

Survey on various Feature Extraction Methods for Breast Cancer Detection using Soft Computing Approach

¹P.EstherJebarani* ²Dr.N.Umadevi

¹Assistant Professor, Department of Computer Science, KovaiKalaimagal College of Arts and Science, Coimbatore.

²Associate Professor, Department of Computer Science, Sri JayendraSaraswathyMahaVidyalaya College of Arts and Science, Coimbatore.

*estershine.7401@gmail.com, umadevigayathri@gmail.com

ABSTRACT

Breast cancer is one of the foremost relentless malignant growth found in women and it has also become a major reason of death ratio. However, the recognition of breast cancer at the premature stage is reliant on both the ability of the radiologist's and the superiority of the images. Treatment for the breast cancer is more successful only when the detection is done at the premature stage. Screening of breast cancer is done in two conditions, first one is when diagnosed at an early stage and the second one is at the later stage with some symptoms and in each stage requires a different type of treatments. Although there are different types of techniques, among which mammography is frequently used technique for discovery of tumors in breast. Mammography is the most effective method to trace the abnormal cancer cells. Screening is one of the key factors to diminish the death rates. It is achieved through the classification algorithm which in turn identifies the severity of lymph's in the breast. During the process of classification, the similar patterns are recognized and extraction of features from the region of the image. Digital mammography is the frequently handled technique for premature detection but these images are very complex to reduce false positives. Feature performs the considerable function in the area of digital image processing. There are enormous image preprocessing techniques adopt various progressions like resizing, thresholding, binarization, normalization etc which are applied on the retrieved images. These techniques are applied to massive image processing applications like character recognition. The feature rules the activities of the image, competence in classification, the storage place and also the time utilization. The different kinds of feature extraction methods based on image processing were studied. It is also suggested for the best features extraction technique that would be applicable for application development.

Keywords: Breast Cancer, Computer-Aided Diagnosis, Local Binary Pattern, Gray Level Co-occurrence Matrix, Fuzzy local Binary Pattern

Introduction

One of the major aggressive diseases which are commonly identified one in eight among women in their lifetime [1]. There are different ways that people can be screened for breast cancer, they are Ultrasound, Mammography, MRI, Thermography and the emerging technologies

such as molecular breast imaging and Digital Breast Tomosynthesis (DBT) [2]. When comparing all the diagnostic methodologies which are available for detecting breast cancer, one of the dependable and sensible techniques is the Mammography, which is to examine cancer at the early stage [59].

The World Health Organization (WHO) for the year 2012 has estimated those 70218 deaths in India [3] and 521,907 deaths worldwide among woman is due to breast cancer [4]. According to the study in 2016, 61,000 new cases are identified as the patients affected with breast cancer [5]. And also in the year 2018, it is noted that 6,27,000 women died due to breast cancer. Breast cancer rates are increasing and higher among women in every region globally. These investigations have indicated that detecting cancer at the early stage and proper treatment at the stipulated time prompts an expansion in the endurance proportion of the patients. Thereby improving the sensitive capturing of the mammographic images can achieve the major goal of the examination while also eradicating the number of needless procedures or surgical operations.

Early detection is critical to progress breast cancer outcomes and survival. There are two main strategies for breast cancer during the case of early detection: early diagnosis and screening [54]. Early diagnosis deals with three integrated steps and must be provided in a timely manner. The following steps are awareness, clinical evaluation, and treatment. Screening is performed to identify the abnormalities of a specific type of cancer or pre-cancer that has not developed with any other symptoms and referring to the subsequent check for diagnosis and treatment.

Mammographic screening programs which are led in all the nations are a successful technique to identify the disease at the underlying stage. There are two categories of mammography available namely, digital and film. When evaluating both the mammography, digital is better than film hence the radiation emission is diminished up to 50 % and still, it can sense the cancer cells [6]. Screening is difficult only when the interpretation of a mammogram image is not properly handled by the radiologist. The radiological interpretation is difficult in the case where the appearance of the normal tissue is variable, complex and the manifestation of the cancer cells is really very small or imprecise [7]. Radiologists possibly get benefited from mammography which in turn builds the strong association between abnormal and normal breast cancer cells. In order to progress in this process, mammography is adopted to trigger out masses with clarity nearly two years before the manifestation of cancer cells [8] [38].

Most of the dataset collected from the database is noisy, partial, and inconsistent, so preprocessing becomes mandatory. Image preprocessing is the initial step which is extremely requisite to ensure greater accuracy of the succeeding steps [55]. A CT and MRI image usually includes some of the artifacts for instance patient-specific, equipment based artifacts, and so on. Patient-specific artifacts consist of metal and motion beam hardening. Remaining consists of volume effect, ring, and staircase. Before starting for analysis the artifacts must be removed by preprocessing. Image enhancement processing is also used to eradicate film, label, and filtering the images [9].

In imaging stipulations, the feature is depicted as a spatial arrangement and dissimilarity in intensities within images [45]. The region of interest within the particular image is identified with the help of the feature extraction. It is a significant and crucial step in image processing as it marks the transition points from the pictorial to non-pictorial representation [56]. The substantial representation can be used as an input for further processing like classification [60], pattern recognition, and so on. This leads to classifying the labels or the contents of the objects. A feature assumes an essential job for distinguishing micro-calcification. Image procurement is done in two distinct locales. The principal area is the center region of the bosom where the profundity is in a uniform way and generally alluded as consistent thickness locale. The subsequent locale is close to the edge of the bosom where the thickness goes to a point because of the bosom geometry. To obtain better results, the texture must be demoralized from the divided image. The region of interest (ROI) of the image is determined by utilizing a network with a pair of pixels. Furthermore, ROI is utilized to distinguish favorable and harmful cells. The above-said texture obtains a distinctive component with various qualities [10].

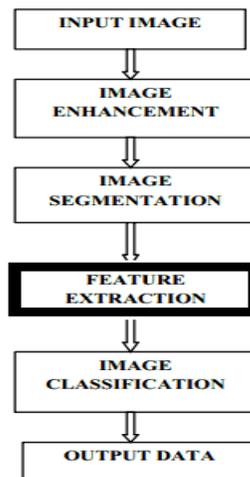


Fig.1 Structure of Image Processing

Fig.1 shows the structure of Image Processing. In this diagram original image is considered as an input. The first step is the image enhancement is towards the enhancement of the image by improving the quality like brightness, contrast etc of the original data. The second step segmentation is to acquire the object that is the Region of Interest to be identified individually. The third step feature extraction is to reduce the image dimension which paves the way to completely depict the original data set. The fourth step classification is to portray the quantitative spectral information contained in an image. Finally, the desired output data is obtained after several processing of image data. Here, the main focus is on the feature extraction which exhibits a foremost task in this current manuscript.

Methodology

The principle idea is to extricate the impossible to highlight from the first image to identify the harmful cells and furthermore to perceive the specific region of interest (ROI) from the influenced cells.

