International Journal of Zoology and Applied Biosciences ISSN: 2455-9571
Volume 10, Issue 6, pp: 211-228, 2025
https://doi.org/10.55126/ijzab.2025.v10.i06.024 http://www.ijzab.com

Research Article

BREAST TUMOUR DETECTION AND CLASSIFICATION: A SURVEY

[1]Sangeeta Sharma, [2]Megha Shrivastava, [3]Ved Kumar Gupta, [4]Neha, [5*]Neeraj Shrivastava, [6]Yagyapal Yadav

[1,4] Department of CSE, National Institute of Technology Hamirpur, Himahcal Pradesh, India
[2]Department of Mathematics, Medi-caps University, Indore, Madhya Pradesh, India
[3,5*,6]Department of CSE, IPS Academy, Institute of Engg. & Science, Indore, Madhya Pradesh, India

Article History: Received 11th September 2025; Accepted 26th October 2025; Published 1st November 2025

ABSTRACT

Cancer considered as the primary wellspring of death nowadays. Cancer is by and large the unusual, uncontrolled development of cells that structure the tumour. One such cancer type is breast cancer. Early discovery, conclusion, and therapy of cancer have diminished the danger of death. Automated system for breast cancer detection will support for accurate diagnosis. The most common screening methods utilized are mammogram and Magnetic Resonance Imaging (MRI). The influence of Machine Learning (ML) in our life and our society plays an important role. To classify normal and abnormal cells for breast cancer, methods are partitioned into picture procurement, pre-handling, division, include extraction and the determination. It is possible to improve the biopsy process by which is beneficial for mammography and physical examination. This paper is centered on the survey of the cutting-edge work strategies and methods that are utilized for the identification and arrangement of breast tumour.

Keywords: Breast cancer, Deep Learning, Mammogram Images, Ultrasound Images, Thermography Images.

INTRODUCTION

World Health Organization (WHO) has revealed that the breast cancer development is the most broadly perceived sort of cancer growth among the world globally. It is the most elevated positioned cancer growth. American cancer society has reported that around 90 percent of the women are known to suffer from the breast cancer. The lifestyle changes of the people in the developed countries increase the occurrence of breast cancer among the women for the most part in the age of 35-55. Breast tumour describes the abnormal growth of cells onto the breast. The breast masses in the breast region appear as lump with different shapes and sizes. The tumours are of two primary sorts: Benign tumours and malignant tumours. The benign tumours are non-cancerous and does not create and spread to different parts of the body dissimilar to in malignant tumour. These tumours can be dealt easily with the surgery or radiation therapy, if any treatment is required for them. The benign cell masses are oval, round. Malignant tumours are the cancerous cell that grow fast and spread out to the other part of the body, if treatment is not provided in the right way and right time and can lead to severe death of the patient. These cell masses are rough, blurred, and irregular in shapes. Symptoms of such tumour are pain, swelling, tenderness, redness in the breast, skin irritation, etc. Breast cancer represents 33% of all cancers studied among women and is responsible for 18% of cancer-related deaths. Although it has become the leading cause of cancer death among women, it was ranked second to lung cancer until 1985. The exact cause of breast cancer remains unclear, making it impossible to prevent. However, modern treatment technologies are highly effective in combating the disease. These advancements allow for the detection of breast cancer at its early stages. The most effective approach to managing breast cancer today is to surgically remove the tumour in its early phase. Mammography remains the most reliable method for detecting breast cancer. However, a negative mammogram is observed in approximately 10 to 30% of women with breast cancer who undergo screening. Around 66% of these bogus negative

*Corresponding Author: Neeraj Shrivastava, Department of CSE, IPS Academy, Institute of Engg. & Science, Indore, Madhya Pradesh, India Email: neeraj0209@gmail.com

mammograms could not be distinguished from cancer on qualms. For radiologist it is a complex task to retain interest in understanding images showing only a few abnormalities out of thousands. Thus, computer-supported frameworks are needed to diagnose breast cancer in mammograms.

As indicated by the American cancer society, the Surveillance, Epidemiology, and End Results (SEER) [39] information tracks 5-year relative perseverance rate for breast cancer and it bundle tumours into restricted, territorial, and removed stages: Restricted methods there is no sign that the cancer cell has fanned out of the breast. Local suggests that the cancer has fanned out of the breast to the nearby constructions and the lymph centres. Far Off suggests that the cancer has fanned out to inaccessible portions of the body such as lungs, liver, and bone. Table 1 describes the approximate 5-year relative survival rate of the breast cancer.

Early discovery of breast cancer can diminish the danger of death by giving legitimate therapy with the goal that cancer cells may not develop and impact in an abnormal way. Cancer reasons can be external or internal. Using proper screening method can help to save life. The regular screening technique utilized is mammogram, ultra sonography, MRI, Mammography (additionally called as Mastography) is a clinical imaging that utilizes low portion X-beam for checking the breast. Mammogram strategy utilizes different information base for example, Mammography Image Analysis Society (MIAS), Digital Database for Screening Mammography DDSM (DDSM) which are investigated for unusual discoveries. Ultrasound is utilized for additional analysing of masses found on mammography or for the majority not seen on the mammogram. MRI are utilized for assessment of the problematic discoveries and furthermore for screening pre-surgical assessment in patients with known breast cancer. Researchers consider the three features to detect and classify breast tumour: intensity, shape and texture. These features are additionally utilized in AI procedures for recognition and classification of breast tumours. This paper gives the per user the comprehensive survey of the methodology that has been utilized in the breast tumour recognition and classification.

Table 1. Estimated 5- year relative endurance pace of the breast cancer.

SEER Phase	5-yearRelativeEndurance pace(inpercent)
Restricted	99
Local	86
Far Off	27
Every SEER phases together	90

×

This survey presents an outline of Artificial Intelligence (AI) approaches such as machine and deep learning used to detect and classify breast cancer using images of tomography, ultra- sound, MRI, mammograms, etc. This paper is organized as follows: Section 2 will cover the basics and background for various Ultrasound Techniques, Calcification, ML algorithms, and Deep Learning techniques. Section 3 will deal with issues and research carried out to date in breast cancer detection and classification. Section 4 will address challenges and trends in the future. Finally, Section 5 brings the research review to a conclusion.

BASICS AND BACKGROUND

Ultrasound Techniques

Automated breast discovery techniques have been developed using several ultrasound techniques. A few ultrasound techniques are discussed in the following section among such strategies.

Automated Breast Ultrasound (ABU) Procedure

The ABU is the primary technology devised for the detection of breast cancer in adults, particularly targeting women with heavy breast tissues or no history of surgical procedures, biopsies, or breast surgery. High frequency and ultra-broadband capabilities will enhance details' resolution and contrast differentiation.

Breast Specific Gamma Imaging (BSGI)

The non-invasive diagnostic technology BSGI also known as Molecular Breast Imaging (MBI), has been brought into use for differentiating the tissues in the breast that retain relatively higher amounts of the radiotracers which emit gamma rays. Inject the radiotracers in the bloodstream following that the imaging tests will be performed with the gamma camera.

Scintimammography

CAD-with-MRI can be an alternate to mammography and other diagnosis on patients with abnormal results, to provide the power of detecting breast lesions. The operation of tactile breast imaging involves placing a sensor against the breast while it measures the response of the material underneath. The signals are then processed into multidimensional colour images as the clinician moves the handheld sensor from the underarm onto the breast. These images also let the clinician visualize on a computer screen how any suspicious masses look in shape and size and hardness and their location.

Sonography

Breast ultrasound, better known as sonography, is a way to image an area without using ionizing radiation. A probe is placed against the skin and generates sound waves while collecting the echoes reflected from body tissues. These echoes are further processed into binary images and displayed on a computer screen. Ultrasound is certainly the prime modality to characterize certain breast masses and ascertain whether certain areas are cysts. Cysts cannot be accurately diagnosed through physical examinations alone. Breast ultrasound can also assist doctors in guiding biopsy needles into a few breast lesions.

Computer-Aided Detection (CAD) systems

CAD systems rely on pattern recognition mechanisms to help radiologists interpret images and formulate reports. These systems have served to provide a standardization in the reporting of breast ultrasound.

Computer-Aided Detection with MRI

CAD has assisted radiologists in interpreting breast MRI images, particularly those enhanced with contrast. With such property, CAD coupled with MRI would serve as an alternative for mammogram and other diagnostic modalities. Advanced breast imaging techniques involved inserting a sensor into direct contact with the breast for data acquisition. Through the movement of the hand-held sensor over a woman and under her arm, signals would be processed to create multi-dimensional coloured images. Such images allow the clinician to visualize easily and assess the size, shape, consistency, and location of any suspicious masses put up on a computer screen.

Electrical Impedance Scanning (EIS)

EIS is used along with mammography as a confirmatory test. Abnormal breast tissue is indicated when small electrical currents are sent into the tissue. This application generates a conductivity map for the breast using electrical currents that malignant tissues generate since malignant tissues conduct electricity better than those that are normal. This automatically aids in identifying suspicious areas. A small electrode is attached to a handheld device, moved back and forth across the breast for sending electrical currents into the body (Kim Y, Kang BJ *et al*., 2016). A probe is placed on the skin surface measure.

Magnetic Resonance Elastography (MRE)

MRE of the breast is a specialized MRI technique used to examine the mechanical properties of tissues.It takes advantage of the increased stiffness of malignant tissues compared with normal tissues, particularly noting that breast cancers are firmer than healthy breast tissue. This difference is quantified by applying a known stress to the tissue and measuring the resultant deformation.

Magnetic Resonance Imaging (MRI)

MRI is a non-invasive imaging technique using magnetic and radio recurrence fields. To image the body tissues, MRI creates extremely detailed cross-sectional images of the intruder. MRI does not use ionizing radiation, unlike computed tomography (CT), and this process is known to be safe.

Calcification

It is also common for women, especially with age to discover calcium deposits within their breasts, which are referred to as "calcifications." It is so small that skin can't feel them; they are only visible on mammogram films as small, bright white spots. Such findings may lead to anxiety and additional tests; tight groups or lines of very small calcifications are frequently counted among breast cancer symptoms. About half of all mammograms done on women older than 50 show some calcification. Calcifications may result from a multitude of reasons, including advanced age, previous injuries, or infections in the breast. Calcification is not caused by calcium intake in a female's diet. Women with calcifications may also have undergone earlier breast surgeries or radiation treatments.

