Mammogram Breast Tumor Abnormalities Detection Using DeepCNN with Discrete Cosine Transform Features

Suriya Priyadharsini M[1], Dr. J. G. R. Sathiaseelan[2]

Abstract- In the past few years, especially among younger people, breast cancer has become one of the most dangerous diseases. The goal of this study was to find out if DBT could be used in the clinic to measure the growth of breast cancer tumors. First, malignant and healthy mammograms are separated, and then, in stage 2, the size of the tumor is determined. Radiologists want to know how far along the cancer is so they can find the best way to cure and treat the person. This can be done with mammography by finding the right kind of abnormality to measure how bad the cancer is. With IBT, they may be able to get a clearer picture of the kind of abnormality. In this study, we show a brand-new way to divide malignant mammograms into six different groups. To figure out if a mammogram is malignant and how big a tumor is, features are taken from pre-processed images and run through different classifiers. The results of the best method, in this case DeepCNN, are then taken into account for further analysis. DeepCNN has divided mammograms that show cancer into six different groups. Again, the classifier (DeepCNN) is used for multi-classification using the "one against all" method. Using the maximum, median, and mean, the average of all the results from each category is found. It has been pointed out that the results are very encouraging and that the max rule is more than 96% accurate at classifying anomalies. For an experiment, a set of data from MIAS is used.

1. Introduction

The most prevalent type of cancer among women is breast cancer. Early tissue diagnosis is crucial since breast cancer tumors are often tiny at first and can develop into big glands over the period of time. Breast self-examination, doctor-performed breast exams, and mammography are all approaches for finding breast cancer; mammography is the most reliable. Finding a way to automate the technique was suggested as a result of radiologists' failures in diagnosing mammography images owing to physician weariness or optical illusion. In 10% to 30% of cases, doctors miss tissue.

Several recognised and suspected causes of breast tumor exist. These might be broken down hooked on7 primary types: stage, genetic influences, proliferative breast illness, family background of breast cancer, early - stage lung irradiation exposure, lifestyle factors, and past background of malignancy. In fact, statistics show that doctors miss 10% to 30% of breast tumors during standard testing [1]. Errors in identification orcategorization carry a heavy price. A suspicious region

cannot be proven to be cancerous or benign by mammography alone. The tissue must be taken out for evaluation using breast biopsy techniques in order to make that decision. An unneeded biopsy could resultfrom a false positive finding. Only 20% to 30% of mammary biopsies cases have been proven to bemalignant, according to data. A real tumor could go undiscovered in a false negative finding, costing more money or perhaps costing a person their time. Through the advancement of system method, doctors now have the opportunity to enhance their image processing and analysis skills by utilising computer features that canenhance mammography picture quality. Since digital breast plate mosynthesis (DBT) scans may collectimages of breasts from a variety of angles, it has helped with the detection of breast illnesses as imaging technology has advanced [2]. By creating computer- aided tools for mammography analysis, computer engineers have worked hard over the past 20 years to assist physicians in the detection and treatment of malignant tumors.

Two significant computer technology mainstreams that have been consistently investigated in the creation of computer-aided radiology technologies are image processing and intelligent systems. The anomalies of breast cancer had been the subject of extensive inquiry. However, nobody has treated it as a comprehensive issue in some of them, such as [4] discussing calcification, [3]

[1]*Affiliated to Bharathidasan University, Dept Computer Science*
Bishop Heber College (Autonomous)
Tamilnadu, India
julimca.sigc@gmail.com

[2]*Affiliated to Bharathidasan University, Dept Computer Science*
Bishop Heber College (Autonomous)
Tamilnadu, India
jgrsathiaseelan@gmail.com

discussing lumps in the breast, and [5] illustrating simply imbalance brought on by breast cancer. The unique strategy that is suggested in this study provides a comprehensive computer-aided health monitoring system while simultaneously addressing all forms of breast cancer anomalies. The identification of malignant pictures and its categorization on the basis of anomalies visible in the image are the two most challenging aspects of computerized mammography diagnostics. Various image-processing techniques have been created to address this issue. The main challenge with this endeavor is that there isn't a universal method that works well for all photographs. In order to distinguish between benign and malignant mammography pictures, we have looked into the efficacy of a screening test that feeds discrete cosine transformation (DCT) characteristics into a variety of classifiers, including DeepCNN, KNN, ANN, Bayesian, and Support Vector Machine. Then, images of cancer are fed through a bank of DeepCNN that have been trained to classify various abnormalities. For multi-classification, one against all is the strategy utilized. DeepCNN trained for one irregularity in each layer. It gives the aberration for which it was trained a likelihood. Binary classifier results are combined using median, mean, and composition criteria. The pre-processing of the mammograms is done using a very effective method that we devised in [6]. Pre-processing include automatically resizing the mammograms, deleting the breast area from the mammography, and deleting other regions that are not associated with the breast.

2. Related Work

YuGu [2021] describe existing CAD methods for lung segmentation, false positive reductions, lesion identification, edge detection, categorization, and recovery using deep learning on Computed tomography data. We go through the history of deep learning, go over some of the key components of tumor segmentation CAD systems, and evaluate the effectiveness of the chosen research using data from the, LUNA16, LIDC- IDRI, LIDC, TianChi, DSB2017, NLST and ELCAP data.

DL-based object detection and segment techniques, in example, may be capable of performing as well as doctors in closed condition by generating accurate and dependable results, according to Rodrguez [2019] Ruiz's research. However, some difficulties and problems must be fully resolved for effectively using such systems to identify breast cancer, particularly when using positive impact on online mammography pictures.

Li Shen et al. [2019] establish an algorithm for deep learning that can detect different types of breast cancers on routine mammograms by using a "end-to-end"

teaching technique that makes efficient use of trained data either with comprehensive medical analysis or only the cancer situation (label) of an entire image. After combining four models, the top individual prediction in the CBIS-DDSM database was able to achieve an AUC of 0.91, improving upon the previous best of 0.88. The average area under the curve (AUC) for the best single simulation on a collection of test full-field digital mammogram (FFDM) images from the breast data was 0.95, while for 4 combinations, the AUC climbed to 0.98.

Asma Baccouche [2022] introduced the use of a stacked ensemble using ResNet models for the last stage for breast mass classification & detection. The aim in this piece is to determine whether identified and divided breast masses are malignant or benign, and to diagnose the Breast Imaging Reporting and Data System (BI- RADS) evaluation group with a mark between two and seven, as well as whether the mass is oval, spherical, lobed, or uneven. The outcomes show that our suggested effective system could exceed the most recent deep learning techniques by taking advantage of all computerized phases.