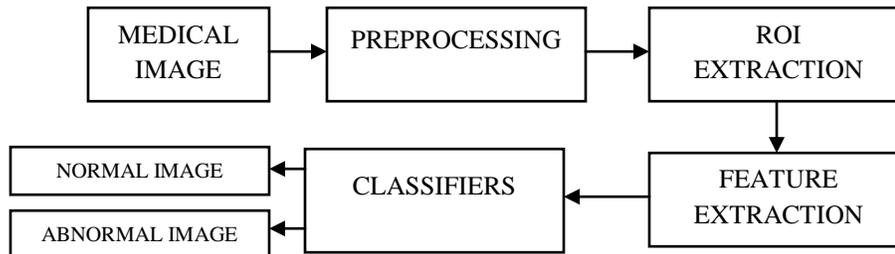


Fig.2 Flow Diagram of medical image processing

The above Fig.2 describes the progression of clinical image processing. The development of the image illustrated fewer than five basic steps. The first step is to collect the image from the database available. The second step is to apply the contrast adjustment and histogram equalization which is the preprocessing of the image. The third step is to find out the particular ROI. The fourth step is to apply the different patterns for the feature extraction. The final step is to classify the texture from the image to find out the existence of normal and abnormal images.

Related works

The preliminary step is to obtain the object of interest to locate the extraordinary point of small masses. In the subsequent step matches the micro-calcifications seen in the mammograms regarding the different shapes, unique sizes and variables [58]. The suspicious ROI is refined and also the surrounding thin tissues with low contrast were identified. When operated in thick tissues cause uncertain territory to be virtually undetectable in young women [61]. Finally, dense tissues sometimes are confused as calcifications which lead to false-positive cases. The identification of malignancy is prepared in different manners as per the component extraction draws near namely, SGLDM, CAD, SFUM, CADx, Gray-Map, MRELBP, Sobel Filter, LFLBP, GLCM, LBPV, CADe, ELBP, AFUM, LBP, and CLBP.

Gray-Map

Gray-map functions over the pixels are very simple to coordinate with feature vectors [62]. The operation is carried out using Principal Component Analysis (PCA) which dominates all the pixel neighborhood of the image and gathers as prototypes. The vectors are predicted as lower level dimensions through PCA. The calculated results are compared with the similar complicated feature extraction methods. This method is evaluated using $n2$ -dimensional vectors formalized in connection with gray-level values [11] [12].

Sobel Filter

The Sobel Filter is the procedure utilized for recognizing the edges by computing the angle of image force in each pixel contained in the image. To find the path of the lump the operation is carried out by adjusting the contrast from the light to dark. And the vice versa dark to light is evaluated across the edges. The image transforms at every pixel smoothly and therefore the pixel symbolizes an edge. The orientation of every pixel is traced by the edges and the filter is applied that results in the region of constant intensity is a zero vector [13].

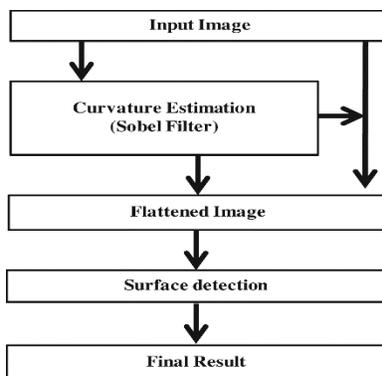


Fig 3 Processing of Sobel Filter

In Fig.3 the sobel channel handling is done by 3 x 3 parts, the earlier one is for the adjustments in the horizontal and the later one for the adjustments in the vertical directions. The two portions are convolved with the representation to operate the approximations of the derivatives. This technique attains little computation time, robustly susceptible to noise and can create disconnected contours.

Spatial Gray Level Dependence Matrices

Flourishing contact to tissue division in clinical images are upheld surface descriptors like Fractal features, Spatial Gray Level Dependence Matrices (SGLDM, additionally alluded to as Co-occurrence Matrices) and different types of textural features [14]. SGLDM may be a statistical procedure which stays in building co-event lattices to mirror the spatial dispersion of dark levels within the region of intrigue. SGLDM is predicated on the induction of the subsequent request with the unexpected likelihood thickness is $g(i,j,d,\Theta)$.

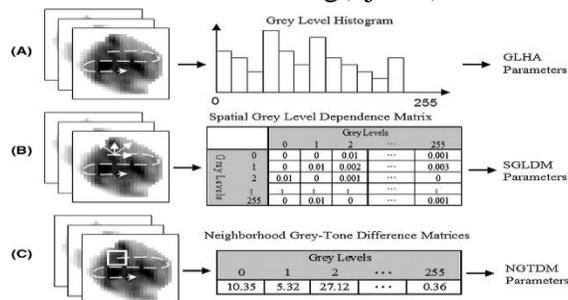


Fig.4 Matrix of SGLDM

In the above stated Fig.4 the SGLDM framework indicates that the component at area (i, j) is likelihood that of two distinctive goals cells which are during a specified direction Θ from the flat and determined separation d from one another, will have dark level qualities i and j separately. The angle is employed to gauge the bearing of texture, and in this manner the application of a few separation esteems can give a significant portrayal of the dimensions of the periodicity surface. Thus for various Θ and d esteems, diverse SGLD frameworks result. The point Θ is typically restricted to estimations of $0, 45, 90,$ and 135° , and thus the distance d is restricted to values limited to basic products of pixel size.

Average Fraction Under the Minimum

In breast cancer detection this methodology is for a mass discernment calculation called Average Fraction Under the Minimum (AFUM). This algorithm rather than generating highlight vectors, gives one characteristic for every pixel, which is utilized straightforwardly as a certainty incentive. This method emphasis on vector and the dimensionality of the factors is more significant. The measurement of the element vectors is decreased by utilizing Principal Component Analysis (PCA) [15]. In Fig.5 considering the pixel p_{ij} , the lower intensity value is received while connecting the distance between r_2 and p_{ij} by the pixel fractions. This process is described for the radius r_1 and focus (i, j) . This computation is completed to a variety of r_1 , and r_2 . Therefore, the standard among the values is registered. The major benefit of this technique is invariant to rotations and there is no need for parameter training process.

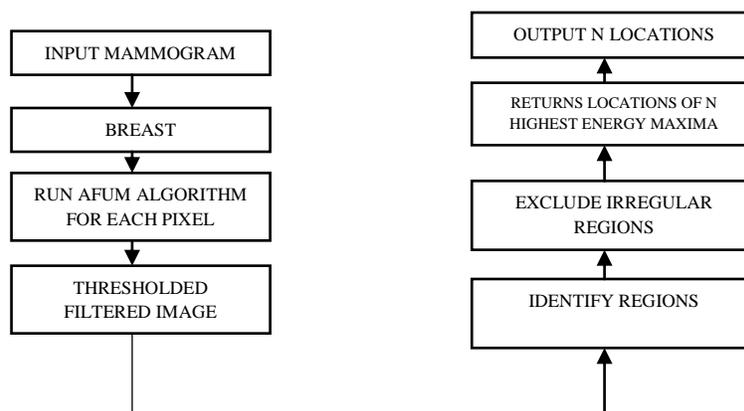


Fig.5 Different levels of AFUM image processing

The initial step is to apply the algorithm for each pixel. The next step is to identify the threshold filtered image which leads to utilize the regions. By acquiring the region of interest, it is extremely simple to categorize the range of the mass detection along with the optimized energy.

Set of Fractions Under the Minimum

In the above-mentioned strategy AFUM, rather than figuring the entire values only the set of values are considered which is termed as Set of Fractions Under the Minimum (SFUM). Here, a component vector is considered rather than one scalar. On the off chance that the dimensionality of this vector is taken into account are excessively high, it is often reduced by methods for Principal Component Analysis (PCA) [16]. Despite the fact that it is often a boon to possess one feature for every pixel expressed in AFUM calculation and a loss of data occurs when the typical group of highlights is registered. To preserve a tactical distance, a variation of the SFUM calculation is proposed.