When non-palpable breast cancer begins to grow, calcifications are the first signs to appear. Calcifications are for the most part identified with the Ductal Carcinoma in Situ (DCIS). Notwithstanding, calcification may likewise happen in intrusive cancers. A breast imaging treatment defines it as a main sign of malignant tumours. According to screening programs, calcifications are the only sign of cancer growth in 41.2% of women. A greater portion of the calcifications can be seen on ultrasound images. Ultrasound RF signals in detecting some subtle patterns give an edge in less sensitivity to such patterns. These tend to show recurrence and scattering but offer data useful in the reconstruction of calcification features. Breast calcifications are indicators of tumour presence more likely found within the radiological evaluation of breast lesions. The American College of Radiology invented the Breast Imaging Reporting and Data System (BI-RADS), which standardized the reports of breast imaging to explain what breast imaging means. The BI-RADS atlas implements a set of rules to decipher calcifications. As a result of these guidelines, the radiologist was able to differentiate doubtful calcifications from normally benign variations. Among the examples are calcifications of the vascular system and the skin. A biopsy was recommended for patients with doubtful calcifications. Calcifications associated with benign diseases can unintentionally lead to bogus positive results. The most common type of cancer detected by calcifications is DCIS. Less than 20% of these cancers are low grade.

Types of Calcifications

Calcifications can be categorized in two types. These types are known as micro-calcification and macro-calcification.

Below is a brief description of these types of calcifications: *Macro-calcifications:* Macro-calcifications appear round and large on mammograms. Cancer is generally not associated with these types of calcifications. As a result, reporting is usually not necessary. *Micro-calcifications:* Micro-calcifications are very small in size. Clusters of these calcifications are possible. Such calcifications normally are benign or non-cancerous. This can be a sign of breast cancer. Generally, doctors can detect changes in the calcification in future. So, it is necessary to take tests for the detection of the cancer. It is never simple to determine whether micro-calcifications are benign (non-cancerous) or a symptom of breast cancer. As a result, multiple images must be obtained, usually from another mammography. These images help in classifying micro-calcifications as benign, probably benign, or perhaps cancerous. The area is not malignant if there are micro-calcifications that are benign or most likely benign. Contrarily, more follow-up tests are needed for micro-calcifications that seem suspicious.

ML Algorithms

ML is widely utilized in breast cancer conclusion throughout the long term. These algorithms identify several imminent data patterns. ML consciousness is a component of AI. A variety of statistical, probabilistic, and optimization tools are used in ML to automatically learn new skills and boost performance. Algorithms for ML can be used to both fresh input and historical data without explicit programming instructions. Both ML and statistics are used to analyse the data. Large, multidimensional, and complex data are dealt with through ML. In general, ML algorithms can extract key traits and probable rules. With the aid of traditional statistics, it is incredibly difficult to distinguish between these highlights and rules [64]. ML techniques are better suited for breast cancer datasets than statistical methods are. AI techniques are coordinated into scientific classification based on required after effect of the calculation. Some basic kinds of techniques are: 1. Supervised learning is that kind of learning where mapping input to expected output is achieved with the help of the learning algorithm. A common example of supervised learning is a classification task where, given a vector, a learner must attempt to understand a function that maps that vector to one of several classes by looking at many pairs of input-output values of that function. 2. Unsupervised learning: It performs the displaying of a bunch of inputs. In these calculations, named models do not occur. 3. Semi-supervised learning: Semi-supervised learning merges both labelled and unlabelled examples for generating a suitable function or classification model.

Following are some popular ML algorithms:

Naive Bayesian (NB) Networks

In simplest terms, Naive Bayes classifiers are Bayesian networks that consist of directed acyclic graphs with a single parent and many children. The children are observed variables, while the parent is unobserved; hence, these networks assume very strongly that its children are conditionally independent of their parent node. As a result, the freedom model (Naive Bayes) depends on evaluation [57]. Bayes classification models are typically less precise than other, more complex learning procedures, like ANNs.

Support Vector Machines (SVMs)

SVMs are among the latest supervised AI algorithms, closely related to traditional multilayer perceptron Neural Network (NN). These algorithms work by creating a margin on either side of a hyperplane that separates two classes of data. The goal is to maximize this margin, ensuring the greatest possible distance between the separating hyperplane and the data points on both sides, thereby minimizing the risk of generalization error.

Decision Tree

Using includes values as a basis for arrangement, decision trees (DT) classify instances. A DT hub each describe a component of the incident that was chosen for classification. A value that is addressed by each branch may be accepted by the hub. Instance classification is carried out from the root node's beginning. Then, these events are planned based on their elemental values. In data mining and AI, DT learning is used. This methodology uses a DT as a predictive model in this instance. The more descriptive names for these kinds of classification models are regression trees or classification trees.

k-means algorithm

A way of average grouping is the k-means. This calculation falls under the heading of a clustering technique based on partitioning. In several disciplines, the K-means clustering algorithm is used. Data recovery, computer vision, and pattern recognition are some of these disciplines. The mean (also known as the centroid) of each cluster's data points is used to calculate the k means for each cluster. This calculation's goal is to divide a set of n items into the necessary k clusters, which are supplied as input to the approach. High and low inter-cluster similitude is the result in the high intra-cluster likeness. For measuring similarity, the mean estimation of the articles in a cluster is used. K is randomly selected as the initial cluster centre. The items in clusters are relocated iteratively using this strategy, starting with one cluster and going on to the next in consideration of the Euclidean distance between the group centroid and the article. Following this transfer of articles inside clusters, the mean estimation of the items is once more registered for each cluster. This method produces a brand-new cluster centroid. This system is used repeatedly until the work on the guidelines converges.

Deep Learning Techniques

Deep learning is a segment of AI and ML which, in contrast to conventional ML extraction algorithms, concentrates an intricate hierarchy of image characteristics. In many recent computer models, deep learning algorithms are used to make significant progress in computers extracting image information. In many medical specialties, in particular pathology and radiology, such algorithms were used and, in some instances, performed comparably with human experts. Whether it is for extracting the prognosis from a given image, analysis of the molecular states within the image, or finding out how sensitive an image is to therapy, deep learning can retrieve much information from medical images that may be unreachable by human analysis. Deep Learning generally refers to multi-layered neural networks that intake the raw image input and analyses them to build a hierarchical structure of features. But modern advanced graphic processing has made the algorithms so powerful that they can train with millions of images and be robust against variations in images. Because of its success in imaging, deep learning has now become a hot topic in all literature on segmentation and classification. Many DL schemes are implemented for a plethora of applications, like object recognition, speech recognition, and disease, genotype, and phenotype classification. Stack auto encoders, Boltzmann deep machines, deep neural networks, and CNN are some of the most common DL methods.

Convolutional Neural Network (CNN)

CNN especially for the recognition and detection of face, language, human bodies, and biological images, has become a popular image processing technique. CNNs are the image algorithms most used. Since their inception, Image classification and segmentation are some of the prominent activities in which CNNs are adopted. Being a deep neural network, CNNs primarily act on visual images for classification. They are feed-forward networks that retain information regarding topological features of the image. CNNs are based on the multilayer perceptron model. As the name suggests, a multilayer perceptron is a fully connected network in which every neuron in a layer is connected to every neuron in the layer above it. A CNN consists of three main layers: convolution, pooling, and fully connected layers. Each layer serves a particular purpose: the convolutional layer extracts feature from the input, the fully connected layer classifies using these features, and the pooling layer acts to reduce the size of feature maps to save on computation cost and enhance the efficiency of the network.

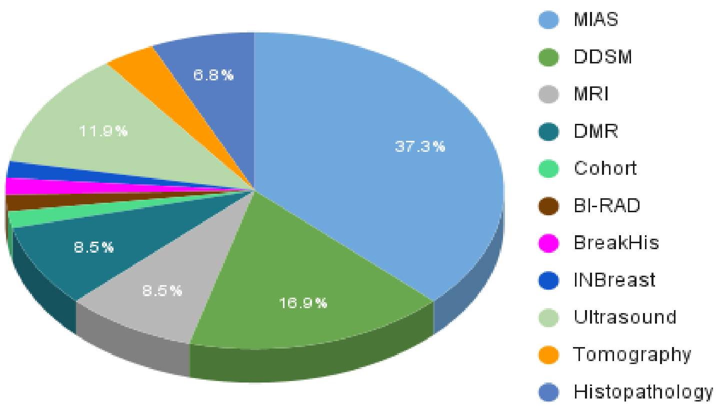

Figure 1. The data sets utilized in the studies are presented as pie chart.

Dataset

An outline of the available datasets utilized for numerous breast cancer classifications in this review is incorporated in this section. The INBreast, BI-RADS, MRI, Database of Mastology Research (DMR), DDSM, Breast Cancer Histopathological Image (BreakHis), and MIAS datasets were among the public datasets utilized for classification of breast cancer shown in Figure1.

Performance Metrics

The performance measures utilized to assess Imaging techniques are explained in this section. If breast cancer is detected correctly, it can be classified as true-negative (TN) or true- positive (TP), and if it is erroneously detected, it can be classified as false negative (FN) or false-positive (FP). The evaluation metrics used for the classification of breast cancer are Accuracy (Equation 1), Precision (Equation 2), Recall/Sensitivity/TPF (True Positive Fraction) (Equation 3), specificity/TNF (True Negative Fraction) (Equation 4), F-measure (F1-score) (Equation 5), AUC (Area under the curve) (Equation 6). These metrics are summarized as follows:

Accuracy (ACC)

$$ACC = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

Precision (PR)

$$PR = \frac{TP}{TP+FP} \quad (2)$$

Recall/TPF/Sensitivity (SN)

$$SN = \frac{TP}{TP+FN} \quad (3)$$

TNF/Specificity (SP)

$$SP = \frac{TN}{TN+FP} \quad (4)$$

F-measure (F1-score)

$$F - measure = \frac{2 \times PR \times SN}{PR+SN} \quad (5)$$

$$Weighted\ F - measure = \frac{1+\beta^2}{\beta^2} \times \frac{PR \times SN}{PR+SN} \quad (6)$$

Area Under the Curve (AUC)

$$AUC = \frac{\sum R_i(I_P) - I_P(I_P+1)/2}{I_P+I_n} \quad (7)$$

WhereI_n and I_P denote the number of negative and positive breast cancer images, respectively, while R_i is ith positive image rating.