Subasish Mohapatra [2022] developed new components of several Architectures and used transfer learning in which was before valued on others in order to compare their efficiency. This included the AlexNet, VGG16, and ResNet50. Using mammography pictures from the mini-DDSM data, which is freely accessible, the aforementioned system classifications were evaluated and trained. Therefore, a 90:10 ratios are employed as the testing approach. When adjusted using pre-trained weights, AlexNet demonstrated precision of 65%, while VGG16 and ResNet50 shown accuracy of 65% and 61%, correspondingly. When learned from beginning, VGG16 fared noticeably worse whilst AlexNet beat another. The performance of VGG16 and ResNet50 was good when transfer learning was used.

3. Proposed System

The suggested method offers a DBT system for detecting breast cancer, as well as the kind and stage of abnormalities, in mammograms. Digital mammography images are inputted into the program and then processed. The algorithm notes the type of abnormal in malignant mammography and detects them. Processes including pre-processing, feature extraction, determining cancer presence, classifying abnormalities, and a composite of classification model make up the system's five main building components. Figure 1 describes the entire system.

The mammography picture is fed to the pre-processing systems during the initial block. Discrete Cosine

Transformation is used for extracting features. The chapter on extracting features discusses the specifics of DCT. By using the top seven frequencies from the two-dimensional DCT, the approach is rendered moment. The classification accuracy of DCNN, KNN and SVM outperforms totally further reported algorithms when using DCT information. The identification of anomaly types is the most significant and original effort done in this research. This block uses a one versus all strategy to correctly classify several categories. This block demonstrates the creativity of the approach because it has been noted in the research that no one approach has been published that can recognise all different kinds of anomalies. Furthermore, methods for detecting a particular anomaly at a period have been presented.
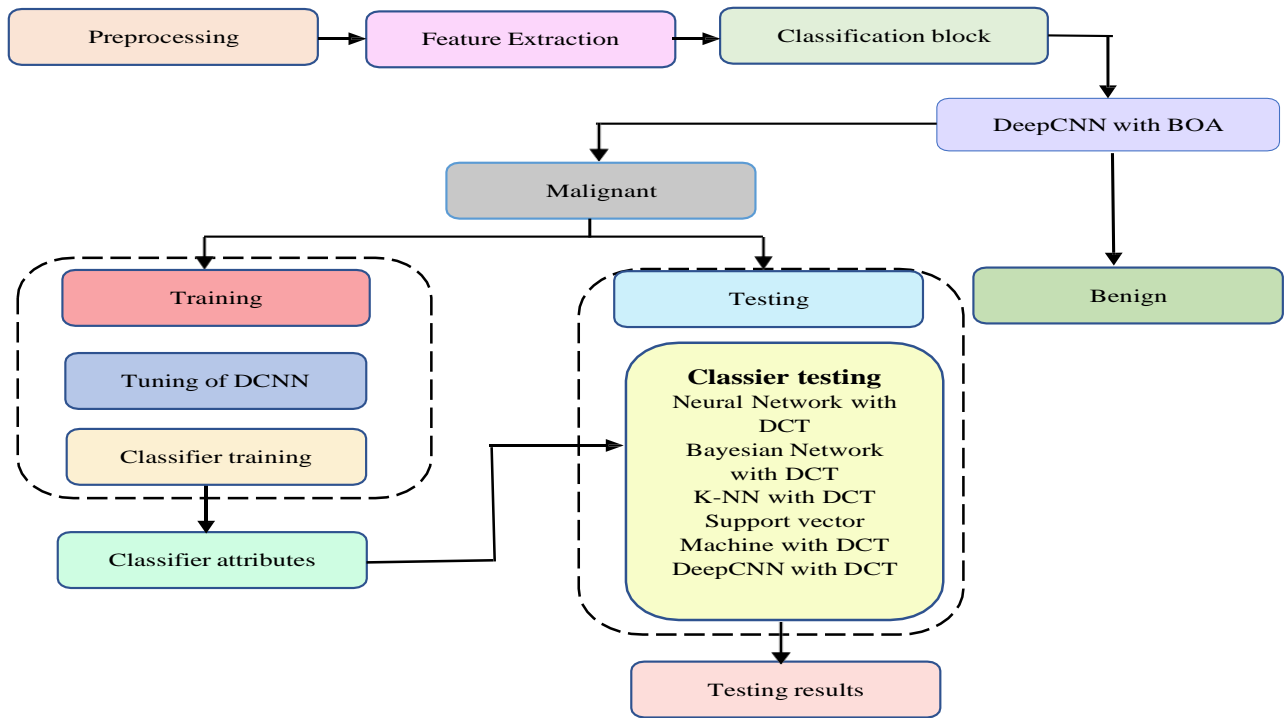

Fig.1 A Block Diagram of Proposed Architecture

How to integrate the output from all of these simultaneous binary classifiers was covered in theprevious block. There are numerous methods, including the majority vote, weighted majority vote, minimum, maximum, produce, and median instruction. The min,item, and median rules have been contrasted. The median rule yields superior outcomes.

a) Feature Extraction: "Feature extraction" refers to the process of isolating specific characteristics from apicture. The picture's relevant characteristics are identified for categorization purposes. Extraction of auseful feature set for categorization is a difficult task. Several more methods exist for extracting features, including feature extraction [11], gabor features [12], features based on wavelet transform [4], primary module analysis, lowest clatter portion renovate, choice edge feature abstraction, discriminant analysis, non-parametricslanted extraction of features, and power spectrum evaluation. There are many uses for discrete cosine transformation (DCT), including image matching. Forour suggested system, we leveraged the DCT function.

b) Discrete Cosine Transforms: (DCT) features. The signal is broken down into its spectral analysis using the separate cosine transform (DCT). DCT makes an effort to conditions and higher the picture data when processing images. The visual data can be compressed using DCT into the least number Transform coefficients while introducing any deformation. DCT has the potential to be separated and is symmetric. Calculating the variance and mean of a radial distance spectrum (equation (3)) and the compact (equation (1)) of the tumor yields the features. If the tumor is not perfectly round or oval but has more of a lobulated shape, it maybe malignant.

The compactness index evaluates how closely a breast tumor fits into a specified circle. Malignancy risk is inversely proportional to the compactness value, which approaches 1 for benign tumors:

$$C = \frac{A}{4\pi L^2} \qquad (1)$$

where A is the tumor's area and L is the breast tumor's perimeter contour.