Gray level Co-occurrence Matrix

The Gray Level Co-occurrence Matrix (GLCM) might be an excellent system for the extraction, which is practical in image retrieval, image recognition, texture analysis, image segmentation, and image classification methods. To deal with the feature extraction a heading measure that supports the directionality of texture is made available for extraction. The resultant is to measure and to have combination using the GLCM. For the most part surface estimations are weighted midpoints of the standardized GLCM [17] cell substance. A weighted normal increases each an incentive to be employed by an element before adding and separated by the number of values. The results of a surface calculation may be a single number representing the whole window. This number can be placed in the middle pixel of the window, at that point the window is moved one pixel and consequently the process is rehashed for calculating a replacement GLCM and a replacement texture measure. The above advances are shown in Fig.6.

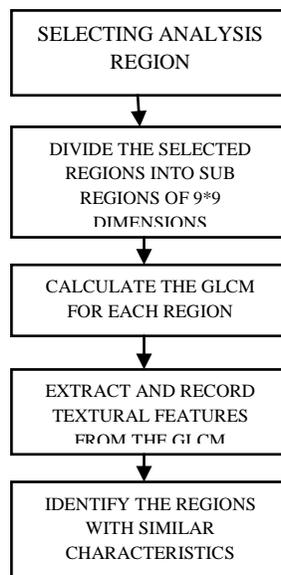


Fig.6 Identifying ROI using GLCM

The entire image is developed using the texture values. The utilization of GLCM to remove the element through the joint condition likelihood appropriation of the picture melancholy level to complete surface and processes the local association of pixels to obtain the component assessment. GLCM is widely used in various fields and has been ceaselessly redesigned. It also provides a superior level when contrasted with the intensity histogram and intensity feature.

Computer-Aided Diagnosis

There is a significant requirement for Computer-Aided Design (CAD) structure is to aid radiologists to recognize and analyze new cases. An entire breast image is equipped for liberating limits of breast and the pectoral muscle. This section is separated into various parts like fatty locale, thick region, pectoral region, and tape foundation. CAD system identifies the initial stage of breast tumor with the assistance of image processing and AI gives 85% -93.1% accurate results. The upgraded CAD system identifies the initial stage tumor with 99.9% accuracy for carcinoma [40]. CAD system combines the segmentation algorithm to urge the highest success rate. The association among the carcinoma and abnormalities depicts that radiologists could obtain assistance from the CAD system [18] to computerize the breast tissue classification. The framework of the CAD scheme is progressed to support radiologists to recognizable the proof of benevolent and dangerous masses. It also serves as a tool for clinician to diagnose and generate computer related outputs. This technique is distinguished from further automated diagnosis techniques in which the investigation is based on the computer algorithms. It is frequently used with two other methods known as Computer-Aided Detection (CADe) and Computer-Aided Diagnosis (CADx). CADx supports the practitioner to analyze and evaluate pathology in medical images [41]. CADe detects the particular area of images that may appear unusual. Fig.7 depicts the four foremost stages of CAD diagnosis namely segmentation, classification, preprocessing, and feature extraction. The image processing improves the quality of an image by eradicating the inappropriate image data from image in numerous applications and domains. In medical field the scanned image consists of a set of irrelevant and unnecessary parts. The first stage is to eliminate such kind of parts; it is necessary to preprocess the image for enhanced version of images before obtaining the diseases in critical [50].

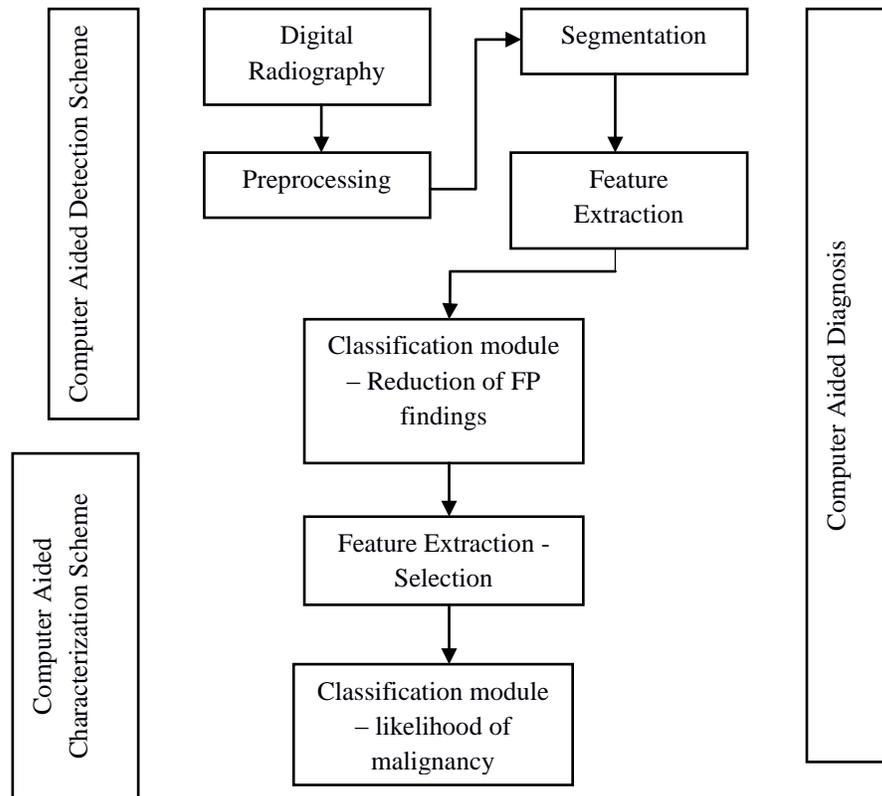


Fig.7 CAD Scheme for the breast cancer detection

The second stage is the segmentation which aid to highlight important areas and obtain different structures so as organs or tumors for more analysis. One of the common and prime algorithms used for CAD system is the region-based algorithm [44]. The feature descriptors of an image are to decrease the dimensions of the original image which further implies feature extraction. Features are properties of the complete image or choose region of interest. Features are extracted depending on the impacts such as computational cost of classification, the accuracy of classification, memory size, and robustness. Depending on the features selected the suspected areas are classified under malignant or benign. The fourth stage is classification, where the strategy is classified into two combinations namely, unsupervised and supervised. In the field of image processing commonly the supervised classification techniques are utilized [46]. CAD framework can build a lively area of analysis and progression in radio analysis.

Computer-Aided Detection and Computer-Aided Diagnosis

To develop the conviction and competence of mammogram assessment, CAD has been initiated within the screening procedure to help radiologists during the investigation [49]. Generally, CAD frameworks do strengthen the translation of clinical images, and two fundamental designs are often determined as CADx and CADe. The method CADe is concentrated lying on the situation of suspect areas while the CADx [19] is focused on characterization [46]. During this investigation, a CADe consolidates the automated

mammogram surface division for the distinguishing proof of starting time tumors [39]. The calculation is experienced upon the various images from the advanced information base, which are used for screening mammography by examination and finding of the sickness [47].