Literature Survey

The following literature survey is divided into several parts based on the datasets used in the research works:

MIAS (MIAS)

Shrivastava and Bharti (2020) proposed a technique automatic pre-processing and efficient Seeded Region Growing (SRG) that extricate tumours from the images. In the initial step, for the pre-processing, noise reduction, contrast enhancement, and automatic extraction of ROI are finished. In the subsequent advance area distinguishing proof of the seed point utilizing the thickness of the pixels is finished. In the last, Seeded Region Growing (SRG) is calculated. The outcome estimated SRG is the extricated tumour. This yields the classification accuracy of 91.4%. Yousefikamal (2019) proposed a technique that comprises two principal parts: - Image classification and tumour region segmentation are the primary tasks. In the first step, the images are divided into two main categories: normal and abnormal. This classification is carried out using a CNN. The proposed method next emphasizes the analysis and segmentation of tumours in mammography images. Pre-processing is the primary step to eliminate noise and artifacts, while segmentation is performed by fuzzy c-means clustering. The method yielded an accuracy of 78%.

Wang et al. (2018) proposed research that utilizes gestalt psychology, in which they join human intellectual attributes with the information on radiologists in clinical image investigation. They isolated the proposed technique into three parts. This is another technique for the programmed recognition of lesions with an accuracy of 92%. Singh et al. (2018) researched a technique that depends on Content-based image retrieval (CBIR). In the pre-processing stage, an adaptive median filter is used for the image enhancement and removal of pectoral muscles and artifacts. In the segmentation stage, the seed region growing calculations are utilized to segment which finds co-occurrence thresh- olds. Moreover, a lot of features were extricated from these segmented regions. Then the feature selection is performed utilizing a minimum redundancy maximum relevance (MRMR) technique. At last, relative images were recuperated utilizing the Euclidean distance resemblance measure. This strategy yields an accuracy of 72%.

Htay and Maung (2018) performed feature extraction utilizing Grey-Level Co-occurrence Matrix (GLCM) and first order statistics. The pre-processing phase used some median filter for the devoicing of breast image and cropped the denoised image. After this, the breast area located on an uneven background was segmented using Otsu's thresholding approach. At last, this work used the K-nearest neighbour (KNN) classifier for classifying the breast images into regular or irregular based on extracted attributes. In the results, the presented algorithm validated on the mini - MIAS's database outperformed the existing breast cancer detection techniques by obtaining 92% accuracy. Kavya et al. (2018), authors could develop a cancer detector system that employs imaging methods such as mammography and thermography. The use of the CAD (computer-aided diagnostic) tool as a reliable method of segmenting and categorizing digital images. This method was used to examine the data that were taken from the hospital. The Cyber-Physical System (CPS) was employed to collect data and transmit it to designated systems. The network was integrated, where one could interact with the system. Thus, with the support of CPS, the system became more flexible, scalable, and optimized. The suggested method might identify breast cancer while ensuring high patient safety because mistake rates were reduced, and the data was being watched.

Varma and Sawant (2018) A technique is needed for quantifying texture complexity from images and for the digital analysis of mammograms. After that, the attributes were successfully extracted, leading to improved identification and appropriate action being taken to reduce the risks associated with breast cancer. The introduced approach could also be modified to shorten processing times and speed up processing. The final output images showed how the suggested method effectively outlined the breast tissue abnormalities and identified breast cancer. Lenzi et al. (2017) provided an introductory evaluation study of a novel mm-waves Modulated Gaussian Pulse

(MGP) radar imaging method depending upon the application of ANN for detecting breast cancer. The majority of presented Ultra-wide Band (UWB) radar imaging methods had the tendency to function by means of the pulse at the central frequency of few gigahertz, where this translated in a suboptimal imaging resolution. This work improved the resolution using an mm waves Modulated Gaussian Pulse (MGP) cantered at 30 GHz. Further, this work used ANN for the processing of the measured radar signals. The ANNs provided more efficient results with extremely low computation overhead and in quasi-real time.

Hepsag et al. (2017) researched a deep learning technique that utilized datasets of mini-MIAS in which a classification performance is made like each dataset. Firstly, pre-processing is performed on datasets that include cropping, augmentation, and balancing data. The CNN is utilized for classification. This yields an accuracy of about 72%. Sangeetha and Murthy (2017), the new mechanism which is developed is based on the integration of different highly performing digital image processing (DIP) techniques that had never been put into use in this area. The developed system enables early breast cancer detection from the microcalcifications and classifies the cells into normal or cancerous category. Greater accuracy was generated using the developed DIP mechanism which provided promising results concerning TP and TN.

Nezhadian and Rashidi (2017), the emphasis was on extracting features by not excising the pectoral muscle in preprocessing. In this study, the new and most efficient technique was used to train an SVM. The block diagram, implemented in the second section, was used for preprocessing. As an additional matter, DWT features were extracted from the images. Zigzag scanning would convert the approximation matrices into a one-dimensional metric. The study patients were classified as benign or malignant based on a MIAS image database. The obtained results revealed that the accuracy of the proposed algorithm roughly reached 95.80 percent. Vikhe and Thool (2016) proposed a procedure that utilized wavelet processing and adaptive thresholding. The pre-processing, mass enhancement, artifact suppression, and pectoral muscle removal are finished utilizing wavelet processing. At that point, mass segmentation for detection is performed utilizing an adaptive threshold procedure. The method yields an accuracy of 93.2%.

Singh et al. (2016) researched a CAD technique that utilized MIAS (MIAS) dataset. In the pre-processing phase, images are cropped to remove pectoral muscle and unwanted tags, and then the noise is removed utilizing a Gaussian filter. Further, features are extricated through the GLCM. Then feature selection is done utilizing the AdaBoost method. Finally, Random Forest (RF) Classifier is used for classification. This yields the maximum accuracy of 93.90%. Al-Hadidi et al. (2016), an acceleration technique was developed to segment mammography images using a Graphics Processing Unit (GPU). The commonly used Single Pass FCM method was modified to detect breast cancer more effectively. This technique was applied to a set of mammogram images to differentiate between malignant and benign cases. A comparative analysis was conducted to evaluate the performance of the developed technique, focusing on execution time and speedup parameters.

Al-Ayyoub et al. (2016) In this work, a mammographic segmentation acceleration scheme was developed employing Graphics Processing Unit (GPU) hardware. Efficiency in breast cancer detection was enhanced by modifying an existing Single Pass FCM technique. Thus, the modified technique was applied to a set of mammograms to differentiate between malignant and benign cases. Performance evaluation of this proposed technique was carried out by way of comparison based on execution time and speedup parameters. Saubhagya et al. (2016), In this proposal, an ANN model is suggested for the breast cancer diagnosis from mammograms. Firstly, microcalcifications were enhanced based on their illumination and irregularity. Then an iterative threshold selection method has been applied for the locating of microcalcifications while the shapes of micro-calcifications would get modified and then isolated pixels deleted mathematically. Finally, the extracted features relevant to breast cancer were identified and enabled mammography images for detecting breast cancer to be analysed.

Shirazi and Rashedi (2016), This implementation uses SVM and MGSAS algorithms in detection of breast tumours in mammographic images. The feature extraction is done using GLCM, while the MGSAS-SVM method rather considers the number of attributes involved in improving the performance of SVM-based classification. Selected features with optimal SVM algorithm will assist in tumour detection. The experimental results have shown that the suggested approach effectively optimizes both feature selection and SVM parameter tuning in breast tumour detection methods. Swetha and Bindu (2015), a hybrid image segmentation technique was proposed that merges a fast-sweeping algorithm and dual front evolution with Laplacian or gradient-based methods along with Otsu's thresholding using ten threshold levels. Tu and Wong's methods can locate the breast tumours very well. The process proves to remove artifacts and detect breast tumours accurately. Moreover, the size and stage of the tumour were assessed using this method. The method suggested was very competent, less complex computationally, and helped diagonal edge detection in mammography images.

Nasiri et al. (2015), It is observed that detection of breast tumours in mammographic images has become a Herculean task for any radiologist. A technique, based on SVM, has been employed to classify mammograms, wherein the features have been extracted through image processing. Wavelet and contourlet transforms were used to extract the feature vectors from the mammograms. The classification process employed the proposed technique. Experiments showed that the method was more efficient than acknowledged algorithms. Basheer and Mohammed (2013) proposed a technique of segmenting wherein the contour of the image is obtained using an adaptive median filter. The selection of contour satisfies the Region Of Interest (ROI). This research yields an accuracy of 92.3%.

Gorgel et al. (2009) researched a methodology that depends on the blend of SVM and wavelet-based sub-band image disintegration which utilized a dataset of 66 digitized MIAS images. Feature extraction is done by figuring wavelet coefficients and classification is performed utilizing SVM classifier with Radial Basis Function (RBF). Then the segmentation is performed manually. The classification accuracy yield is 84.8%. Adams and Bischof (1994) research a technique for segmenting of intensity images which utilized seeded region growing. It requires the contribution of various seeds, either individual pixels or regions, which control the development of regions where the image will be segmented. The strategy creates an accuracy of 78.2%.

Digital Database for Screening and Mammography (DDSM)

Wang et al. (2021) recommended to use End-to-End (E2E) training for Deep Convolution Neural Network (DCNN) to detect breast cancer. First, the image attributes were extracted using a pre-train algorithm having custom layers. Thereafter, a feature fusion module was implemented for quantifying the weight of every feature vector. This method provided diverse sensibility to distinct instances of every case. In the end, the fused attributes were classified through a classifier module. The developed clinical dataset and 2 publicly available datasets were utilized for conducting the experiments. The results depicted that the suggested techniques were efficient. Shrivastava and Bharti (2020) proposed a technique automatic pre-processing and efficient Seeded Region Growing (SRG) that extricate tumours from the images. In the initial step, for the pre-processing, noise reduction, contrast enhancement, and automatic extraction of ROI are finished. In the subsequent advance area distinguishing proof of the seed point utilizing the thickness of the pixels is finished. In the last, Seeded Region Growing (SRG) is calculated. The outcome estimated SRG is the extricated tumour. This yields the classification accuracy of 91.4%.