The elliptic compact is indeed the ratio of a circumference of the fitted ellipse to a circumference of aoriginal tumor contour. It's inversely associated to tumor malignancy. The goal of the elliptic fitting technique isto locate an ellipse that closely follows the contours of a tumor. More generally, the contour lines of the tumor arefitted with elliptic equation also as model such that a specific elliptic equation can satisfy those points to the greatest extent possible, so each parameter of the elliptic equations is obtained. Figure 2 displays the result ofusing an ellipse fit. The white hue is the outline of the tumor, as well as the yellow line seems to be the fitting

ellipse. Based on the fitting ellipse produced, the features are computed as follows:

$$C = \pi \frac{(a+b)}{D} \qquad (2)$$

where an is the ellipse's semi major axis, b is is the ellipse's semi minor axis, and D is the breast tumor contour's perimeter. The degree of tumor margin roughness was measured using the radial distance spectrum approach, which involved statistically assessing the radial distance out of each point on the tumor margin to the tumor's centre. When all is said and done, we use the mean and variance of the logarithmic amplitude spectrum's harmonic components as defining features. The following is a formula for determining radial distance:

$$D(t) = \sqrt{(p_t - x_o)^2 + (q_t - y_o)^2} \qquad (3)$$

where the tumor edge points are denoted as $P_t(p_t, q_t)$ and the center point is denoted as (x_0, y_o)

Fig.2 The examples of the fitting ellipse that transformed from breast tumor contour

c) **Malignancy Detection**: Classification is used to identify malignancies. The technique of grouping input sequence into related classes is known as categorization. When choosing a positiveness, classification accuracy, technique efficiency, and high computational should be taken into account. Unsupervised and supervised classification are the two main categories of categorization. Within multi-spectral information, unsupervised categorization uncovers innate groups or patterns. Unsupervised categorization has the benefit of not requiring considerable prior knowledge of the area, therefore the algorithm is not explicitly directed. The method of classifying samples with uncertain identities is known as supervised classification. Information sequences are given along with the classifications, and supervised classification has the characteristics of requiring in-depth domain expertise. Therefore, supervised classification is more focused and under supervision, which unquestionably improves accuracy.By evaluating training data to assess how well they have

been accurately categorised, it ought to be able to spot serious mistakes.

d) **Digital breast tomosynthesis (DBT):** Digital breast tomosynthesis (DBT) is replacing conventional mammography as the gold standard for imaging breast cancer due to improved screening and diagnostic imaging. Acquiring tomosynthesis helps with identifying and characterising lesions as well as pinpointing their exact location, as it eliminates the confounding effect of overlapping tissue. Furthermore, a reconstituted DBT information set's quasi three-dimensional data enables a more efficient imaging work-up in comparison to relying simply on two-dimensional purely digital mammography during imaging.

e) **Support Vector Machine:** The Support Vector Machine (SVM) is an extremely effective technology in use for classification. SVM has also been implemented to tackle a range of real-world issues, including gene expression on microarray data processing, face recognition, cancer detection, text categorization, and

glaucoma identification. The suggested technique employed SVM to classify mammography pictures as benign or cancerous on a binary basis. For multi-classification once again, binary SVM is employed to identify anomalies in a malignant mammography. A one-versus-all strategy is used to convert a binary SVM into a multiclassifier. In essence, SVM seeks to segment the input data into selection surfaces. A hyperplane called a selection plane splits the data into two groups. The supporting vector that defines the hyper plane is made up of training points. Maximizing the margins among 2 classes of the hyper plane is the central goal of SVM.

f) K-nearest Neighbors: Supervised instance-based learning classifier K-nearest neighbors (KNN). The preponderance of K-nearest neighbors' categories are used to classify the results of the new example request. A new instance is categorized using its characteristics and training examples. Any instance's closest neighbours are calculated using the ""city block distance, "Euclidean distance," "cosine distance," "correlation distance,". According to the "nearest," "random," or "consensus" rules, classifications are made.A tie is decided at chance. KNN's suitability for cross categories is one of its benefits because its categorization judgement is based on a relatively small neighborhood of comparable items. This means that even if the destination class comprises of entities whose predictor variables have various properties for distinct subsets, it can still produce results with high accuracy. The fact that the KNN similarity measure considers all attributes identically when calculating commonalities is a significant downside. When only a limited fraction of the characteristics is relevant for categorization, this might result in low cosine similarity and categorization mistakes.

g) Neural Network (NN): A biological neural network consists of a collection of neurons, each of which might have either a literal or metaphorical significance. There may be many connections between neurons in a biological network, and each neuron may have manyconnections with other neurons. Although symptomatology terminals and other links are feasible, axonal and oligodendrocytes often link at terminals. Transport of neurotransmitters results in additional types of communication besides physical transmission.

The way that biological neural processes data set serves as inspiration for pattern recognition models such as artificial intelligence, cognitive modelling, and neural networks. Cognitive models and AI strive to simulate thefunctioning of the brain's neural circuits. In the area of artificial intelligence, artificial neural networks were used to successfully develop intelligent agents and robotic systems capable of voice recognition, picture identification, and adaptive control.

Historical digital computers, which evolved from von Neumann paradigm, access memory via a number of processors in order to carry out predefined guidelines. However, neural networks were developed as an offshoot of research into how biological systems process information. In contrast to the nonlinear paradigm, neural network technology does not separate memory and processing. The concept of a neural network, which has helped researchers better understand how brain neurons behave, is the basis for efforts to construct artificial intelligence.

h) Deep CNN Using DCT: In this instance, the DCT technique is used to provide the straight photo collection to the classifier's training mechanism. The BOA variable is selected at random during the initialization of the direct values. The planner offers the most iterations atthe time of variable creation. As an outcome, in direction to progress the graphics performance, the suggested filteris functional during the start-up phase. These processed pictures are forwarded to the convolution layer, which applies a convolution filter to the picture to activate the breast picture's characteristics. The picture will be normalised by the convolution 2D layer before the ReLU layer's training method is carried out. Through a sample design, the pooling layer chooses the picture's greater value. Deep CNN uses a single GPU unit to do categorization. The Deep CNN receives a direct training image from the DCT approach. The BOA provides the assumption accuracy based on the training picture.

Utilized is a convolution neural network with three layers and deconstructing. The engagement system made the feature map available for each picture's geographical size. The highest values of the educated picture is then sent to the pooling layer for selection. The aforementioned layers in the convolution neural network carried out various tasks to trigger the operation. In the suggested procedure, the 90 input photos are recovered for the CNN input layer's classification of breast cancer. In this convolution layer, it has been normalised. The MIAS dataset, which has a resolution of 1024×1024 pixels, is obtained for breast cancer pictures. Because thepicture is shrunk to extract features, a single GPU processor can evaluate the 64×64 picture data. The breast cancer picture in greyscale is treated using the convolution 2D layer to eliminate the characteristics.