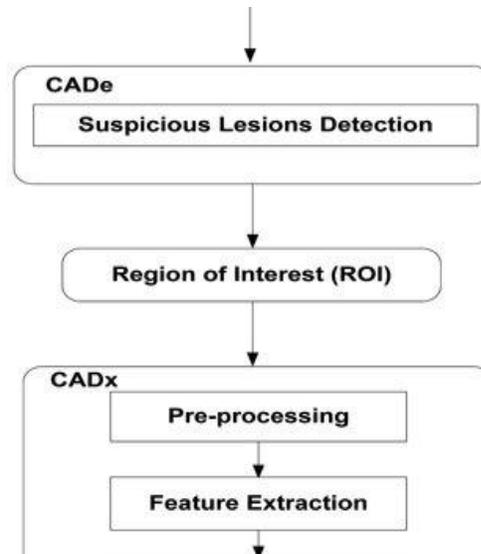


Fig.8 Operation with CADe and CADx system

In the above stated Fig.8 the most function of a carcinoma CADe conspires [20] to isolate the regions which surround masses as the foundation for characteristic tissue of an organ [48]. Such regions are partition into a few non-meeting regions then extract ROI where suspicious masses probably found to be in ultrasound image [42]. A CADx framework is to forecast the abnormal area inside mammograms. Finally, this method suggests to focus on the interior representation for preprocessing the image and also to deal with the extraction of the particular objects supported by texture segmentation [43].

Gray Level Run Length Matrix

The Gray Level Run Length Matrix (GLRLM) is where the surface feature can be construed to include assessment. The GLRLM is predicated on manipulating the number of the gray level depending on the length of the runs. A gray level run might be a lot of progressive and successive pixels to compare the gray level evaluation [36]. GLRLM are proficient and attained to employ the classification precision rate in correlation to different pixel points. GLRLM [21] is that the group of constant pixels contains similar dimension levels. The duration of the run is obtained from the neighboring gray level values. The number of times is calculated along with counting of runs within the image. In the equation.1 GLRLM lattice the pixel $P(i, j|\theta)$, this segment deciphers the measure of runs alongside the dark level i and the length j happen inside the image irrelevant θ . The gray level is calculated by using the run length that are often defined as,

$$N_g \sum_{i=1}^{N_r} \sum_{j=1}^{N_p} P(i,j|\theta) \text{ and } 1 \leq N_z(\theta) \leq N_p \text{ ----- (1)}$$

Where, N_g is the quantity of force esteems

N_r is the quantity of length of runs

N_p is the quantity of pixels inside the image

$N_z(\theta)$ is the quantity of the runs along with image point

$P(i,j|\theta)$ is the length of the run grid for a discretionary bearing θ

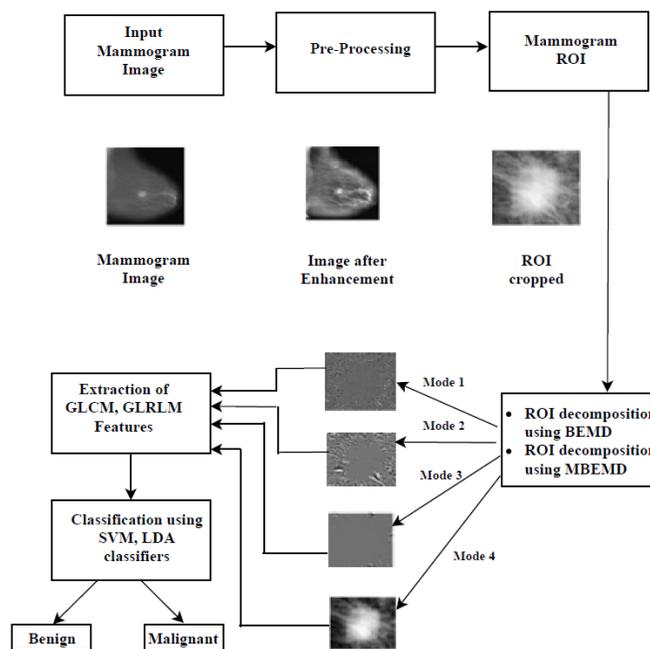


Fig.9 Block diagram for GLRLM method

The square graph of the GLRLM strategy is represented in Fig.9. The proposed strategy includes extraction technique comprises of the subsequent steps: (i) The initial step is to pre-process the specified region. (ii) The next step is to find the ROI to trim the mass area. (iii) To decompose of ROI the supported BEMD and the MBEMD are split into many BIMFs or several modes. (iv) Using the methods called GLCM and GLRLM is to extract the BIMFs. (v) Finally the ROI are classified using the SVM, and LDA classifier. For arrangement of Region of Interest (ROI), SVM and LDA classifiers are employed. The result is obtained by BEMD and MBEMD. Using the classifier, it is very simple to segregate whether it is benign or malignant.

Local Binary Pattern

Local Binary Pattern (LBP) is an efficient procedure for imparting highlights of the pixel by moving an area operator on every pixel area of the gray scale. This reflects the spatial structure of an image by predetermining the uniqueness among pixel estimation of the focal point and its neighbors. The subsequent decimal estimation of the created parallel is to label the specified pixel [51]. This method has been productively applied to a few differing issues including dynamic feature acknowledgment, remote detecting, unique finger impression coordinating, visual review, image recovery, biomedical image investigation, face image investigation, movement investigation, edge discovery, and condition modeling. An outsized number of LBP [22] variants are developed to upgrade its vigor, discriminative force, and materialness. It has appeared to be the foremost prominent surface descriptors, pulling in critical attention within the field of the analysis [52]. The remarkable point of interest is to have simple implementation over the invariance to monotonic light changes, and to lower computational multifaceted nature. In the equation.2 LBP histogram is designed with the assistance of (x, y) pixels by holding the size of $M \times N$.

$$LBP_{P,R}(x,y) = \sum_{b=0}^{P-1} s(I_n - I_c)2^b, \quad \text{----- (2)}$$

$$s(z) = \begin{cases} 1, & z \geq 0 \\ 0, & z < 0 \end{cases}, \quad \text{----- (3)}$$

The power estimation of every neighboring pixel based on I_c showing the pixel estimation of the middle pixel at the point (x, y); P alludes to the quantity of the adjacent pixels and R signifiesthe sweep of region.

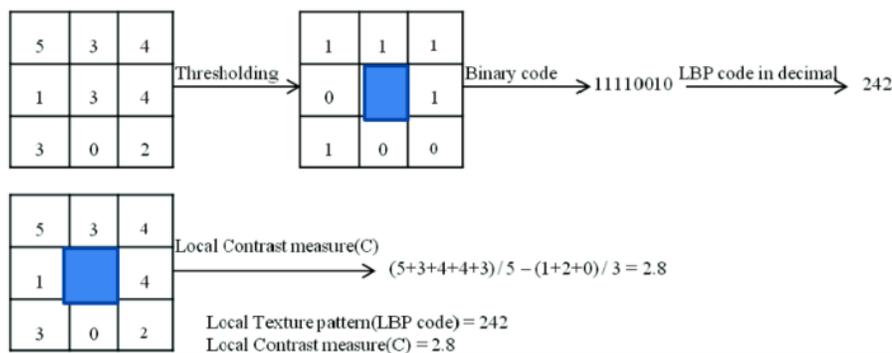


Fig.10 Working of Local Binary Pattern

Fig.10 showcases the LBP calculation over the pixels to obtain the determined pattern. The initial process is to attain the thresholding value of the pixels in the matrix. With the help of the

base thresholding obtain the binary code. Further the binary codes are converted into decimal values to broadcast the final LBP code.

