A Review on Region of Interest Segmentation Based on Clustering Techniques for Breast Cancer Ultrasound Images

Abstract

The most prevalent cancer amongst women is woman breast cancer. Ultrasound imaging is a widely employed method for identifying and diagnosing breast abnormalities. Computer-aided diagnosis technologies have lately been developed with ultrasound images to help radiologists enhance the accuracy of the diagnosis. This paper presents several ultrasound image segmentation techniques, mainly focus on eight clustering methods over the last 10 years, and it shows the advantages and disadvantages of these approaches. Breast ultrasound image segmentation is, therefore, still an accessible and challenging issue due to numerous ultrasound artifacts introduced in the imaging process, including high speckle noise, poor contrast, blurry edges, weak signal-to-noise ratio, and intensity inhomogeneity.

Keywords: *Breast cancer, Computer-aided diagnosis, Ultrasound image, Segmentation, Clustering.*

I. Introduction

Breast cancer disease is a standout amongst the most dangerous illnesses faced by ladies in advanced countries [1]. The most familiar type of cancer in ladies is breast malignancy, and over 8% of females get sick with that illness in their lifespan [2,3]. Although the main factors of that disease are still unidentified; however, immediate discovery is beneficial in reducing the number of people who die from that tumour, and maybe over 40% survive [4]. On-time diagnosis of chest tumour raises the chances of curing and fighting against that disease. Nevertheless, on-time recognition needs a precise and dependable finding that must differentiate the unnaturally grown tumours from serious malignancy. An effective way of recognition could cause positive and adverse conditions [5,6].

Though there are few limitations, mammography is utilized as one of the most important and practical techniques in detecting premature breast tumours [2,7]. Non-malignant damage (FP value) might be taken seriously. On the other hand, tumours (FNvalue) could be ignored when using mammography. Consequently, several redundant tissue procedures (65-85%) are applied [8], and radiologist doesn't succeed in detecting 10-30% of chest tumours among women. Furthermore, breast tumours in grown-up females with other supportive tissues can be barely detected [9,10].

Meanwhile, radioactivity used through mammography may put the life of the diagnosed people in jeopardy. For that reason, ultrasound can be more productive and safer option in contrast to mammography because of the accompanying benefits: no radioactivity, quicker screening, superior effectiveness, and precision, and less expense [11-13]. Additionally, it demonstrates an organic matter in using ultrasound to find breast tumour [14,15]. Findings have indicated that ultrasound images are being used widely by scientists, and the number of its users is increasing noticeably [16]. Researches have shown that applying ultrasound could separate gentle tumours from cancerous ones more precisely [11, 17], grow a complete cancer recognition by 17% [18]. Clinically, the separation is mainly done by physically locating that is boring and taking too much time and too many skills,

as well as very much experience. So, for developing the computerization and strength of Computer-Aided Diagnosis (CAD) structures, dependable and programmed or unprogrammed separation techniques are selected to group the region of interest from the ultrasound images. However, experienced and well-trained radiologists are needed to interpret ultrasound images produced by Breast Ultrasound (BUS). Thus, it is crucial to have computer systems that are able to assist junior radiologists in terms of image processing in the detection of breast cancer. CAD systems were developed to produce non-biased diagnosis, reliable, and accurate diagnosis as a secondary judgment to form strong reasoning for doctors to single out benign breast tumors from malignant breast tumors [19, 20]. The most critical and overwhelming stage in the processing of images is image segmentation. The stage is a vital element towards successfully analyzing images and recognizing patterns [21]. Image segmentation is a set of procedures used to split a screened image into several regions. In processing medical images, image segmentation plays a pivotal role in segmenting the representation of tissues of interest from the background. As ultrasound (US) images are always crowded with image noises, it becomes practically challenging to segment tumors in BUS images [22, 23]. Specifically, image noises that pose challenges to segment BUS images as they effectively include intensity inhomogeneity, low signal-to-noise ratio, and high speckle noise [24]. In the medical examination of image segmentation, the task is typically accomplished by labor-intensive human efforts to perform tracing. Such a strategy is laborious, involves specific skillset, requires substantial experience, and consumes a significant amount of time. Therefore, developments in CAD systems are also deemed very important because radiologists can accurately and efficiently diagnose breast cancer [25, 26]. Existing CAD systems are divided into two types, namely completely automatic and semi-automated, based on the level of human intervention in image segmentation tasks. In semi-automated systems, the tasks of specifying a region of interest (ROI) such as a seed in the lesion, an initial boundary, or lesions are accomplished with the expertise of radiologists. “Fig. 1”. shows an example of manually segmented ROI by an expert highlighted as a red line border of breast tumor in an ultrasound image.