Deep Deb et al. (2020) presented an approach which deals with extracting features from dissimilar convolutional layers of the Deep CNN (DCNN) based on Global Average Pooling (GAP). This approach concatenated all the extracted attributes prior to the ultimate classification process. In this work, six different already trained DCNN models were used. These networks were trained using imagnet dataset and revealed the features extracted from diverse convolutional blocks to improve the classification efficiency. The tested results depicted that the role of ROI was quite significant in both the finer and coarser details. Singh et al. (2020) proposed the technique to segment the image utilizing conditional Generative Adversarial Network (cGAN) which finds the ROI by segmenting. The adversarial network discovers how to find ground truth and binary masks are formed by segmenting. At that point for the shape classification descriptor, CNN is utilized which produced four shapes of tumours: oval, round, lobed, and irregular. The strategy yields an accuracy of 80%.

Samala et al. (2019) proposed the strategy that utilizes CNN prepared on multi- stage adjusted exchange learning for 2-stage learning. In the single exchange learning, CNN is prepared on ImageNet information which is adjusted straightforwardly with the digital breast tom synthesis (DBT) which classifies the tumours. In the multi-stage transfer learning, Image-Netare fine-tuned twice firstly with the mammography information and secondly with the DBT information. At that point, it uses 2- transfer learning. The strategy yield accuracy of 91%. Ragab et al. (2019) proposed a strategy in which two segmentation approaches are utilized. The primary methodology includes deciding the ROI physically, while the subsequent methodology utilizes the procedure of thresholding and region based. In feature extraction, a DCNN is utilized and is calibrated to classify two classes. Lastly, SVM classifier is used. The strategy yield accuracy of 80.5%.

Wang et al. (2018) proposed research that utilizes gestalt psychology, in which they join human intellectual attributes with the information on radiologists in clinical image investigation. They isolated the proposed technique into three parts. This is another technique for the programmed recognition of lesions with an accuracy of 92% with the MIAS dataset and 93% with the DDSM dataset. Vikhe and Thool (2016) proposed a procedure that utilized wavelet processing and adaptive thresholding. The pre-processing, mass enhancement, artifact suppression, and pectoral muscle removal are finished utilizing wavelet processing. At that point, mass segmentation for detection is performed utilizing an adaptive threshold procedure. The method yields an accuracy of 76%.

Liu et al. (2015) proposed a technique that uses a sliding window scheme to scan the breast area segmented. For every current window, the Histogram of Oriented Gradient (HOG) is extracted and fed to a supervised algorithm, Kernel-Based Supervised Hashing (KSH) to obtain the corresponding compact binary code. A specific decision rule to classify the current window in hamming space is used. Then a flexible sliding window fusion algorithm is proposed to label the detected mass region more accurately. The method yields a TPF value of 94%. Hiba et al. (2016) proposed a method which first extract features using Local Binary Pattern (LBP) then the SVM is used to classify the images. The accuracy yield is 91%.

Magnetic Resonance Imaging (MRI)

Zhang et al. (2020), utilizing a Mask R-CNN, the entire image set was processed to detect suspicious lesions. The tumour segmentation was done through the FCM algorithm, thereby confirming unilateral mass lesions for each of the patients. There were two datasets from DCE-MRI; the training part consisted of 241 patients with a non-fat sequence, while the test part was comprised of 98 patients with a fat-sat sequence. To test the method above, we employed a free-response receiver operating characteristic (FROC) analysis. The algorithm has been currently associated with an average accuracy of about 0.86.

Shrivastava and Bharti (2020) proposed a technique automatic pre-processing and efficient Seeded Region Growing (SRG0 that extricate tumour from the images. In the initial step, for the pre-processing, noise reduction, contrast enhancement, and automatic extraction of ROI are finished. In the subsequent advance area distinguishing proof of the seed point utilizing the thickness of the pixels is finished. In the last, Seeded Region Growing (SRG) is calculated. The outcome estimated SRG is the extricated tumour. This yields the classification accuracy of 91.4%.

Table 2. Summary of publications describing the various strategies used to analyze images from the MIAS dataset.

Author	Techniques	Parameters
Aqsa et al (2022) [6]	Modified YOLOv5 Network	Accuracy = 96.5 %
Shrivastava and Bharti (2020) [47]	Automatic Pre-processing and Efficient SRG	Accuracy = 91.4 %
Yousefikamal (2019) [66]	CNN	Accuracy = 78%
Wang et al (2018) [60]	Gestalt Psychology	Accuracy = 92%
Singh et al (2018) [52]	Automated CBIR	Accuracy = 80%
Htay and Maung (2018) [17]	GLCM, k-NN	Accuracy = 92%
Kavya et al (2018) [21]	CLAHE, LSVM	Accuracy = 93.3 %
Varmaand Sawant(2018)[58]	Automatic pre-processing and Efficient SRG(Image Processing)	Accuracy = 86%
Lenzietal.(2017)[25]	MGP,ANN	Accuracy = 86%
Hepsagetal. (2017)[14]	CNN	Accuracy = 72%
Sangeethaand Murthy (2017) [43]	CLAHE, Otsu, Bayes Classifier (Simulation)	Accuracy = 86%
NezhadianandRashidi (2017) [32]	CLAHE, SVM	Accuracy = 86%-95%
VikheandThool(2016)[59]	Wavelet Processing and Adaptive Thresholding	Accuracy = 71%
Singhetal.(2016)[51]	GLCM, AdaBoost, RF	Accuracy = 93.90%
Al-Hadidietal.(2016)[4]	LR, BPNN	Accuracy=93%
Al-Ayyoubetal. (2016) [2]	FCM	Accuracy=87%
Saubhagyaetal. (2016) [44]	Sobel, Thresholding, NN	Accuracy=87%
ShiraziandRashedi(2016)[46]	GLCM, MGSA-SVM (Mixed Gravitational Search Algorithm with SVM)	Accuracy=93%
SwethaandBindu(2015)[55]	Hybrid Image segmentation, Otsu Thresholding	Accuracy=88%
Nasirietal.(2015)[31]	Wavelet and Contourlet Transform, PCA, SVM	Accuracy=82%
Basheerand Mohammed (2013) [7]	CLAHE, Adaptive Median Filtering	Accuracy=92%
Gorgeletal. (2009)[12]	Wavelet Based Subband Image Decomposition, SVM with RBF	Accuracy=84.8%
Adams and Bischof (1994)[1]	Seeded Region Growing (SRG)	Accuracy=79%

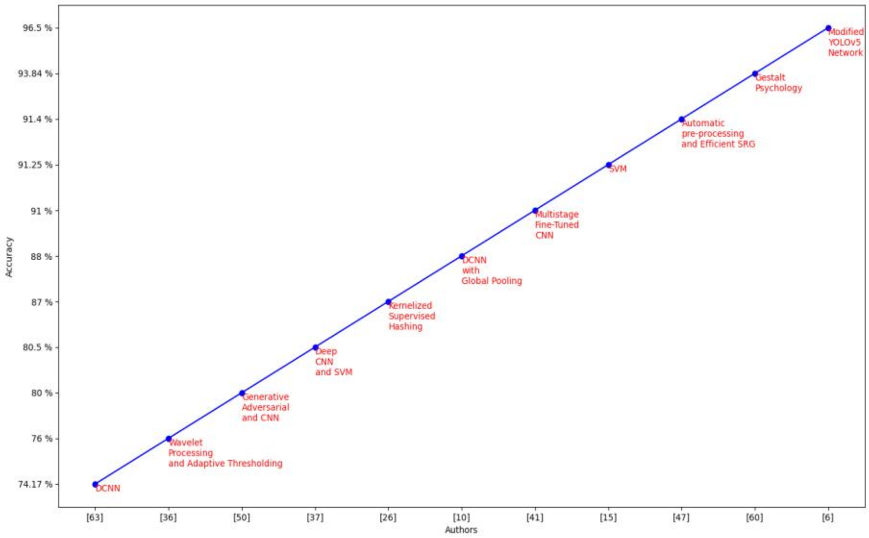

Figure 2. Summary of publications describing the various strategies used to analyze images from the DDSM dataset.

Chiang et al. (2018) intended a fast and efficient CAD technique planned based on 3-D CNN. Initially, the implementation of an effective sliding window technique was done for extracting the Volumes Of Interest (VOI). Subsequently, this system assisted in estimating every VOI and choosing the VOI having a higher estimated probability as tumour candidates. The issue of over-aggregation was addressed by prioritizing candidates based on the estimated tumour probability. A test set of 171 tumours was used to evaluate the proposed technique. The results demonstrated that the intended technique provided a sensitivity of around 95% and worked more quickly than other methods.

Gamil et al. (2018), an automatic CAD technique has been proposed that detects early breast tumours from ultrasound images. The method consists of several modules with its schemes and mechanisms. Initially, an input ultrasound image pre-processes and filters and then the set of features computed and extracted after segmenting the ROI from background noise. Tumour classification using ML classifiers. The method developed has been implemented up to the present time with improvements in runtime and about 10% accuracy gain as compared to previously existing methods.

Ponnusamy and Babu (2016), Joint Probabilistic Seed Selection and Fuzzy Region Growing (JPSS-FRG) presents a method based on an improved region-growing algorithm. A seed point is selected within the mass region, aided by the Joint Probabilistic Relevance Factor, which reduces the breast segmentation time. Afterward, a fuzzy logic-based region-growing algorithm is applied from this seed point. A Gradient Vector tumour segmentation technique is finally proposed to facilitate the early detection of tumour cells. This method obtains a True Positive Fraction (TPF) value of 82.3%. Al-Faris et al. (2014) proposed a strategy that utilizations modified automatic seed growing region dependent on Particle Swarm Optimization (PSO) for image clustering. For pre-processing, images are split into two sub-images, If the image is Axial: right and left breast image done by tracking down the centre of the X-coordinate of the image and parting the image vertically starting there and afterward noise is removed utilizing the median filter. For breast skin detection and deletion, incorporation of Level Set Active Contour calculations with Morphological Thinning calculations is utilized. At last, for Tumour segmentation, the SRG algorithm is utilized that utilizes the image clustering based on PSO. The method yields a TPF value of 79%.