Local Binary Pattern Variance

Normally LBP histogram contains local spatial information but to order to manipulate with the features there must be more protection to deal with the rotation variations and contradiction [53]. To execute the functionality further some of the instructions has been included by characterizing a combined histogram of LBP and turn in-variant changes measure the value as (VAR) LBPV, R/VAR_{P,R} [23]. Be that as it may, to defeat the disadvantages of LBP the elective technique is known as LBPV. In LBP, the weights of the pixels are calculated using the LBP codes and to attain LBPV histogram it is scientifically communicated as:

$$LBPV_{P,R}(l) = \sum_{x=1}^M \sum_{y=1}^N w(LBP_{P,R}(x,y),l), l \in [0,L] \quad \text{-----(4)}$$

$$w(LBP_{P,R}(x,y),l) = \begin{cases} VAR_{P,R}(x,y), & LBP_{P,R}(x,y) = l \\ 0, & otherwise \end{cases}, \quad \text{-----(5)}$$

$$VAR_{P,R} = \frac{1}{P} \sum_{b=0}^{P-1} (I_n - I_{av})^2, \quad \text{-----(6)}$$

$$I_{av} = \frac{1}{P} \sum_{b=0}^{P-1} I_n, \quad \text{-----(7)}$$

Where, L expresses the highest value obtained for LBP codes.

Completed Local Binary Pattern

Completed Local Binary Pattern (CLBP) is like the LBP strategy and developed with the sign variety. The size variety of the middle pixel and the following neighboring pixels are inspected to find the nearby administrator. Additionally, the conversion from the encoded code into a code of the middle pixel is executed using thresholding.

Finally, CLBP histogram [24] is encircled by connecting the three dissimilar pixel values. In the expressed strategy, the differentiation between intensity value of any nearest pixel (I_n) and pixel value of the center pixel (I_c) is experimented to recognize the local area to accomplish a distinction vector d_b which is additionally partitioned into the two segments.

$$d_b = S_b * M_b \quad \text{and} \quad \begin{cases} S_b = \text{sign}(d_b) \\ M_b = |d_b| \end{cases} \quad \text{-----(8)}$$

Where S_b point towards the sign of d_b known by $S_b = 1$ if $d_b \geq 0$, otherwise 0 and M_b symbolizes magnitude difference.

Median Robust Extended Local Binary Pattern

Median Robust Extended Local Binary Pattern (MRELBP) is examining among the primary computationally proficient superior texture appearance. Be that as it may, the LBP form is extremely receptive to noise of the image and powerless to obtain macrostructure information. To address these issues of LBP, a new methodology presented is totally exclusive descriptor for texture analysis known as the MRELBP [25]. The normal LBP and the LBP variants are evaluated by the MRELBP to obtain the local images to the pixel rather than unprocessed image sharpness. A multi-scale LBP type descriptor is processed effectively at image midpoints covering a unique sampling system that captures both macrostructure and microstructure information. A complete assessment on the data set discloses that the MRELBP works on elevated appearance to deal with the gray scale varieties, pivot changes and noise. MRELBP produces the highest accuracy with 99.82%, 99.38%, and 99.77%.

Extended Local Binary Pattern

LBP method embraces the central and neighboring pixels whereas Extended Local Binary Pattern (ELBP) is designed to establish the connection between the local regions and also to attain the spatial information. ELBP [26] consist of three methods such as LBP-like descriptors Radial Difference-based LBP (ELBP_RD), Neighborhood Intensity-based LBP (ELBP_NI), and Center Intensity-based LBP (ELBP_CI), which investigate information. The ELBP approach is parallel to the unique LBP.

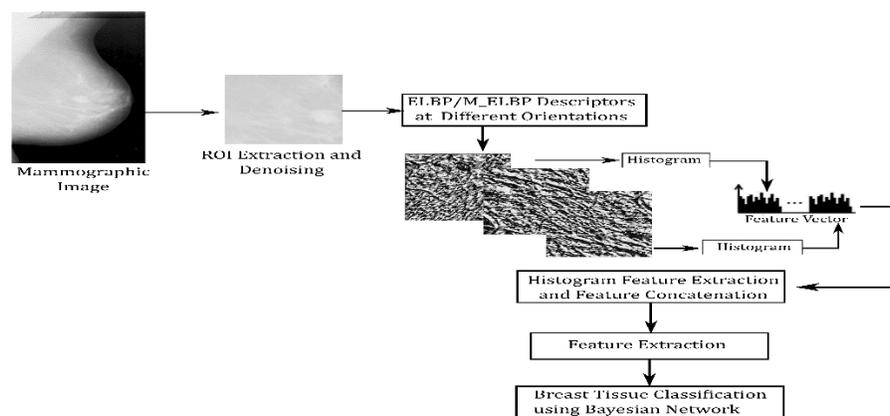


Fig.11 Analysis of ROI selection, article eradication and allocation using ELBP variants

The above-mentioned Fig.11 depicts the arbitrary region of interest selection and the character eradication is evaluated using ELBP variants. ELBP/M-ELBP orientations are evaluated by the combinations of feature concatenated and feature extraction. The histogram labels for the ELBP/M-ELBP caption in the least directions are incorporated to acquire

the feature vector for the ROI. ELBP formed good texture analysis conduct; however, there remain some symbolic disadvantages like affectability to image obscure and commotion, neglecting to grasp texture macrostructure and high detail dimensionality.

Intuitionistic Fuzzy Local Binary Pattern

Intuitionistic Fuzzy Local Binary Pattern (IFLBP) technique is to extricate the texture section from the input image. This methodology proposes that the Fuzzy Local Binary Pattern (FLBP) procedure includes intuitionistic fuzzy inside the statement of neighborhood examples of surface specifically images. This technique is measured as the support for reviewing the IFLBP. FLBP for texture revelation is strong and vulnerable to noise. Most of the applications to be originate using fuzzifications such as LBP, BP capability, LBP/C, MBP, and LEP. The texture expressions are considered for an extensive and competent investigation based on a typical collection of textures. The entire IFLBP histogram is determined by expanding the IFLBP codes as

$$H_{IFLBP}(i) = \sum_{x,y} C_{IFLBP}(x,y,i), i=0,1,\dots,2^k-1 \quad \text{-----(9)}$$

x,y

IFLBP [27] histogram acquires data than firm LBP and FLBP histograms. IFLBP include extraction structure to encode nearby quality by checking intuitionistic fuzzy pure mathematics within the depiction of neighborhood designs. This system is to extend the methodology of FLBP. It has supplementary advantage to contribute the allocation of the IFLBP values. The feature vectors are frequently operated in the diverse field of the image processing that includes image recognition, classification of pattern, image de-noising problems and so on.