Therefore, the CNN is fed with the idealized views from the ReLu layer; else, it produces zero. The pooling layer plays a key role in turning on CNN. The maximum number of images from the ReLu layer are chosen in this case using the bigger pooling layer. The densely integrated level provides the picture identification result from feature extraction. The pooling level is where the extract is initiated. The SoftMax layer is used to produce the classification operations where the picture

probabilities are implemented. To choose the training image with maximum accuracy in this, the Butterfly optimization technique is used. The three procedures,such as feature selection, feature extraction and classification, are all carried out by a single Deep CNN algorithm. To evaluate the effectiveness, the sensitivity, accuracy, specificity, and calculation period are evaluated. The datasets were chosen for them

trustworthiness and anonymity (de-identification) were performed to ensure complete privacy. There are 10 attributes, and a single class stands in for the dependent variable, for a total of 11 fields that represent the estimation produced by the machine learning method. As seen in Table 1, all domain attributes have numeric value

Table 1: Attributes of Breast Cancer

Number	Attribute	Domain
1	Sample code number	1 to 10
2	Clump Thickness	1 to 10
3	Uniformity of Cell Size	1 to 10
4	Uniformity of Cell Shape	1 to 10
5	Marginal Adhesion	1 to 10
6	Single Epithelial Cell Size	1 to 10
7	Bare Nuclei	1 to 10
8	Bland Chromatin	1 to 10
9	Normal Nucleoli	1 to 10
10	Mitoses	1 to 10

i)Abnormality detection. Mammogram classification into malignant and benign images has previously been covered in earlier research. The diagnosis of breast cancer's intensity is a significant problem. The type of irregularity found in the data set reveals information about its intensity. The next sections go into detail about these anomalies. Therefore, as shown in Table 1, the MIAS dataset divides breast cancer abnormalities into seven categories. Breast cancer that is aggressive as well

as in situ also can have micro-calcification clusters. Presently, the presence of micro-calcifications allows for the timely identification of several breast malignancies. They are still only medically dangerous when they emerge as groups of three or more benign tumors. Mammograms can detect microcalcifications that rangein diameter from around 0.1 mm to 0.5 mm. A sample of micro calcification clusters is shown in Figure 2(a).

Table 2 Abnormal Mammogram Results

Period Number	Abnormality term	Abbreviation
One	Calcification	CALC
Two	Well-defined/circumscribed masses	CIRC
Three	Speculated Masses	SPIC
Four	Other, ill-defined masses	MISC
Five	Architectural distortion	ARCH
Six	Asymmetry	ASYM
Seven	Normal	NORM

Both malignant and benign groups may exist. It might be challenging to distinguish between benign and cancerous clusters solely on mammographic characteristics. To

avoid having too many false positives while remembering all groups, classification for micro-

calcifications is essential given that 80% of all groups are created by harmless activities.

Calcium formations that resemble small granules are known as calcifications. These plaque build-ups seem rather bright (dense) when compared to the healthy tissues around them. Mammogram calcifications are a key sign of malignant breast cancer. Malignant benign growths typically have a spreading direction, are numerous, clustered, tiny, variable in shape and size and pointed in shape. Malignancy-related calcifications are often smaller than benign calcifications. They are significantly easier to detect on a mammogram because they are typically coarser, frequently circular with straight edges, fewer in quantity, more evenly

distributed, and more uniform in shape and size. The harshness of their structure is one of the primary distinctions among malignant and benign abnormal cells. Skin benign growths, arterial calcifications, big shaft calcifications, circular calcifications, viridian calcifications, eggshell or edge calcifications, milk of calcium abnormal cells, seam calcifications, and dysplastic calcifications are examples of common benign calcifications. Amorphous and coarse heterogeneous calcifications are indications of malignancy. Fine pleomorphic, fine-linear, and fine linear-branching calcifications are malignancy highly suspect calcifications. It's crucial to take their dispersion into account while evaluating calcifications (diffuse, regional, cluster, linear, segmental).
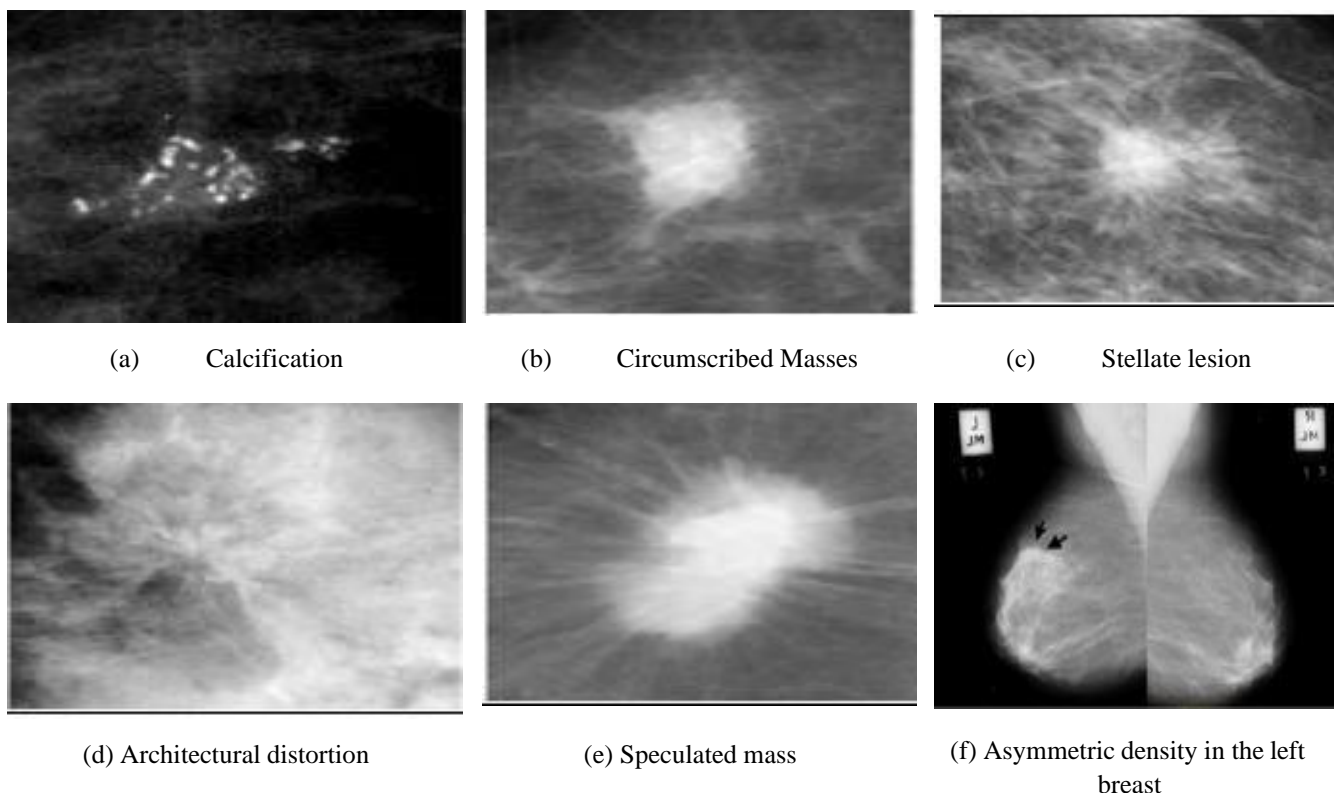