Ultrasound

Wei et al. (2020), an algorithm was created to detect and classify breast tumour lesions, utilizing Faster R-CNN (Region Convolution Neural Networks) on ultrasound images. The algorithm can detect breast lesions with a bounding box and classify them as either cancer or normal. The algorithm was trained and tested on an open-access database of ultrasound images of breast lesions. The results of the test indicated that the algorithm successfully identifies and classifies breast lesions with a success rate of over 95%. Khasana et al. (2020). The goal is to implement a watershed transform algorithm to aid in the segmentation process, which will help locate the tumour and differentiate the objects from the background. After that, the thresholding binaries were used to divide the image of the tumour into its component parts. The area of the cancer was calculated at the last stage. The results showed that the anticipated algorithm had produced an ER of 11.3% and an accuracy of about 88.6% for all the data analysed. The results of the tests indicated that the anticipated algorithm might successfully identify a breast tumour from an ultrasound image.

Wang et al. (2019), To improve the detection sensitivity effectively, a densely deep supervision technique was proposed. Multi-layer properties were exploited to achieve this. In addition, a threshold loss was introduced to provide a voxel-level adaptive threshold, enabling the difference between malignant and normal images. As a result, low FPs and great sensitivity were achieved. To determine the anticipated technique's accuracy, a dataset composed of 745 cancer regions and collected from 219 patients was used. The results of the trial showed that the projected technique's sensitivity was 95% accurate with 0.84 false positives. The proposed method of analysing the breast with the use of ABUS resulted in an efficient cancer detection system. Hoque et al. (2019), an automated method is developed for breast cancer detection using an ultrasound greyscale image. To segment unwanted noise removed with a median filter, Gray images for testing were processed with Otsu's thresholding method. Finally, key attributes typically used by pathologists were employed to define the ROI in the breast tissue for diagnosis. This technique had a specificity of 93.11%. Results from experiments showed that the proposed method surpassed traditional methods in terms of accuracy.

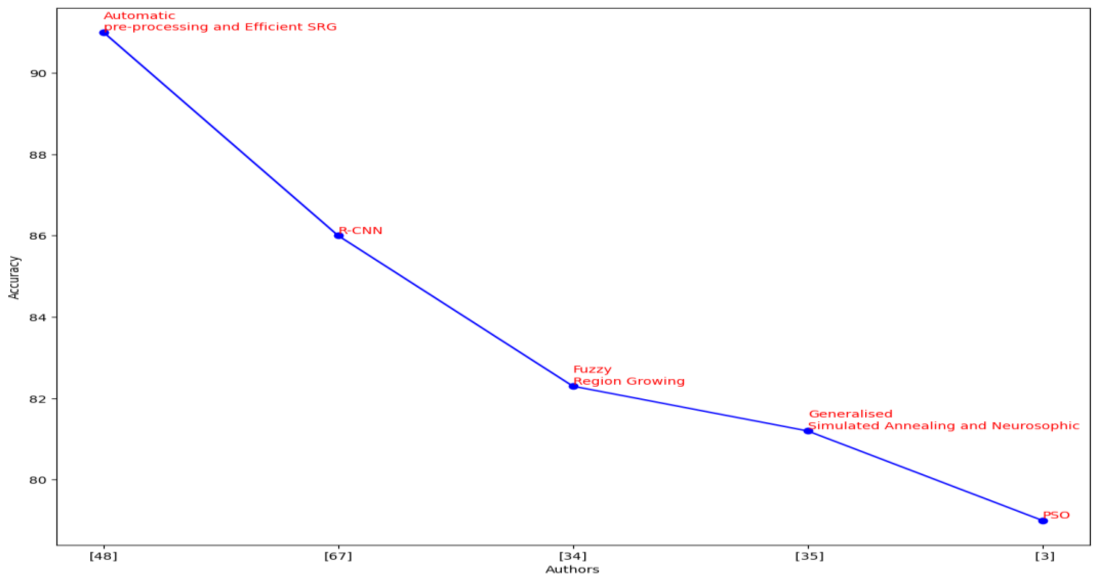

Figure 3. Summary of publications describing the various strategies used to analyse images from the MRI dataset.

Prabhakar and Poonguzhali (2017), an algorithm has been proposed to automate the segmentation and classification of breast lesions from ultrasound images. To decrease speckle noise, the Tetrolet filter was applied. Segmentation of lesions was done using an active contour method based on statistical attributes. Fifteen textural features, twenty-one morphological features, and four fractal features were extracted from the images during the segmentation process. For classification purposes, a SVM with a polynomial kernel was employed for classification based on the extracted attributes. This method assists radiologists with automatic detection and classification of lesions.

Tomography

De Jesus Aragao et al. (2020). An advanced co-focal imaging algorithm was developed for detecting breast cancer using radar-based Ultra-Wideband (UWB) technology. By using an example of a mixed breast, which had a tumour located in a ductal area near the glandular tissue, waveform sets created from an artificial simulation system have been used to demonstrate the robustness of the proposed algorithm. Results revealed that the system could recognize and localize the breast tumour more efficiently but at only a 2% additional cost of computation, even when faced with a heterogeneous, thick breast. Ronagh and Eshghi (2019). A new hybrid optimization technique proposed for the solution of the inverse scattering problem is the combination of Binary Genetic Algorithm (BGA) with Binary PSO (BPSO). The hybrid method converged at a rate four times faster than that of the ordinary BGA. It was able to reconstruct not just heterogeneous breast tissues but also gave a quantitative image of the breast's permittivity and conductivity profiles. The method thus was seen to be effective in size and localization of tumours even when surrounded by benign and fibro-glandular tissues.

Database of Mastology Research (DMR)

Kiymet et al. (2019), A novel technique for breast cancer detection was developed, based on thermal imaging. The detection procedure analysed thermographic breast images with the help of four different deep learning (DL) algorithms. The experimental results demonstrated that the ResNet50 network was efficiently utilized, attaining a testing accuracy of 88.89% for breast cancer detection. The technique developed has been shown to be flexible and effective for breast cancer detection.

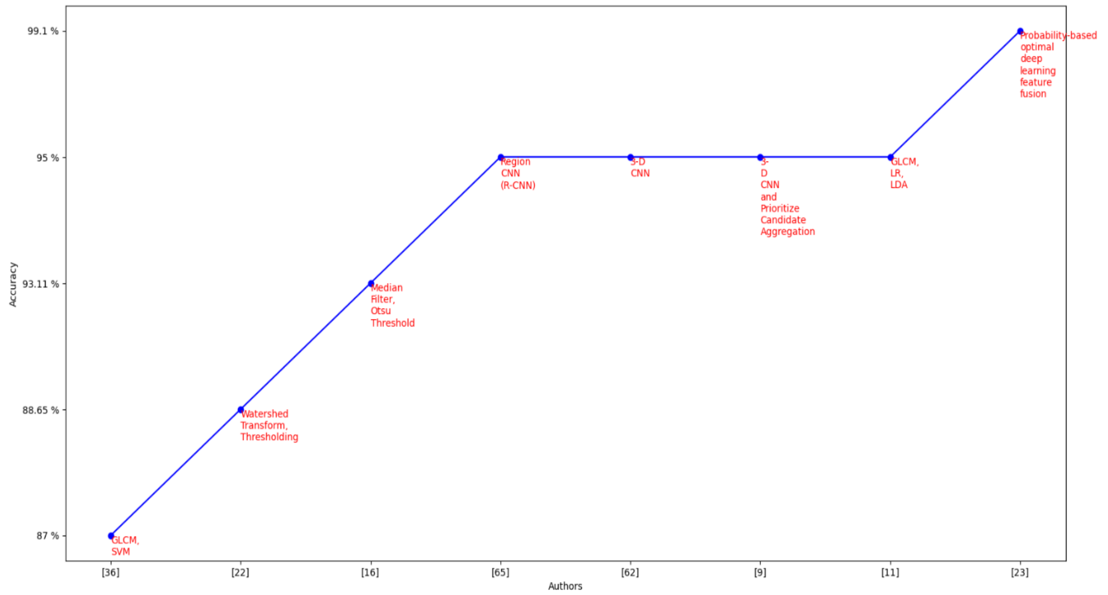

Figure 4. Summary of publications describing the various strategies used to analyse images of Ultrasound.

Table 3. Summary of publications describing the various strategies used to analyze images of Tomography.

Author	Dataset	Techniques	Outcome
De Jesus Aragao et al (2020) [19]	Tomography	Delay and Sum (DAS) Confocal Algorithm	With 2% increase in Computation cost accurately identify tumour
Ronagh and Eshghi (2019) [39]	Tomography	Hybrid BGA with PSO (HBGAPSO)	Improved performance compared to BGA based algorithm

Ismail and Sovuthy (2019) A deep learning (DL) method was developed that utilizes the VGG16 and ResNet50 networks to detect normal and abnormal breast cancer. The process involved image pre-processing, followed by classification, and the performance was evaluated in this approach. In this evaluation, three indicators were used to gauge the performance of classification techniques. The results showed an accuracy of 94% for VGG16 and 91.7% for ResNet50. Kavya et al. (2018) The present system for breast cancer detection comprised different kinds of imaging methods such as mammography and thermography. This was well adopted by the CAD tool for segmenting and classifying the digital images. This method was used to examine the data that were taken from the hospital. The CPS was utilized to gather the information and distribute it to certain systems. Additionally, the network was integrated, people interacted with the system, and with the aid of CPS, the system was made adaptable, scalable, and optimized. The approach was said to be capable of identifying breast cancer while offering patients a high level of safety because error rates were reduced, and the data was being monitored.