Table.1 Comparison of different Feature Extraction Techniques / Methods

Author and Year	Feature Extraction Technique s / Methods	Advantages / Characteristics	Limitations
Guangxing Guo & Navid Razmjoo (2019)	Sobel Filter [28]	<ul style="list-style-type: none"> • Short computation time • Harmony to distinguish edges three pixels extensively. 	<ul style="list-style-type: none"> • Exposed to noise. • Use to provide irregular contours.
Michael D. Heath and Kevin W. Bowyer. (2000)	AFUM [29]	<ul style="list-style-type: none"> • Constant to rotations 	<ul style="list-style-type: none"> • Information loss happens

		<ul style="list-style-type: none"> • No need for parameter training or learning process. • Single feature for each pixel 	<p>during the time of calculating the set of features for the average value.</p>
Xin Zhang, Jintian Cui et.al (2017)	GLCM [30]	<ul style="list-style-type: none"> • Primitive or low level image features can be extracted • Motion estimation of images. • Better performance when compared with intensity histogram and intensity features 	<ul style="list-style-type: none"> • GLCM matrix contains tiny undetectable pixel components that can't be extricated from the mammogram image especially reflecting on the ROI. • Small inconsistency in direction • Clustering effect is affected or (robustness of the texture direction change)
Sumit Kumar Sourav, Rajeev Ranjan, et.al(2018)	CAD [31]	<ul style="list-style-type: none"> • Identifies the early stage tumor with accuracy • Highest success rate. • The powerful relationship between breast cancer and abnormalities. • Radiologists perceive advantage from the CAD (Computer-aided 	<ul style="list-style-type: none"> • Features separated legitimately from ROIs may not give strong and precise execution.

		<p>Diagnosis) system with capabilities of automatic breast tissue classification.</p> <ul style="list-style-type: none"> • Advanced to support radiologists for the identification of benign and malignant masses. 	
Jinshan Tang, Rangaraj M Rangayyan et.al (2009)	CADe [32]	<ul style="list-style-type: none"> • The interior replica is the wrenching of the objects of attention from the structure dependent on surface selection. • Helps in discovering and recognizing feasible abnormalities found in the image 	<ul style="list-style-type: none"> • The tumor's characteristics are not provided to the radiological by the CADe systems.
Monica EzzatGamil, Mariam Mohamed Fouad, et.al (2018)	CADx [33]	<ul style="list-style-type: none"> • Targeted to characterization (i.e., malignancy versus benignity). 	<ul style="list-style-type: none"> • This method includes constraints to identify nodules and an elevated automation level.
Hui Zhou *, Runsheng Wang, Cheng Wang (2008)	LBP [34]	<ul style="list-style-type: none"> • The sequence of calculations is essential and manageable to explore images in real-time applications. • Ease of usage • Invariance to 	<ul style="list-style-type: none"> • On-site adequacy assessment cannot be performed • Receptive to the noise of the image • Incompetent

		monotonic light changes	to obtain information of macrostructure
AlimaDamakMasmoudi; NorhenGargouri Ben Ayed et.al (2015)	LBPV [35]	<ul style="list-style-type: none"> • A commutable weight is allocated to the LBP codes to get the LBPV histogram 	<ul style="list-style-type: none"> • Nil
RinkuRabidas, Abhishek Midya, et al (2016)	CLBP [22]	<ul style="list-style-type: none"> • Has large feature size 	<ul style="list-style-type: none"> • The performance of this method is lower than LBP
Li Liu, Songyang Lao et al (2016)	MRELBP [25]	<ul style="list-style-type: none"> • Extremely strong to image clamor this incorporates salt-and-pepper noise, Gaussian noise, random pixel corruption, and Gaussian blur. • Strong discriminativene ss • Grayscale and rotation invariance • There is no need to have parameters tuning and pretraining • Computational efficiency • Computational complexity is slower 	<ul style="list-style-type: none"> • Accuracy is less in object recognition
Hui Zhou, Runsheng Wang, Cheng Wang (2008)	ELBP [27]	<ul style="list-style-type: none"> • Good texture classification 	<ul style="list-style-type: none"> • Responsive to noise and image blur

		performance	<ul style="list-style-type: none"> • Losing to gain extraordinary feature and texture macrostructure
MohdDilshad Ansari, Satya Prakash Ghrera (2016)	IFLBP [37]	<ul style="list-style-type: none"> • When comparing FLBP histograms and crisp LBP, the current method captures more information. • Obtains continuously higher than or equivalent to the entropies acquired by FLBP and LBP. • Supplement more powerful than one bin in sharing the IFLBP standards 	<ul style="list-style-type: none"> • Nil

In the above Table.1 depicts the various techniques which are discussed in this study. The different techniques are highlighted along with the advantages and limitations [57].

Conclusion

In this survey, diversity of feature extraction methods of breast cancer have been studied and discussed. Breast Cancer is the resulting primary reason for the death in women, so it's important to investigate different strategies. If breast malignant growth is distinguished at its untimely stage it can expand the pace of endurance. Cancer is one of the oldest ailments and plenty of research has been conducted in this area. In spite of the huge detriment of Principal Component Analysis is to decrease a measurement which is appropriate for recognizing real cases from dismissing cases. The feature extraction models such as Gray Level Co-occurrence Matrix and Local Binary Pattern performs best for extraction of image features, in which features were trained using several classifiers to separate normal and abnormal cells. But the technique Local Binary Pattern produces larger detection error. Computer-Aided Diagnosis method provides assistance to the doctors in the detection/diagnosis of malformations faster than the traditional methods but do not provide robust and accurate performance. Intuitionistic Fuzzy Local Binary Pattern is the expansion of Fuzzy Local Binary Pattern procedure which adds more than one canister in the appropriation of the Intuitionistic Fuzzy Local Binary Pattern esteems.

Median Robust Extended Local Binary Pattern is pleasant in computational simplicity, rotation invariance, and gray scale but accuracy is very less in image recognition. To overcome all the drawbacks faced by the different techniques, a novel technique would be proposed to have robust and versatile processing which enhances the discrimination of the image in finding the region of interest clearly and effectively.

References

- [1] Mussarat Yasmin, Muhammad Sharif and SajjadMohsin (2013), Survey Paper on Diagnosis of Breast Cancer Using Image Processing Techniques. *Research Journal of Recent Sciences*, 2(10), 88-98.
- [2] Adam, B., Jagtap, S., Anjum, W., Yegingil, H., Wan Y. Shih., Wei-Heng Shih. and Ari D. Brooks (2009), Modern Breast Cancer Detection: A Technological Review. *International Journal of Biomedical Imaging*, 2009: 1-15.
- [3] Arpana M.A, Prathiba Kiran (2014), Feature Extraction Values for Digital Mammograms, *International Journal of Soft Computing and Engineering*, 4(2): 2231-2307.
- [4] Kaur Prabhjot, Kaur Amardeep (2016), Review of Different Approaches in Mammography, *International Journal of Advance Research, Ideas and Innovations in Technology*, 2:1 - 4.
- [5] Sumit Kumar Sourav, Rajeev Ranjan, Manjunath M, Fazeel Ur Rehman Khan and Savitha (2018), A Survey on Computer Aided Diagnosis System for Mammogram Classification, *International Research Journal of Engineering and Technology*, 5 (5): 2799 - 2802.
- [6] Hala Al-Shamlan and Ali El-Zaart (2010), Feature Extraction Values for Breast Cancer Mammography Images, *2010 International Conference on Bioinformatics and Biomedical Technology*, Chengdu, China, 2010, (pp. 335-340), doi: 10.1109/ICBBT.2010.5478947.
- [7] Llobet R., Paredes R., Pérez-Cortés J.C. (2005), Comparison of Feature Extraction Methods for Breast Cancer Detection. In: Marques J.S., Pérez de la Blanca N., Pina P. (eds) *Pattern Recognition and Image Analysis. IbPRIA2005. Lecture Notes in Computer Science*, vol. 3523. Springer, Berlin, Heidelberg. https://doi.org/10.1007/11492542_61.
- [8] Monika Sharma, R. B. Dubey, Sujata, S. K. Gupta (2012), Feature Extraction of Mammograms, *International Journal of Advanced Computer Research*, 2 (3): 201-209.
- [9] Shameena N, RahnaJabbar (2014), A Study of Preprocessing and Segmentation Techniques on Cardiac Medical Images, *International Journal of Engineering Research & Technology*, 3 (4): 336-341.
- [10] PrannoyGiri and K Sara Vanakumar (2017), Breast Cancer Detection using Image Processing Techniques, *Oriental Journal of Computer Science & Technology*, 10 (2): 391-399.
- [11] R. Llobet and J. C. Perez-Cortes (2004), "Breast cancer detection in digitized mammograms using non-parametric methods", *Proceedings of the 2nd International Conference on Advances in Biomedical Signal and Information Processing (MEDSIP)*, 1: 281-287, Sliema, Malta.
- [12] Llobet R., Toselli A.H., Perez-Cortes J.C., Juan A. (2003), Computer-Aided Prostate Cancer Detection in Ultrasonographic Images. In: Perales F.J., Campilho A.J.C., de la Blanca N.P., Sanfeliu A. (eds) *Pattern Recognition and Image Analysis. IbPRIA 2003. Lecture Notes in*