(a)   Calcification   (b)   Circumscribed Masses   (c)   Stellate lesion

(d) Architectural distortion   (e) Speculated mass   (f) Asymmetric density in the left breast

Fig.3 Breast abnormalities

Calcium deposits are evenly distributed throughout the breast. Greater dense breast size (>2 cm3) is dispersed with calcifications in a gets affected, and these alterations are frequently benign. If there are five or more calcifications in a tiny breast tissue area (less than1 cm3), that is a sign of a cluster of calcifications. The presence of calcification in a linear pattern is a sign of cancer. One can determine which calcifications are benign, which ones need to be closely monitored, and which ones require biopsy by examining the location, thickness, size, or morphology, variation, number, and existence of accompanying abnormalities, such as duct dilation or a mass. The shape of particular calcifications determines their classification and possible causation more so than its magnitude. Micro-calcifications can

vary in size, structure, and intensity, which is a worrying characteristic, however variation should be evaluated in combination with morphological.

Dynamically formed and shaped calcifications with smoother margins are far less probable to be malignant compared to those with pointed, sharp edges and varied morphology. The visual indicators that doctor look for during screening mammography can be divided into three fundamental groups, excluding micro-calcification clusters: masses, architecture abnormalities, and asymmetric density. These anomalies could be symptoms of metastatic breast cancer. Clearly delineated masses (also known as confined masses) are typically benign. A lump, though, is more probable to be cancerous if it has a slight jagged edge. If a tumor is surrounded by a

radiating pattern of scutes, it is called a lobulated mass and stipe lesion. Parenchymatous lesions are extremely doubtful signs of breast cancer. A spatially lesion that is visible in at least two separate perspectives is referred to as a mass. Before its 3 is established, a prospective mass must be referred to as "imbalance" or "non - symmetric density" if it can only be observed in a singular projected. Varied masses have various forms, varying edges (delimited, micro lobular, veiled, unclear, activities that are performed), and various densities (morbidly obese, densities, isodense, and high density) (curved, elliptical, lobular, uneven). The presence of well-defined borders, isodense and low density, and a fatty halo are all characteristics of benign lesions, however none of these features should be taken as absolute guarantees of benignity. When a narrow radiopaque band completely encircles a contained tumor, doctors call it the "crowned sign," and it's sure-fire sign that the mass in question is harmless. Constricted (very well delineated) edges define

the boundary between both the lesion and the surrounding cells. None suggests invasion lacking extra factors. In Figure 2a mass with a bounded border is seen (b). The outlines of lesions containing micro-lobular borders are undulating. Due to overlaying with the interstitium, the edges of the mass that are opaque (erased) are also erased. This phrase is employed when a doctor is certain that a mass has rough points but unnoticed margins. Penetration by the lesion may happen due to the imprecise edges, which are not most likely the result of the normal breast tissue placed on top. As seenin Figure 2, lines spreading from a mass's borders identify lesions with lobulated margins (e). A lesion witha history of non- or malignant behavior is cause for concern. The physical characteristics of a mass may help determine whether it is malignant or benign. Figure 3 shows how tumors with irregular forms are a telltale signof cancer. Round or oval masses, the most common shape for benign tumors, are a good sign.
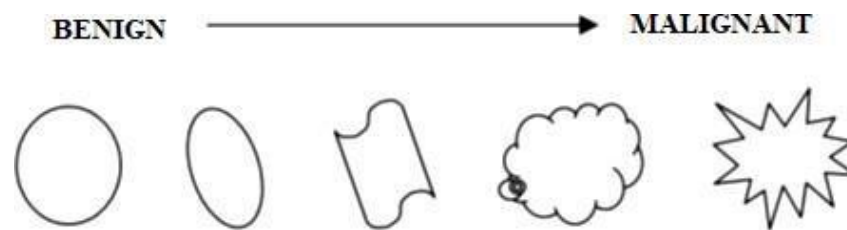

Fig.4 Morphologic spectrum of masses

Architectural distortion is the disruption of the radial ductal pattern. These lesions might have benign or malignant mechanisms and are sometimes relatively inconspicuous. Figure 2 displays a mammography with geometric deformation (d). Architectural distortion, which includes spiculations spreading from a spot and focused contraction or deformation at the margin of the tissue, is described as deformations of the typical structure with no identifiable mass apparent. Architectural deformation of breast tissue may be a sign of malignancy, particularly when combined with other obvious lesions such masses, asymmetries, or calcifications. When trauma-related scarring and soft- tissue injury is taken into account, architectural distortion can be categorized as benign. On mammograms, aggressive cancers may show up as a lump or an architecture deformation. A tumor is referred to as a stellate lesion if silica particles are arranged in a radiate arrangement around it. Not every tumor has a center mass, and spheric cancers in particular arefrequently only visible due to a factor associated of the thyroid gland. Because of the imbalance in the left and right breast patterns, some tumors can be found by doctors.
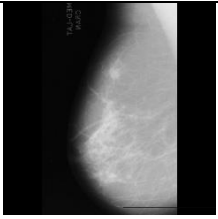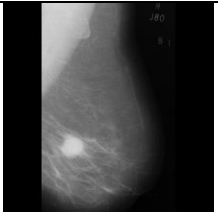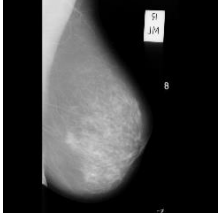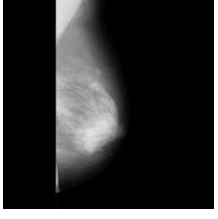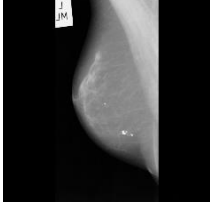
Furthermore, in actual practice, a wide range of manifestations scratches deprived of a central mass to lesions that have both a mass and spiculesis observed. Maybe an imbalance between the left and right breasts is the sole reason a growth can be diagnosed as a malignancy. A concentration evident in only one picture is suspicious; mammograms of both breasts must be remarkably similar. Mammograms from earlierscreenings sessions are utilized to boost the sensitivity and specificity in order to find differences between the old and new pictures. Several benign processes generate difficult-to-distinguish benign from malignant tumors. Masses and micro-calcifications frequently coexist in a single mammography, simplifying classification and identification. Among the best indicators for spotting initial breast cancer is the imbalance of the breast between the 2 sides. Asymmetric density is presented in Figure2(f).