Soliman et al. (2018) An image processing system followed by image analysis towards the proper diagnosis and correct treatment of breast cancer has been framed. The ROI, whose segmentation was carried out from the thermal input image, was used to extract the key breast features. Based on these attributes, a NN (neural network) classification technique was then used to classify the image as malignant or normal. The designed system was quantified using a benchmark dataset, and a success rate of roughly 96.51% was found. The results showed that the system was effective as intended. Al Rasyid et al. (2018), the intensity values of the grayscale thermograms were analysed for benign and malignant breasts. Histogram statistics and GLCM were used for extracting features, while experiments were performed on 18 healthy and 21 cancerous breast images taken from the DMR dataset. The results showed that common statistical attributes such as mean, entropy and skewness were good in showing whether the thermograms pointed to the presence of cancerous cells.

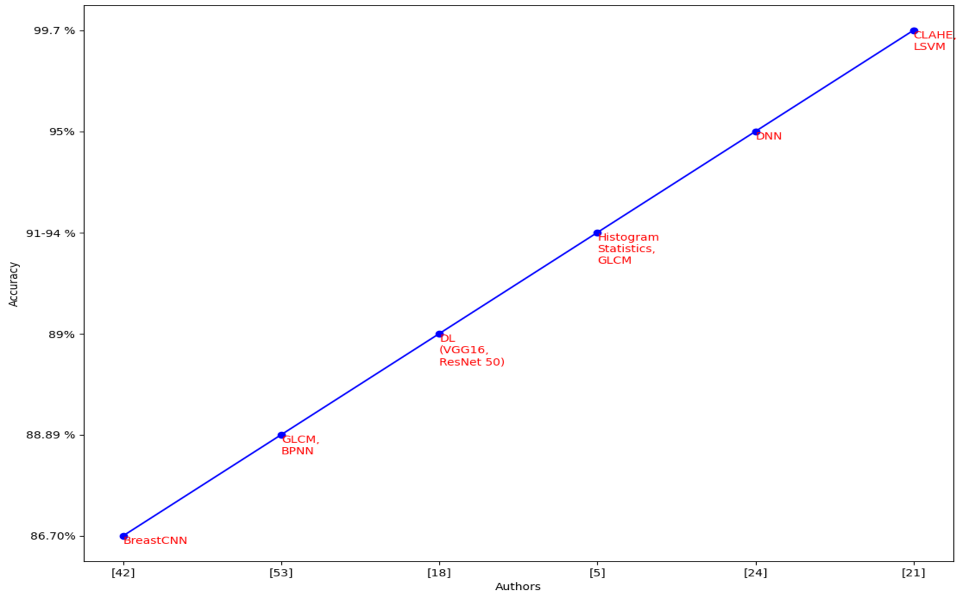

Figure 5. Summary of publications describing the various strategies used to analyze images of Database of Mastology Research (DMR) dataset.

Others

Wang et al. (2021) recommended to use End-to-End (E2E) training for DCNN to detect breast cancer. First, the image attributes were extracted using a pre-train algorithm having custom layers. Thereafter, a feature fusion module was implemented for quantifying the weight of every feature vector. This method provided diverse sensibility to distinct instances of every case. In the end, the fused attributes were classified through a classifier module. The developed clinical dataset and 2 publicly available datasets were utilized for conducting the experiments. The results depicted that the suggested techniques were efficient. Singh and Kumar (2020), diagnosis and classification of breast tumours depended significantly on histopathological attributes. Classification using KNN, Random Forest, and six types of SVM algorithms were used in the tests. The public Break His dataset was employed, considering various metric parameters for experimentation. The results indicated an accuracy of almost 92.3% achieved by the proposed method using cubic SVM in the detection and classifying of breast cancer.

Lu et al. (2019) discussed the relevance of CNN for classifying breast cancer. This work performed pre-processing of more than 9,000 mammogram images using median filter, data augmentation, and Contrast Limited Adaptive Histogram (CLAHE). Further, this work used CNN to train a classifier model. The tested outcomes depicted that the presented model performed considerably better with pre-processed images than without pre-processed images. Cai et al. (2019), an effective approach for mitosis detection in breast cancer histology images was devised using advanced version RCNN. The generalization performance of the technique was established by cross validating the datasets. The accuracy and recall obtained were 0.76 and 0.72, respectively, which were better than the previous results. Further, the proposed technique involving RCNN was proved to be efficient and qualified for clinical use.

Narayanan et al. (2019), a novel DCNN architecture was proposed for the classification and detection of breast carcinoma. The classification process was carried out using the Breast Histopathology Images dataset available on Kaggle. The correct regions for invasive ductal carcinoma (IDC) were extracted and marked in the images of the dataset. The dataset consisted of 277524 patches, with 198738 from the negative class and 78786 from the positive class. Results show an improved performance of the proposed model in relation with other existing state-of-the-art algorithms.

Mohebian et al. (2017) proposed a method that uses Clinicopathologic attributes of 579 breast cancer patients. First, the images were investigated then statistical feature selection technique is utilized for feature selection. Then additional unwanted features are removed utilizing PSO for absolute feature selection then classification is done using ensemble learning (BDT) with hold out and four-fold cross-validation. Further for comparison, three distinct classifiers

DT, Multilayer Perceptron neural network, Support Vector Classifier (SVM) is utilized. This yields an accuracy of 85%. Johra and Shuvo (2016), a single pipeline was investigated that can achieve breast cancer detection together with the feature extraction using the open-source image analysis software Cell Profiler. An algorithm based on a fuzzy inference system classifies the tumour as benign or malignant. Metrics such as accuracy, sensitivity, and specificity were used to compare the proposed approach with ANN and SVM. The results reported the approach investigated as superior to the others with a sensitivity of 95.6% and a specificity of approximately 90.63%. Paul and Mukherjee (2015) proposed a fast and efficient technique for mitotic detection in histopathological images: the cell segments using area morphological scale space with a novel scheme devised to generate scale space. The scales were restricted based on maximizing the relative entropy between the cells and their background for cell segmentation accuracy. The Random Forest (RF) algorithm classified the segmented cells. Our experiments demonstrated that the technique improved the F1-score by 12% over more than 450 histopathological images at 40× magnification.

Table 4. Summary of publications describing the various strategies used to analyse images of Others Publicly Available Datasets.

Author	Dataset	Techniques	Parameters
V. K. Reshma et al (2022) [57]	Histopathology Images	Deep Learning Techniques	Accuracy = 89.13 %
Wang et al (2021) [64]	INBreast	DCNN	Accuracy = 74.17%
Singh and Kumar (2020) [50]	BreakHis	GLCM, multi-SVM	Accuracy = 92.3%
Lu et al (2019) [27]	BI-RADS	Contrast-Limited Adaptive Histogram Equalization (CLAHE) Transfer Leaning	Accuracy = 88%
Cai et al (2019) [8]	Histopathology Images (TU-PAC 2016, ICPR2014)	RCNN	F1 Score = 0.731-0.736
Narayanan et al (2019) [30]	Histopathology image	DCNN	AUC = 0.935
Mohebian et al (2017) [29]	Cohort Dataset	PSO for feature selection, Bagged Decision Tree (BDT) or Ensemble Learning for classification	Accuracy = 85%
Johra and Shuvo (2016) [20]	Histopathology	Fuzzy Logic (Crisp)	Accuracy = 94%
Paul and Mukherjee (2015) [33]	Histopathology	Automatic Segmentation, RF	Precision = 0.82
Vinit et al, 2025 [68]	DDSM	CAMR_GF_BCC	Accuracy= 99.48%
Vinit et al, 2025 [68]	MIAS	CAMR_GF_BCC	Accuracy= 99.20%
Prinda et al, 2023 [69]	Augmented Dataset	Deep Learning Technique	Accuracy = 95.07%
YagedAlashban, 2024 [70]	TCIA	Deep Learning Technique	Accuracy = 96.21 %

Challenges and Future Trends

Future research directions to detect and classify breast cancer are described in this section, and considerable endeavours are necessary to improve breast cancer classification efficiency. This review finds out some challenges and limitations that need to be resolved to detect and classify breast cancer. Several future works and challenges are described as follows: Approaches and ideas that worked well can be extracted combined with further improvement. Especially when combined with the deep learning methods used for precise features extraction. Many new approaches of deep learning, ML can also be applied that yield good results that are trustworthy enough to be applied in the real life. Most of the reviewed works use the supervised learning methods for breast cancer classification in which for training annotated images are used and can achieve good results through these methods. However, gathering breast cancer images with proper labels which are classified by expert doctors is difficult in real life. So, the model based on unsupervised learning methods is frantically required to classify breast cancer that can be trained using an unlabelled images dataset. The limited availability of data is the main hurdle to realizing Deep Learning's effectiveness in medical imaging. Deep Learning classifiers are dependent on a good quantity and quality of the dataset. For medical imaging, the absence of sufficient training datasets is a vital barrier in the deep learning model training, according to the studies we reviewed. Furthermore, annotating vast amounts of medical imaging data is challenging since it takes a lot of effort and time not only from one person but from numerous specialists to eliminate human error. It's also tough to put together vast amounts of medical imaging data. Empowering a ML model to learn from the surrounding environment is also a big challenge. The main issue is obtaining sufficient breast cancer images sample to symbolize all breast cancer. To overcome this challenge methodology of reinforcement learning can be used, thus using this methodology for medical imaging can significantly improve the efficiency and efficacy of breast cancer classification systems.

CONCLUSION

Various approaches used in breast tumour detection and classification have been summarized and presented in the paper. In the review, strengths of various popular algorithms, such as Naive Bayes, KNN, DT, and SVM, are discussed. In addition, CNN and other Deep Learning techniques for the detection and classification of breast cancer through various image modalities are elaborated upon. This review discusses algorithms employed for optimizing the segmentation and classification of images into benign and malignant tumours. Many of these methods attained acceptable results with satisfactory accuracy, earning over 80%.

ACKNOWLEDGMENT

The authors express sincere thanks to the Head of the Department of CSE, IPS Academy, Institute of Engg. & Science, Indore, Madhya Pradesh, India for the facilities provided to carry out this research work.

CONFLICT OF INTERESTS

The authors declare no conflict of interest

ETHICS APPROVAL

Not applicable

FUNDING

This study received no specific funding from public, commercial, or not-for-profit funding agencies.

AI TOOL DECLARATION

The authors declares that no AI and related tools are used to write the scientific content of this manuscript.