- Computer Science*, vol 2652. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-44871-6_48
- [13] GuangxingGuo Taiyua, Taiyuan & NavidRazmjoooy (2019). A new interval differential equation for edge detection and determining breast cancer regions in mammography images. *Journal Systems Science & Control Engineering*,7 (1): 346-356.
- [14] Rui Xu, Shoji Kido, Kazuyoshi Suga, Yasushi Hirano, Rie Tachibana, Keiichiro Muramatsu, KazukiChagawa, Satoshi Tanaka (2014), Texture analysis on 18F-FDG PET/CT images to differentiate malignant and benign bone and soft-tissue lesions. *The Japanese Society of Nuclear Medicine*, 28(9), 926- 935.
- [15] Michael D. Heath and Kevin W. Bowyer (2000). Mass detection by relative image intensity. *In Proceedings of the 5th International Workshop on Digital Mammography (IWDM-2000)*, 219-225.
- [16] Sunil Arya, David M. Mount, Nathan S. Netanyahu, Ruth Silverman, Angela Y. Wu (199). An optimal algorithm for approximate nearest neighbor searching. *Journal of the ACM*. 45 (6): 891 - 923.
- [17] Monika Sharma, R., Dubey, B., Sujata, S. and Gupta, K. (2012), Feature Extraction of Mammograms. *International Journal of Advanced Computer Research*, 2(5), 201-209.
- [18] Sharma, S. and Khanna, P. (2014). Computer-Aided Diagnosis of Malignant Mammograms using Zernike Moments and SVM. *Journal of Digital Imaging*, 28: 77-90.
- [19] Deserno, T. M. (ed.) (2011), Biomedical Image Processing. *Journal of Digital Imaging*, 25 (6): 689 - 691.
- [20] Guzmán-Cabrera, R.,Guzmán-Sepúlveda, J. R., Torres-Cisneros, M., May-Arrijoja, D. A., Ruiz-Pinales, J., Ibarra-Manzano, O, G., Aviña-Cervantes, G., González Parada, A. (2013), Digital Image Processing Technique for Breast Cancer Detection.*International Journal of Thermophysics*, 34: 1519-1531.
- [21] Chandra Prabha, K. and Bharathi, C. (2019), Texture Analysis using GLCM & GLRLM Feature Extraction Methods. *International Journal for Research in Applied Science & Engineering Technology (IJRASET)*, 7(5), 2059-2064.
- [22] RinkuRabidas, Abhishek Midya, Jayasree Chakraborty, WasimArif (2016). A study of different texture features based on local operator for benign-malignant mass classification. *Procedia Computer Science*, 93: 389-395.
- [23] Guo, Z., Zhang, L. and Zhang, D. (2010), Rotation invariant texture classification using LBP variance (LBPV) with global matchin. *Pattern Recognition*, 43 (3): 706-719.
- [24] Guo, Z., Zhang, D. and Zhang, D. (2010), A completed modeling of local binary pattern operator for texture classification. *IEEE Transactions on Image Processing*. 19 (6): 1657-1663.
- [25] Li Liu, Songyang Lao, Paul W. Fieguth, YulanGuo, Xiaogang Wang, MattiPietikäinen (2016). Median Robust Extended Local Binary Pattern for Texture Classification. *IEEE Transactions on Image Processing*. 25 (3): 1368-1381.

- [26] Ojala, T., Pietikainen, M. and Maenpaa, T. (2012), Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7): 971-987.
- [27] MohdDilshad Ansari, Satya Prakash Ghrera (2016). Feature Extraction Method for Digital Images Based on Intuitionistic Fuzzy Local Binary Pattern. *2016 International Conference System Modeling & Advancement in Research Trends (SMART)*, Moradabad, 2016, (pp. 345-349), doi: 10.1109/SYSMART.2016.7894547.
- [28] GuangxingGuo. &NavidRazmjoooy (2019), A new interval differential equation for edge detection and determining breast cancer regions in mammography images. *Systems Science & Control Engineering*, 7(1): 346-356.
- [29] Gabor Takacs, Bela Pataki (2005), Fast Detection of Mammographic Masses with Difficult Case Exclusion. *2005 IEEE Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications*, Sofia, Bulgaria, 2005, (pp. 424-428), doi: 10.1109/IDAACS.2005.283016.
- [30] Xin Zhang, Jintian Cui, Weisheng Wang, Chao Lin (2017), A Study for Texture Feature Extraction of High-Resolution Satellite Images Based on a Direction Measure and Gray Level Co-Occurrence Matrix Fusion Algorithm. *Sensors*, 17(7): 1 - 15.
- [31] Sumit Kumar Sourav, Rajeev Ranjan, Manjunath M, Fazeel Ur Rehman Khan, Savitha (2018), A Survey on Computer Aided Diagnosis System for Mammogram Classification. *International Research Journal of Engineering and Technology (IRJET)*, 5(5): 2799-2802.
- [32] Jinshan Tang, Rangaraj M. Rangayyan, Jun Xu, Issam El Naqa, and Yongyi Yang (2009), Computer-aided detection and diagnosis of breast cancer with mammography: recent advances. *IEEE Transactions on Information Technology in Biomedicine*, 13(2): 236-251.
- [33] Monica EzzatGamil ; Mariam Mohamed Fouad ; Mohamed A. Abd El Ghany ; Klaus Hoffinan. Fully automated CADx for early breast cancer detection using image processing and machine learning. *2018 30th International Conference on Microelectronics (ICM)*, Sousse, Tunisia, 2018, (pp. 108-111), doi: 10.1109/ICM.2018.8704097.
- [34] Hui Zhou, Runsheng Wang (2008), A novel extended local-binary-pattern operator for texture analysis. *Information Sciences*, 178(22): 4314 - 4325.
- [35] AlimaDamakMasmoudi, NorhenGargouri Ben Ayed, DorraSellamiMasmoudi, RiadAbid. (2015), Robust mass classification-based local binary pattern variance and shape descriptors. *International Journal of Signal and Imaging Systems Engineering*, 8(1): 20-27.
- [36] Preetha, K., Jayanthi. S. K. (2018), GLCM and GLRLM based Feature Extraction Technique in Mammogram Images. *International Journal of Engineering and Technology*, 7(2): 266 - 270.
- [37] MohdDilshad Ansari, Satya Prakash Ghrera (2016), Feature extraction method for digital images based on intuitionistic fuzzy local binary pattern. *2016 International Conference System Modeling & Advancement in Research Trends (SMART)*, Moradabad, 2016, (pp. 345-349), doi: 10.1109/SYSMART.2016.7894547.