Global asymmetries is often a benign shift, but the observation may be important if it coincides with a dense breast cancer. On two aspects, a focused imbalance is a limited area of imbalance, but it lacks the mass's bounds and prominence. The fatty layer that surrounds it is typically overlaid on an area of healthy fibro mammary

glands. Despite potentially being a symptom of malignancy, focal asymmetries are frequently benign. Asymmetries that are altering, expanding, or fresh, those that are perceptible, and those that connected to other discoveries such micro-calcifications or structural

deformation, are of particular significance. The density is viewed with a higher degree of suspicion for malignancy if a perceptible thickness or mass correlates to an asymmetric density.

Table 3 An overview of the various different tissues in the MIAS database is shown.

Reference number	Image	Intestinal tissue's makeup	current type of abnormality	picture-coordinates of the abnormality's center	Estimated total distance of a circle encompassing the anomaly in pixels
mdb023		Fatty-glandular	Circumscribed masses	538, 681	29
mdb028		Fatty	Circumscribed masses	338 ,314	56
mdb058		Dense-glandular	Other, ill-defined masses	318, 359	27
mdb072		Fatty-glandular	Asymmetry	266, 517	28
mdb075		Fatty	Asymmetry	468 ,717	23

4. Implementation Details

Using the pre-processing methods mentioned above, we have retrieved DCT elements out of each picture. Only the top 7 spectral characteristics are being used as the characteristics vectors for categorization. There are 3444 complete features in the set of features since the picture

is divided into 4 components and DCT is performed to each portion of the image. All of the features in thedataset are arbitrarily split in half. There will be a 25% test and 75% training component. Each classifier, including the DCNN, Bayesian, ANN, SVM and KNN, receives 75% of the features as input. 25% of the feature set is fed to classifiers for testing after training is

finished. In this manner, classifier accuracy is confirmed. Now, this very same ratio of 1513 features from 54 malignant mammograms is separated between training and testing to train Deep CNN for detection based. The multi-classification outcomes are also shown here.

5. Experimental Results

The suggested scheme computes the picture's 2-Ddiscrete cosine transform. We only use the top seven greatest frequency content as features vectors for classification in order to reduce computational complexity and increase throughput of the system. We have employed a variety of classifications to categories mammograms, and our results have shown that DeepCNN, Support Vector Machines (SVM), and KNN are the best. The outcomes listed in Table 4 and table 5as well as Figures 4 and 5 of both the plot represent the best of 100 runs. Results vary between 5 and 10% depending on the predictor. Table 6 contains the findingsfrom the chapter on abnormality detection. The Mammographic Institute Society Research served as the project's data (MIAS). The mammography has a resolution of 200 microns and a size of 1024 x 1024 pixels.This collection covers 322 mammograms of the right and left breast from 161 women, of which 54 had cancer, 69 had benign cancer, and 207 had ordinary cancer. This data containing a folder that tilts the mammograms in the MIAS record and offers pertinent information, such as the category of abnormality, x, y, and estimated size (in pixels) of a round surrounding the irregularity. The types of detected anomalies are used to categorize the abnormalities. In Table 3, information is provided by column.

Comparative analysis of clinical data and pathological characteristic data in patient's breast cancer tumor size.

The findings showed that tumor size was greater in the triple-negative and HER2-overexpressed subtypes than in the luminal A, B, C, D, and luminal E in these subtypes (Fig. 4). The tumor diameters were 11.86 ± 0.59, 2.31 ± 0.92,3.74 ± 1.94,3.12 ± 1.74 and 3.26 ± 1.81, in the luminal A, B, C, D, and luminal E, triple-negative, and HER2 over expression, respectively. However, statistical analysis showed no appreciable variations in tumor size between these five groups in fig 4. The HER2 over expression tumor had a considerably greater calcification score than the other tumor.
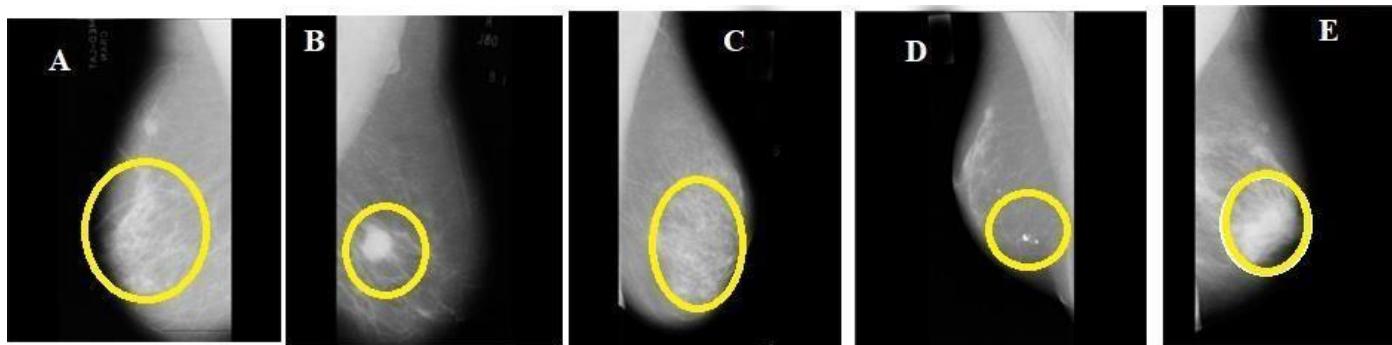

Fig. 5 Tumors in different types of breast cancer. The yellow circles indicate the location of the lesion. (A) Luminal A subtype. Tumor size,1.7x1.5 cm. (B) Luminal B subtype. Tumor size, 2.3 cm. (C) Human epidermal growth factor receptor2-over expression subtype. Tumor size, 7x6 cm. (D). Tumor size, 1.4x1.0 cm (D). (E) Tumor size, 1.7x1.5 cm

Tumor size analysis: In order to further investigate the clinical usefulness of indices produced using digital breast tomosynthesis (DBT) utilizing DeepCNN for the identification of breast cancer, thresholds of indicators were created for things like tumor size, calcification score, or lymph node size. According to the sixth edition of a breast cancer classification scheme published by American Joint Committee on Cancer (22), T1 stage tumors have such a maximum diameter of 2 cm. Since tumor size is linked to prognosis, this value of 2 cm was employed in the current study. Since previous research by Lu et al. indicated that calcification scores between 9 and 12 hint to malignant calcification, this study's threshold value was set at 10. Lymph nodes greater than 1.5 cm have been found to have a significantly different incidence rate than those between 0.6 and 1.4 cm in size.