DATA AVAILABILITY

Data will be available on request

REFERENCES

Adams, R., & Bischof, L. (1994). A new method of interpolation and smooth curve fitting based on local procedures. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 16*(6), 641–647. https://doi.org/10.1109/34.295913

Al-Ayyoub, M., AlZu'bi, S. M., Jararweh, Y., & Alsmirat, M. A. (2016). A GPU-based breast cancer detection system using single pass fuzzy c-means clustering algorithm. *2016 5th International Conference on Multimedia Computing and Systems (ICMCS)* (pp. 650–654). https://doi.org/10.1109/ICMCS.2016.7905595

Al-Faris, A. Q., Ngah, U. K., Isa, N. A. M., & Shuaib, I. L. (2014). Breast MRI tumour segmentation using modified automatic seeded region growing based on particle swarm optimization image clustering. In *Soft Computing in Industrial Applications* (pp. 49–60). Springer. https://doi.org/10.1007/978-3-319-06041-3_5

Al-Hadidi, M., Alarabeyyat, A., & Alhanahnah, M. (2016). Breast cancer detection using k-nearest neighbor ML algorithm. *2016 9th International Conference on Developments in eSystems Engineering (DeSE)* (pp. 35–39). https://doi.org/10.1109/DeSE.2016.8

Al Rasyid, M. B., Arnia, F., & Munadi, K. (2018). Histogram statistics and GLCM features of breast thermograms for early cancer detection. *2018 International ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI-NCON)* (pp. 120–124). https://doi.org/10.1109/ECTI-NCON.2018.8601449

Mohiyuddin, A., Basharat, A., Ghami, U., Vesely, P., Abbas, S., Naeem, O. B., & Rizwan, M. (2022). Breast tumour detection and classification in mammogram images using modified YOLOv5 network. *Computational and Mathematical Methods in Medicine, 2022*, 1–16. https://doi.org/10.1155/2022/1359019

Basheer, N. M., & Mohammed, M. H. (2013). Segmentation of breast masses in digital mammograms using adaptive median filtering and texture analysis. *International Journal of Recent Technology and Engineering, 2*(1), 39–43.

Cai, D., Sun, X., Zhou, N., Han, X., & Yao, J. (2019). Efficient mitosis detection in breast cancer histology images by RCNN. *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)* (pp. 919–922). https://doi.org/10.1109/ISBI.2019.8759461

Chiang, T. C., Huang, Y. S., Chen, R. T., Huang, C. S., & Chang, R. F. (2018). Tumour detection in automated breast ultrasound using 3-D CNN and prioritized candidate aggregation. *IEEE Transactions on Medical Imaging, 38*(1), 240–249. https://doi.org/10.1109/TMI.2018.2860257

Deep Deb, S., Rahman, M. A., & Jha, R. K. (2020). Breast cancer detection and classification using global pooling. *2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)* (pp. 1–5). https://doi.org/10.1109/ICCCNT49239.2020.9225375

Gamil, M. E., Fouad, M. M., Abd El Ghany, M. A., & Hoffinan, K. (2018). Fully automated CADx for early breast cancer detection using image processing and ML. *2018 30th International Conference on Microelectronics (ICM)* (pp. 108–111). https://doi.org/10.1109/ICM.2018.8704097

Gorgel, P., Sertbaş, A., Kilic, N., & Osman, O. (2009). Mammographic mass classification using wavelet based support vector machine. *Istanbul University Journal of Electrical & Electronics Engineering, 9*(1), 867–875.

Hellgren, R., Dickman, P., Leifland, K., Saracco, A., Hall, P., & Celebioglu, F. (2017). Comparison of handheld ultrasound and automated breast ultrasound in women recalled after mammography screening. *Acta Radiologica, 58*(5), 515–520. https://doi.org/10.1177/0284185116664972

Hepsag, P., Ozel, S. A., & Yazici, A. (2017). Using deep learning for mammography classification. *2017 International Conference on Computer Science and Engineering (UBMK)* (pp. 418–423). https://doi.org/10.1109/UBMK.2017.8093429

Hiba, C., Hamid, Z., & Omar, A. (2016). An improved breast tissue density classification framework using bag of features model. *2016 4th IEEE International Colloquium on Information Science and Technology (CiSt)* (pp. 405–409). https://doi.org/10.1109/CIST.2016.7804986

Hoque, M. U., Sazzad, T. S., Farabi, A. A., Hosen, I., & Somi, M. (2019). An automated approach to detect breast cancer tissue using ultrasound images. *2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT)* (pp. 1–4). https://doi.org/10.1109/ICASERT.2019.8930491

Htay, T. T., & Maung, S. (2018). Early stage breast cancer detection system using GLCM feature extraction and k-nearest neighbor (k-NN) on mammography image. *2018 18th International Symposium on Communications and Information Technologies (ISCIT)* (pp. 171–175). https://doi.org/10.1109/ISCIT.2018.8587920

Ismail, N. S., & Sovuthy, C. (2019). Breast cancer detection based on deep learning technique. *2019 International UNIMAS STEM 12th Engineering Conference (EnCon)* (pp. 89–92). https://doi.org/10.1109/ENCON.2019.8920155

de Jesus Aragão, A., de Carvalho, D., Sanches, B., & Van Noije, W. A. (2020). An improved confocal algorithm for breast cancer detection using UWB signals. *2020 IEEE 11th Latin American Symposium on Circuits & Systems (LASCAS)* (pp. 1–4). https://doi.org/10.1109/LASCAS45839.2020.9022941

Johra, F. T., & Shuvo, M. M. H. (2016). Detection of breast cancer from histopathology image and classifying benign and malignant state using fuzzy logic. *2016 3rd International Conference on Electrical Engineering and Information Communication Technology (ICEEICT)* (pp. 1–5). https://doi.org/10.1109/ICEEICT.2016.7873617

Kavya, N., Usha, N., Sriraam, N., Sharath, D., & Ravi, P. (2018). Breast cancer detection using non-invasive imaging and cyber-physical system. *2018 3rd International Conference on Circuits, Control, Communication and Computing (I4C)* (pp. 1–4). IEEE. https://doi.org/10.1109/I4C.2018.8687627

Khasana, U., Sigit, R., & Yuniarti, H. (2020). Segmentation of breast using ultrasound image for detection of breast cancer. *2020 International Electronics Symposium (IES)* (pp. 584–587). IEEE. https://doi.org/10.1109/IES50839.2020.9231629

Jabeen, K., Khan, M. A., Alhaisoni, M., Tariq, U., Zhang, Y. D., Hamza, A., Mickus, A., & Damasevicius, R. (2022). Breast cancer classification from ultrasound images using probability-based optimal deep learning feature fusion. *Sensors, 22*(3), 807. https://doi.org/10.3390/s22030807

Kiymet, S., Aslankaya, M. Y., Taskiran, M., & Bolat, B. (2019). Breast cancer detection from thermography based on deep neural networks. *2019 Innovations in Intelligent Systems and Applications Conference (ASYU)* (pp. 1–5). IEEE. https://doi.org/10.1109/ASYU48272.2019.8946367

Lenzi, C., Pasian, M., Bozzi, M., Perregrini, L., & Caorsi, S. (2017). Mm-waves modulated Gaussian pulse radar breast cancer imaging approach based on artificial neural network: Preliminary assessment study. *2017 Mediterranean Microwave Symposium (MMS)* (pp. 1–4). https://doi.org/10.1109/MMS.2017.8497140

Liu, L., Li, J., & Wang, Y. (2015). Breast mass detection with kernelized supervised hashing. *2015 8th International Conference on Biomedical Engineering and Informatics (BMEI)* (pp. 79–84). IEEE. https://doi.org/10.1109/BMEI.2015.7401569

Lu, H. C., Loh, E. W., & Huang, S. C. (2019). The classification of mammogram using convolutional neural network with specific image preprocessing for breast cancer detection. *2019 2nd International Conference on Artificial Intelligence and Big Data (ICAIBD)* (pp. 9–12). IEEE. https://doi.org/10.1109/ICAIBD.2019.8837000

Ma, Y., Tay, P. C., Adams, R. D., & Zhang, J. Z. (2010). A novel shape feature to classify micro-calcifications. *2010 IEEE International Conference on Image Processing* (pp. 2265–2268). https://doi.org/10.1109/ICIP.2010.5651810

Mohebian, M. R., Marateb, H. R., Mansourian, M., Mañanas, M. M., & Mokarian, F. (2017). A hybrid computer-aided diagnosis system for prediction of breast cancer recurrence (HPBCR) using optimized ensemble learning. *Computational and Structural Biotechnology Journal, 15*, 75–85. https://doi.org/10.1016/j.csbj.2016.12.001

Narayanan, B. N., Krishnaraja, V., & Ali, R. (2019). Convolutional neural network for classification of histopathology images for breast cancer detection. *2019 IEEE National Aerospace and Electronics Conference (NAECON)* (pp. 291–295). IEEE. https://doi.org/10.1109/NAECON46414.2019.9057992

Nasiri, Y., Hariri, M., & Afzali, M. (2015). Breast cancer detection in mammograms using wavelet and contourlet transformations. *2015 2nd International Conference on Knowledge-Based Engineering and Innovation (KBEI)* (pp. 923–926). https://doi.org/10.1109/KBEI.2015.7436167

Nezhadian, F., & Rashidi, S. (2017). Breast cancer detection without removal of pectoral muscle by extraction turn counts feature. *2017 Artificial Intelligence and Signal*

Processing Conference (AISP) (pp. 6–10). https://doi.org/10.1109/AISP.2017.8324112

Paul, A., & Mukherjee, D. P. (2015). Mitosis detection for invasive breast cancer grading in histopathological images. *IEEE Transactions on Image Processing, 24*(11), 4041–4054. https://doi.org/10.1109/TIP.2015.2452392

Ponnusamy, J., & Babu, N. (2016). Joint probabilistic seed selection based fuzzy region growing algorithm to segment tumour on breast MRI. *International Journal of Printing, Packaging and Allied Sciences, 21*, 1428–1437.

Ponnusamy, J. A. R., & Babu, C. N. K. (2016). Breast lesion segmentation using generalised simulated annealing and neutrosophic region growing algorithm in breast MRI. *Academic Journal of Cancer Research, 9*, 75–81.