- [38] Sarmistha Chakraborty, MrinalKantiBhowmik, Anjan Kumar Ghosh, Tannistha Pal (2016),Automated edge detection of breast masses on mammograms. *2016 IEEE Region 10 Conference (TENCON)*, Singapore, 2016, (pp. 1241-1245), doi: 10.1109/TENCON.2016.7848209.
- [39] Dheeba, J., Albert Singh, N. and Tamil Selvi, S. (2014), Computer-aided detection of breast cancer on mammograms: A swarm intelligence optimized wavelet neural network approach. *Journal of Biomedical Informatics*, 49: 45-52.
- [40] Julia, M., Calas, G., Gutfilen, B., Coelho, W. and Pereira, D. A. (2012), CAD and mammography: why use this tool? *RadiolBraseleira*, 45(1): 46 - 52.
- [41] AfsanehJalalian,SyamsiahMashohor ,Rozi Mahmud ,BabakKarasfi ,M Iqbal B Saripan , Abdul Rahman B. Ramli (2017), Foundation and methodologies in computer-aided diagnosis systems for breast cancer detection. *Excli Journal*, 16: 113-137.
- [42] Hamid El malali ,AbdelhadiAssir ,Mohammed Harmouchi ,AissamLyazidi, MouradRattal, AzeddineMouhsen. (2020), Fully Automatic Computer-Aided Detection of Breast Cancer based on Genetic Algorithm Optimization.*2020 1st International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET)*, Meknes, Morocco, 2020, (pp. 1-6), doi: 10.1109/IRASET48871.2020.9092186.
- [43] Grgel, P., Sertbas, A. and Uan, O. N. (2015), Computer-aided classification of breast masses in mammogram images based on spherical wavelet transform and support vector machines. *Expert Systems*, 32 (1),155 – 164.
- [44] Christoyianni, I., Koutras, A., Dermatas, E. and Kokkinakis, G. (2002), Computer aided diagnosis of breast cancer in digitized mammograms, *Computerized Medical Imaging and Graphics*. 26(5): 309-319.
- [45] Chang, Y. H., Bin Zheng, Xiao-Hui Wang, Good, W. F. (1999), Computer-Aided Diagnosis of Breast Cancer using Artificial Neural Networks: Comparison of Backpropagation and Genetic Algorithms. *IJCNN'99. International Joint Conference on Neural Networks. Proceedings (Cat. No.99CH36339)*, Washington, DC, USA, 1999, (pp. 3674-3679), vol.5, doi: 10.1109/IJCNN.1999.836267.
- [46] Sheshadri H.S, Kandaswamy A. (2004), Computer aided decision system for early detection of breast cancer. *Indian Journal of Medical Research*, 124(2): 149 - 154.
- [47] Mehul P. Sampat, Mia K. Markey, Alan C. Bovik (2005), Computer-aided detection and diagnosis in mammography. *Handbook of image and video processing*, 2(1): 1195 - 1217.
- [48] Rangaraj M. Rangayyan, Fabio J. Ayres, and Leo Desautels, J. E. (2007), A review of computer-aided diagnosis of breast cancer: Toward the detection of subtle signs. *Journal of the Franklin Institute*, 3(4): 312 - 348.
- [49] Loris Nanni, Alessandra Lumini, and Sheryl Brahnem (2010), Local binary patterns variants as texture descriptors for medical image analysis, *Artificial intelligence in medicine*, 49(2): 117 - 125.
- [50] Naresh, S. and Vani Kumari, S. (2015), Breast cancer detection using Local Binary Patterns. *International Journal of Computer Applications*, 123(16): 6 - 9.

- [51] Ponraj, N., Poongodi, LourthuHepziba Marcy, M. (2016), Texture Analysis of Mammogram for the Detection of Breast Cancer using LBP and LGP: A Comparison. *2016 Eighth International Conference on Advanced Computing (ICoAC)*, Chennai, 2017, (pp. 182-185), doi: 10.1109/ICoAC.2017.7951766.
- [52] Rajendra Prasad, K., Raghavendra, C., Sai Saranya, K. (2017), A review on classification of breast cancer detection using combination of the feature extraction models, *International Journal of Pure and Applied Mathematics*, 116(21): 203 - 208.
- [53] Ranjeet Singh Tomar, Tripty Singh, SulochanaWadhvani, Sarita Singh Bhadoria (2009), Analysis of Breast Cancer using Image Processing Techniques. *2009 Third UKSim European Symposium on Computer Modeling and Simulation*, Athens, Greece, 2009, (pp. 251-256), doi: 10.1109/EMS.2009.103. Tobias.
- [54] ChristiianCahoon, Melanie A. Sutton, James C. Bezdek (2000). Breast Cancer Detection Using Image Processing Techniques. *Ninth IEEE International Conference on Fuzzy Systems. FUZZ- IEEE 2000 (Cat. No.00CH37063)*, San Antonio, TX, USA, 2000, (pp. 973-976), vol. 2, doi: 10.1109/FUZZY.2000.839171.
- [55] Nithya, R., Santhi, B (2011), Comparative Study on Feature Extraction Method for Breast Cancer Classification. *Journal of Theoretical and Applied Information Technology*, 33(2): 220 - 226.
- [56] Pradeep. N., Girisha. H., Sreepathi. B., Karibasappa. K. (2012), Feature Extraction of Mammograms. *International Journal of Bioinformatics Research*, 4(1): 241 - 244.
- [57] Ireaneus Anna Rejani, Y., ThamaraiSelvi, S. (2009), Early Detection of Breast Cancer using SVM Classifier Technique, *International Journal on Computer Science and Engineering*, 1(3): 127 - 130.
- [58] Robert M. Haralick, Shanmugam, K. and Its'HakDinstein (1973), Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics*,3(6): 610 - 621.
- [59] Bovis, K. J. and Singh, S. (2000). Detection of masses in mammograms using texture features, *Proceedings 15th International Conference on Pattern Recognition. ICPR-2000*, Barcelona, Spain, 2000, (pp. 267-270), vol. 2, doi: 10.1109/ICPR.2000.906064.
- [60] Olfa Ben Sassi, Lamia sellami, Mohamed Ben Slima, Khalil Chtourou, Ahmed Ben Hamida (2012), Improved spatial gray level dependence matrices for texture analysis, *International Journal of Computer Science & Information Technology (IJCSIT)*, 4(6): 209 - 219.
- [61] Than ThanHtay, Su Su Maun(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)*, Bangkok, 2018, (pp. 171-175), doi: 10.1109/ISCIT.2018.8587920.
- [62] MellisaPratiwia, Alexandra, JeklinHarefaa, SakkaNandaa (2015), Mammograms Classification using Gray-level Co-occurrence Matrix and Radial Basis Function Neural Network. *Procedia Computer Science*, 59: 83-91.