To that end, 15 millimeters was selected as the lymph node size threshold for the current study. There were then found to be statistically significant differences in lymph node size and calcification score across the five groups.

6. Performance Measures

We calculated and analysed the DeepCNN classifier's accuracy, sensitivity, and specificity for detect tumor size in order to evaluate its efficiency. The following defines them:

Accuracy: Total mass divided by categorized mass

$$\frac{TP + TN}{TP + TN + FP + FN}$$

Sensitivity: Total number of malignant masses divided by the number of correctly classified malignant masses

$$\frac{TP}{TP + FN}$$

Specificity: Total number of benign masses separated by the number of correctly categorized benign masses

$$\frac{TN}{TN + FP}$$

As shown in Table 4, the accuracy, sensitivity, and specificity for DCNN and SVM are quite high at 98.3%.

Table 4 displays the classification of malignant mammography outcomes.

Method	Accuracy (%)
Neural Network with DCT	63.03
Bayesian Network with DCT	63.01
K-NN with DCT	89.5
Support vector Machine with DCT	98.1
DeepCNN with DCT	98.9

Table 5 shows the outcomes indicators for classifying malignant mammograms.

Method	Accuracy	Sensitivity	Specificity	Precision	F1 Score
Neural Network with DCT	63.03	70.5	43.4	76.3	73.3
Bayesian Network with DCT	63.01	70.5	42.8	77.7	74.01
K-NN with DCT	89.5	90.5	81.8	97.3	93.8
Support vector Machine with DCT	98.1	98.6	66.6	99.3	98.9
DeepCNN with DCT	98.9	99.3	66.6	99.3	99.3

Table 6 DeepCNN results for one against all approach

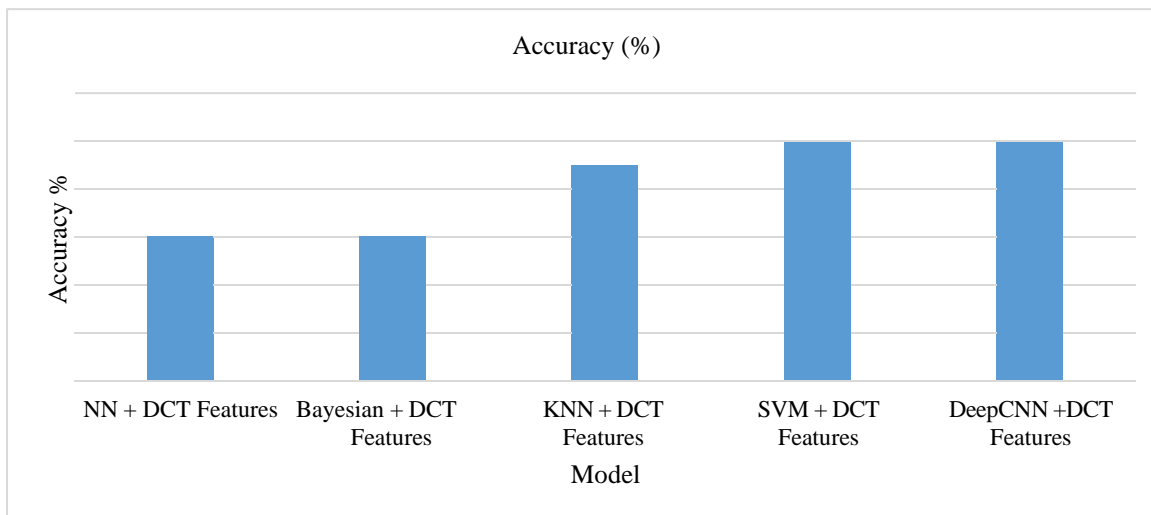| Number	Category of Abnormality	Contraction	Mean Rule %	Median Rule%	Product Rule%
One	Calcification	CALC	97.6	95.6	96.6
Two	Well-defined/circumscribed masses	CIRC	96.1	93.1	97.1
Three	Speculated Masses	SPIC	98.3	94.3	96.3
Four	Other, ill-defined masses	MISC	98.2	96.2	95.2
Five	Architectural distortion	ARCH	97.3	96.3	95.3
Six	Asymmetry	ASYM	98.1	97.1	94.1

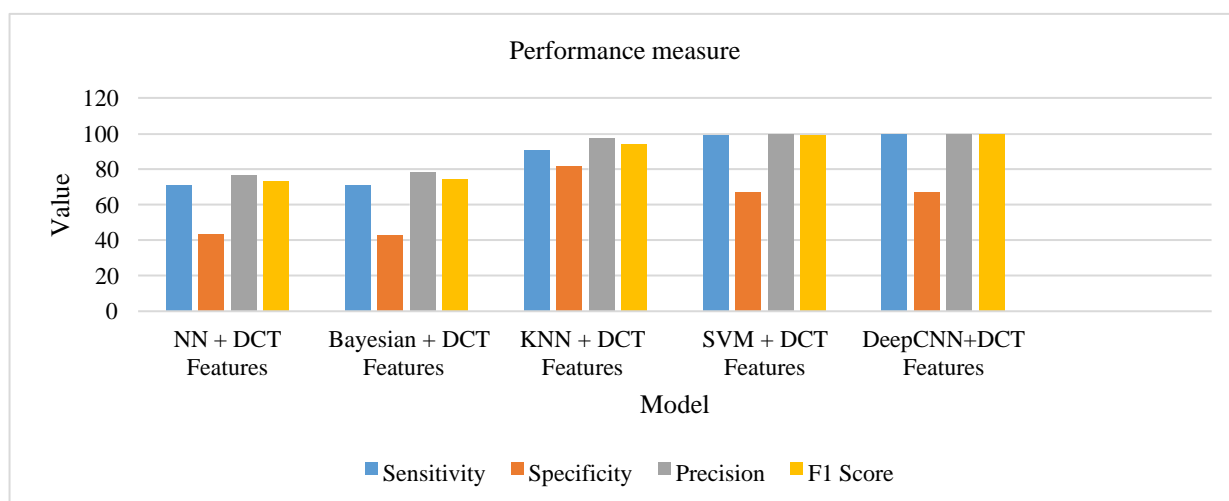**Fig. 6** Graph representation for accuracy results

Fig. 7 Graph representation of malignant mammogram results

7. Conclusion and Future Work

In order to classify breast cancer and estimate tumor size from mammograms, the proposed approach was created. Multiple phases of diagnostics are carried out by this technique. Pre-processing of the mammography picture is performed in the initial stage to reduce computational expense and improve the likelihood of detection. During the second stage, DCT features are retrieved. Using the retrieved properties, mammography can be sorted according to the size of the tumors they reveal. Malignancy in mammograms can be classified using these derived features. In order to identify mammography abnormalities, malignant pictures are reclassified using a one-against-all method that incorporates DCNN with DCT. Data are chosen for their trustworthiness and anonymity is ensured by the process of selecting and masking the data. The DCNN using DCT algorithm's estimation yields 11 fields, where 10are attributes and 1 is the class encoding the dependent variable. Already, classification's efficiency has been proven to be better than that of similar systems. This is because of the consistency throughout the dataset. The

created system favours SVM or DCNN classifiers over others when distinguishing between cancerous and normal mammograms. Excellent outcomes were achieved using one against all methods for multiple classification. All trials demonstrate that, when matched to recently suggested technique, the suggested technology offers remarkably successful outcomes. In order to recognize malignant mammography, we have an accuracy of classification of 98.99% on average.