Prabhakar, T., & Poonguzhali, S. (2017). Automatic detection and classification of benign and malignant lesions in breast ultrasound images using texture, morphological, and fractal features. *2017 10th Biomedical Engineering International Conference (BMEiCON)* (pp. 1–5). IEEE. https://doi.org/10.1109/BMEiCON.2017.8229158

Ragab, D. A., Sharkas, M., Marshall, S., & Ren, J. (2019). Breast cancer detection using deep convolutional neural networks and support vector machines. *PeerJ, 7*, e6201. https://doi.org/10.7717/peerj.6201

Ren, J. (2012). ANN vs. SVM: Which one performs better in classification of MCCs in mammogram imaging? *Knowledge-Based Systems, 26*, 144–153. https://doi.org/10.1016/j.knosys.2011.08.012

Kumar, V., Chandrashekhara, K. T., Jagini, N. P., Rajkumar, K. V., Godi, R. K., & Tumuluru, P. (2025). Enhanced breast cancer detection and classification via CAMR-Gabor filters and LSTM: A deep learning-based method. *Egyptian Informatics Journal, 29*(1), 1–12. https://doi.org/10.1016/j.eij.2025.10.003

Labcharoenwongs, P., Vonganasup, S., Chunhapran, O., Noolek, D., & Yampaka, T. (2023). Asian Pacific Journal of Cancer Prevention, Vol. 24, pp. 1081–1088. *Asian Pacific Journal of Cancer Prevention, 24*(3), 1081–1088. https://doi.org/10.31557/APJCP.2023.24.3.1081

Alashban, Y. (2024). Breast cancer detection and classification with digital breast tomosynthesis: A two-stage deep learning approach. *Diagnostic and Interventional Radiology, 29*(2), 118–125. https://doi.org/10.4261/1305-3825.DIR.2023.23.51

Ronagh, M., & Eshghi, M. (2019). Hybrid genetic algorithm and particle swarm optimization-based microwave tomography for breast cancer detection. *2019 IEEE 9th Symposium on Computer Applications & Industrial Electronics (ISCAIE)* (pp. 244–248). IEEE. https://doi.org/10.1109/ISCAIE.2019.8710787

Ruhl, J., Callaghan, C., Hurlbut, A., Ries, L., Adamo, P., Dickie, L., & Schussler, N. (2021). Survival rates for breast cancer. *American Cancer Society.*

Samala, R. K., Chan, H.-P., Hadjiiski, L., Helvie, M. A., Richter, C. D., & Cha, K. H. (2019). Breast cancer diagnosis in digital breast tomosynthesis: Effects of training sample size on multi-stage transfer learning using deep neural nets. *IEEE Transactions on Medical Imaging, 38*(3), 686–696. https://doi.org/10.1109/TMI.2018.2870343

Alqthani, S. M. (2022). BreastCNN: A novel layer-based convolutional neural network for breast cancer diagnosis in DMR-thermogram images. *Applied Artificial Intelligence, 36*(1), 881–889. https://doi.org/10.1080/08839514.2022.2104523

Sangeetha, R., & Murthy, K. S. (2017). A novel approach for detection of breast cancer at an early stage by identification of breast asymmetry and microcalcification cancer cells using digital image processing techniques. *2017 2nd International Conference for Convergence in Technology (I2CT)* (pp. 593–596). https://doi.org/10.1109/I2CT.2017.8226198

Saubhagya, V. K., Rani, A., & Singh, V. (2016). ANN-based detection of breast cancer in mammograph images. *2016 IEEE 1st International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES)* (pp. 1–6). https://doi.org/10.1109/ICPEICES.2016.7853734

Shen, L., Rangayyan, R., & Desautels, J. (1994). Application of shape analysis to mammographic calcifications. *IEEE Transactions on Medical Imaging, 13*(2), 263–274. https://doi.org/10.1109/42.293919

Shirazi, F., & Rashedi, E. (2016). Detection of cancer tumours in mammography images using support vector machine and mixed gravitational search algorithm. *2016 1st Conference on Swarm Intelligence and Evolutionary Computation (CSIEC)* (pp. 98–101). https://doi.org/10.1109/CSIEC.2016.7482133

Shrivastava, N., & Bharti, J. (2020). Breast tumour detection and classification based on density. *Multimedia Tools and Applications, 79*(33–34), 26467–26487. https://doi.org/10.1007/s11042-020-09355-6

Shrivastava, N., & Bharti, J. (2022). Breast tumour detection in MR images based on density. *International Journal of Image and Graphics, 22*(2), 2250001–2250021. https://doi.org/10.1142/S021946782250001X

Singh, S., & Kumar, R. (2020). Histopathological image analysis for breast cancer detection using cubic SVM. *2020 7th International Conference on Signal Processing and Integrated Networks (SPIN)* (pp. 498–503). IEEE. https://doi.org/10.1109/SPIN48936.2020.9071239

Singh, V. K., Rashwan, H. A., Romani, S., Akram, F., Pandey, N., Sarker, M. M. K., Saleh, A., Arenas, M., Arquez, M., Puig, D., *et al.* (2020). Breast tumour segmentation and shape classification in mammograms using generative adversarial and convolutional neural networks. *Expert Systems with Applications, 139*, 112855. https://doi.org/10.1016/j.eswa.2019.112855

Singh, V. P., Srivastava, A., Kulshreshtha, D., Chaudhary, A., & Srivastava, R. (2016). Mammogram classification using selected GLCM features and random forest classifier. *International Journal of Computer Science and Information Security, 14*(6), 82–88.

Singh, V. P., Srivastava, S., & Srivastava, R. (2018). Automated and effective content-based image retrieval for digital mammography. *Journal of X-ray Science and Technology, 26*(1), 29–49. https://doi.org/10.3233/XST-17297

Soliman, O. O., Sweilam, N. H., & Shawky, D. M. (2018). Automatic breast cancer detection using digital thermal images. *2018 9th Cairo International Biomedical Engineering Conference (CIBEC)* (pp. 110–113). IEEE. https://doi.org/10.1109/CIBEC.2018.8631165

Stojadinovic, A. (2006). *Electrical impedance scanning for the early detection of breast cancer in young women.* Technical report. Henry M. Jackson Foundation for the Advancement of Military Medicine, Rockville, MD.

Swetha, T., & Bindu, C. (2015). Detection of breast cancer with hybrid image segmentation and Otsu's thresholding. *2015 International Conference on Computing and Network Communications (CoCoNet)* (pp. 565–570). https://doi.org/10.1109/CoCoNet.2015.7411244

Tasoulis, M. K., Zacharioudakis, K. E., Dimopoulos, N. G., & Hadjiminas, D. J. (2014). Diagnostic accuracy of tactile imaging in selecting patients with palpable breast abnormalities: A prospective comparative study. *Breast Cancer Research and Treatment, 147*(3), 589–598. https://doi.org/10.1007/s10549-014-3142-3

Reshma, V. K., Arya, N., Ahmad, S. S., Wattar, I., Mekala, S., Joshi, S., & Krah, D. (2022). Detection of breast cancer using histopathological image classification dataset with deep learning techniques. In *Cloud Analytics-Based Applications for Bioinformatics 2021* (pp. 1–13). https://doi.org/10.1002/9781119799201.ch1

Varma, C., & Sawant, O. (2018). An alternative approach to detect breast cancer using digital image processing techniques. *2018 International Conference on Communication and Signal Processing (ICCSP)* (pp. 134–137). https://doi.org/10.1109/ICCSP.2018.8524576

Vikhe, P., & Thool, V. (2016). Mass detection in mammographic images using wavelet processing and adaptive threshold technique. *Journal of Medical Systems, 40*(4), 82. https://doi.org/10.1007/s10916-016-0442-x

Wang, H., Feng, J., Bu, Q., Liu, F., Zhang, M., Ren, Y., & Lv, Y. (2018). Breast mass detection in digital mammogram based on Gestalt psychology. *Journal of Healthcare Engineering, 2018*, Article 3953702. https://doi.org/10.1155/2018/3953702

Wang, T., Wang, K., Yao, Q., Chen, J. H., Ling, R., Zhang, J. L., Dong, X. Z., Fu, F., Dou, K. F., & Wang, L. (2010). Prospective study on combination of electrical impedance scanning and ultrasound in estimating risk of development of breast cancer in young women. *Cancer Investigation, 28*(3), 295–303. https://doi.org/10.3109/07357900903264645

Wang, Y., Wang, N., Xu, M., Yu, J., Qin, C., Luo, X., Yang, X., Wang, T., Li, A., & Ni, D. (2019). Deeply supervised networks with threshold loss for cancer detection in automated breast ultrasound. *IEEE Transactions on Medical Imaging, 39*(3), 866–876. https://doi.org/10.1109/TMI.2019.2936745

Wang, Y., Zhang, L., Shu, X., Feng, Y., Zhang, Y., & Lv, Q. (2021). Feature-sensitive deep convolutional neural network for multi-instance breast cancer detection. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 1–1. https://doi.org/10.1109/TCBB.2021.3060183

Wang, Y., Zhang, L., Shu, X., Feng, Y., Zhang, Y., & Lv, Q. (2021). Feature-sensitive deep convolutional neural network for multi-instance breast cancer detection. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*. https://doi.org/10.1109/TCBB.2021.3060183

Wei, K., Wang, B., & Saniie, J. (2020). Faster region convolutional neural networks applied to ultrasonic images for breast lesion detection and classification. In *Proceedings of the 2020 IEEE International Conference on Electro Information Technology (EIT)* (pp. 171–174). IEEE. https://doi.org/10.1109/EIT49495.2020.9206775

Yousefi Kamal, P. (2019). Breast tumour classification and segmentation using convolutional neural networks. *arXiv preprint arXiv:1905.04247*. https://doi.org/10.48550/arXiv.1905.04247

Zhang, Y., Chan, S., Park, V. Y., Chang, K. T., Mehta, S., Kim, M. J., Combs, F. J., Chang, P., Chow, D., & Parajuli, R. (2020). Automatic detection and segmentation of breast cancer on MRI using mask R-CNN trained on non–fat-sat images and tested on fat-sat images. *Academic Radiology, 27*(11), 1546–1555. https://doi.org/10.1016/j.acra.2020.01.006,