Li, Z.; Liu, F.; Yang, W.; Peng, S.; Zhou, J. A survey of convolutional neural networks: Analysis, applications, and prospects. IEEE Trans. Neural Netw. Learn. Syst. 2022, 33, 6999–7019.

[13] Zuluaga-Gomez, J.; Al Masry, Z.; Benaggoune, K.; Meraghni, S.; Zerhouni, N. A CNN-based methodology for breast cancer diagnosis using thermal images. Comput. Methods Biomech. Biomed. Eng. Imaging Vis. 2021, 9, 131–145.

[14] Erog˘ lu, Y.; Yildirim, M.; Çinar, A. Convolutional Neural Networks based classification of breast ultrasonography images by hybrid method with respect to benign, malignant, and normal using mRMR. Comput. Biol. Med. 2021, 133, 104407.

[15] Huang, Q.; Yang, F.; Liu, L.; Li, X. Automatic segmentation of breast lesions for interaction in ultrasonic computer-aided diagnosis. Inf. Sci. 2020, 314, 293–310.

[16] Huang, Q.; Huang, Y.; Luo, Y.; Yuan, F.; Li, X. Segmentation of breast ultrasound image with semantic classification of superpixels. Med. Image Anal. 2020, 61, 101657.

[17] Zhou, J.; Luo, L.; Dou, Q.; Chen, H.; Chen, C.; Li, G.; Jiang, Z.; Heng, P. Weakly supervised 3D deep learning for breast cancer classification and localization of the lesions in MR images. J. Magn. Reson. Imaging 2019, 50, 1144–1151.

[18] Yurttakal, A.H.; Erbay, H.; Ikizceli, T.; Karaçavus¸, S. Detection of breast cancer via deep convolution neural networks using MRI images. Multimed. Tools Appl. 2019, 79, 15555–15573.

[19] Rahman, A.S.; Belhaouari, S.B.; Bouzerdoum, A.; Baali, H.; Alam, T.; Eldaraa, A.M. Breast mass tumor classification using deep learning. In Proceedings of the 2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT), Doha, Qatar, 2 February 2020; pp. 271–276.

[20] Sun, L.; Wang, J.; Hu, Z.; Xu, Y.; Cui, Z. Multi-view convolutional neural networks for mammographic image classification. IEEE Access 2019, 7, 126273–126282.

[21] Carr, C.; Kitamura, F.; Partridge, G.; Kalpathy-Cramer, J.; Mongan, J.; Andriole, K.; Lavender Vazirabad, M.; Riopel, M.; Ball, R.; Dane, S.; et al. RSNA Screening Mammography Breast Cancer Detection, Kaggle 2022. Available online:

[22] Suckling, J.; Parker, J.; Dance, D.; Astley, S.; Hutt, I.; Boggis, C.; Ricketts, I.; Stamatakis, E.; Cerneaz, N.; Kok, S.; et al. Mammographic Image Analysis Society (Mias) Database v1. 21. Available online: https://www.repository.cam.ac.uk/handle/1810/250394 (accessed on 2 February 2023).

[23] Tan, M.; Le, Q. Efficientnetv2: Smaller models and faster training. In Proceedings of the International Conference on Machine Learning, PMLR, Virtual Event, 1 July 2021; pp. 10096–10106.

[24] Liu, Z.; Mao, H.; Wu, C.Y.; Feichtenhofer, C.; Darrell, T.; Xie, S. A convnet for the 2020s. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition 2022, New Orleans, LA, USA, 19–20 June 2022; pp. 11976–11986.

[25] Vijayarajeswari, R.; Parthasarathy, P.; Vivekanandan, S.; Basha, A.A. Classification

of mammogram for early detection of breast cancer using SVM classifier and Hough transform. Measurement 2019, 146, 800–805.

[26] Arafa, A.A.A.; El-Sokary, N.; Asad, A.; Hefny, H. Computer-aided detection system for breast cancer based on GMM and SVM. Arab. J. Nucl. Sci. Appl. 2019, 52, 142–150.

[27] Diaz, R.A.; Swandewi, N.N.; Novianti, K.D. Malignancy determination breast cancer based on mammogram image with k-nearest neighbor. In Proceedings of the 2019 1st International Conference on Cybernetics and Intelligent System (ICORIS), Denpasar, Indonesia, 22 August 2019; Volume 1, pp. 233–237.

[28] Agrawal, S.; Rangnekar, R.; Gala, D.; Paul, S.; Kalbande, D. Detection of breast cancer from mammograms using a hybrid approach of deep learning and linear classification. In Proceedings of the 2018 International Conference on Smart City and Emerging Technology (ICSCET), Mumbai, India, 5 January 2018; pp. 1–6.

[29] Li, B.; Ge, Y.; Zhao, Y.; Guan, E.; Yan, W. Benign and malignant mammographic image classification based on convolutional neural networks. In Proceedings of the 2018 10th International Conference on Machine Learning and Computing, New York, NY, USA, 26 February 2018; pp. 247–251.

[30] Platania, R.; Shams, S.; Yang, S.; Zhang, J.; Lee, K.; Park, S.J. Automated breast cancer diagnosis using deep learning and region of interest detection (bc-droid). In Proceedings of the 8th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics, Boston, MA, USA, 20 August 2017; pp. 536–543.