If there are many abnormalities present in the mammography image, the suggested algorithm will only be able to identify a single dominant abnormality using the mean rule, despite the fact that this rule is more likelyto correctly identify the presence of any abnormalities. Future research will aim to solve this problem. The next steps in this research are to optimise the characteristics to use only the elements that increase accuracy, and to assess the tumor's thickness and area.

References

[1] Harada-Shoji, Narumi, Akihiko Suzuki, Takanori Ishida, Ying-Fang Zheng, Yoko Narikawa-Shiono,

Akiko Sato-Tadano, RieOhta, and Noriaki Ohuchi. "Evaluation of adjunctive ultrasonography for breast cancer detection among women aged 40-49 yearswith varying breast density undergoing screening mammography: a secondary analysis of a randomized clinical trial." JAMA network open 4, no. 8 (2021): e2121505-e2121505.

[2] Conant, Emily F., William E. Barlow, Sally D. Herschorn, Donald L. Weaver, Elisabeth F. Beaber, Anna NA Tosteson, Jennifer S. Haas et al. "Association of digital breast tomosynthesis vs digital mammography with cancer detection and recall rates by age and breast density." JAMA oncology 5, no. 5 (2019): 635-642.

[3] Yu, Xiang, Qinghua Zhou, Shuihua Wang, and Yu-Dong Zhang. "A systematic survey of deep learning in breast cancer." International Journal of Intelligent Systems 37, no. 1 (2022): 152-216.

[4] Oyelade, Olaide N., and Absalom E. Ezugwu. "A novel wavelet decomposition and transformation convolutional neural network with data augmentation for breast cancer detection usingdigital mammogram." Scientific Reports 12, no. 1 (2022): 1-22.

[5] Hassan, Nada M., Safwat Hamad, and Khaled Mahar. "Mammogram breast cancer CAD systems for mass detection and classification: a review." Multimedia Tools and Applications (2022): 1-33.

[6] Baccouche, Asma, Begonya Garcia-Zapirain, and Adel S. Elmaghraby. "An integrated framework for breast mass classification and diagnosis using stacked ensemble of residual neural networks." Scientific reports 12, no. 1 (2022): 1-17.

[7] Gu, Yu, Jingqian Chi, Jiaqi Liu, Lidong Yang, Baohua Zhang, Dahua Yu, Ying Zhao, and Xiaoqi Lu. "A survey of computer-aided diagnosis of lung nodules from CT scans using deep learning." Computers in biology and medicine 137 (2021): 104806.

[8] Rodriguez-Ruiz, Alejandro, Kristina Lång, Albert Gubern-Merida, Mireille Broeders, Gisella Gennaro, Paola Clauser, Thomas H. Helbich et al. "Stand-alone artificial intelligence for breast cancer detection in mammography: comparison with 101 radiologists." JNCI: Journal of the National Cancer Institute 111, no. 9 (2019): 916-922.

[9] Shen, Li, Laurie R. Margolies, Joseph H. Rothstein, Eugene Fluder, Russell McBride, and Weiva Sieh. "Deep learning to improve breast cancer detectionon screening mammography." Scientific reports 9, no. 1 (2019): 1-12.

[10] Mohapatra, Subasish, SarmisthaMuduly, Subhadarshini Mohanty, J. V. R. Ravindra, and Sachi Nandan Mohanty. "Evaluation of deep learning models for detecting breast cancer using histopathological mammograms Images."Sustainable Operations and Computers 3 (2022): 296-302.

[11] Liu, Xiu, and Chris Aldrich. "Deep Learning Approaches to Image Texture Analysis in Material Processing." Metals 12, no. 2 (2022): 355.

[12] Chowdhary, Chiranji Lal, Tapan Kumar Das, Vijaykumar Gurani, and Abhishek Ranjan. "An improved tumor identification with gabor wavelet segmentation." Research Journal of Pharmacy and Technology 11, no. 8 (2018): 3451-3456.

[13] Zhang, Yu-Dong, Suresh Chandra Satapathy, David S. Guttery, Juan Manuel Górriz, and Shui-Hua Wang. "Improved breast cancer classificationthrough combining graph convolutional network and convolutional neural network." Information Processing & Management 58, no. 2 (2021): 102439.

[14] Alshammari, Maha M., Afnan Almuhanna, and Jamal Alhiyafi. "Mammography Image-Based Diagnosis of Breast Cancer Using Machine Learning: A Pilot Study." Sensors 22, no. 1 (2021): 203.

[15] Basurto-Hurtado, Jesus A., Irving A. Cruz-Albarran, Manuel Toledano-Ayala, Mario Alberto Ibarra-Manzano, Luis A. Morales-Hernandez, and Carlos A. Perez-Ramirez. "Diagnostic Strategies for Breast Cancer Detection: From Image Generation to Classification Strategies Using Artificial Intelligence Algorithms." Cancers 14, no. 14 (2022): 3442.

[16] Surendhar, S. Prasath Alias, and R. Vasuki. "Breast cancer detection using deep belief network by applying feature extraction on various classifiers." Turkish Journal of Computer and Mathematics Education 12, no. 1S (2021): 471-487.

[17] Shen, Li, Laurie R. Margolies, Joseph H. Rothstein, Eugene Fluder, Russell McBride, and Weiva Sieh. "Deep learning to improve breast cancer detectionon screening mammography." Scientific reports 9, no. 1 (2019): 1-12.

[18] Patel, Jalpa J., and S. K. Hadia. "An enhancement of mammogram images for breast cancer classification using artificial neural networks." IAES International Journal of Artificial Intelligence 10, no. 2 (2021): 332.

[19] Zahoor, Saliha, Umar Shoaib, and Ikram Ullah Lali. "Breast Cancer Mammograms Classification Using Deep Neural Network and Entropy-Controlled Whale Optimization Algorithm." Diagnostics 12, no. 2 (2022): 557.

[20] https://www.kaggle.com/datasets/kmader/mias-mammography