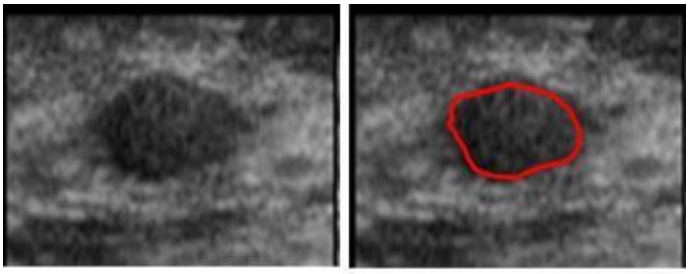

Fig. 1. Illustrates the process of segmentation: a) Original US image b) Segmented US image [27].

In contrast, fully automated systems are independent of human involvement and are capable of modeling breast ultrasound and oncology knowledge as preceding constraints autonomously. The primary advantages of fully automated systems include reproducibility and independency from operators [28]. A standard CAD system performs four primary tasks, as shown in “Fig. 2”. Encompassing: 1) pre-processing, 2) segmentation, 3) Selection of features, and 4) classification. Accuracy of segmentation tasks dramatically affects the results of CAD systems as numerous critical features used for distinguishing malignant and benign tumors rely on texture, contour, and shape of lesions. The features may only be effectively extracted if the segmentation of tumors is performed with high accuracy [29-32].

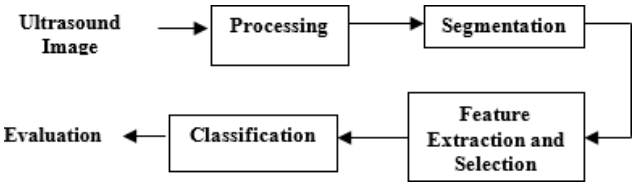

Fig. 2. CAD system for breast cancer detection and classification [33].

Data mining has developed and turned out to be a powerful and robust tool because extracting beneficial records out of tons of engineering and commercial data, the use of multi strategies to analyze data specific as much classification clustering and association [1, 34, 35]. Clustering is an essential method for dealing with comprehensive data that is explosively generated on the Internet [36, 37]. As this paper focuses on clustering techniques image segmentation, we should first give a short explanation of what cluster is, clustering is a critical task of processing images. Clustering, which is an unsupervised learning task, involves having a finite range of categories called pixel identification clusters. A criterion for correspondence is defined between pixels; then, analogous pixels are collectively assembled to build clusters. Pixels grouped into clusters are dependent upon the law of exploiting the intra-class similarity and curtailing the inter-class similarity. The technique of clustering seeks to utilize the association of patterns of the collection by the organism in patterns in clusters or classes such that patterns within a cluster are extra similar to each other when compared to patterns in other clusters. The clustering result consistency is based on the metric of similarity used by the process and its implementation. Superior clustering strategy will build clusters of the highest standard with a robust intra-class similarity and low inter-class similarities. Also, the excellence of a clustering technique is deliberated by its discovery capability. Clustering is nothing more than the categorization of objects in groups based on specific belongings of those objects. In clustering approaches, an attempt is made to find out a vector from local regions of the image. A standard clustering method is to allocate every pixel to the mean of the nearby cluster [2]. Later in this paper, we will explain the most clustering method that has been used for breast cancer ultrasound image segmentation. Thus, this paper concentrates on the summarizing of the developed algorithms for breast cancer segmentation using ultrasound images. Based on our review, there are only a few surveys in the literature for breast cancer segmentation based on the clustering of ultrasound images. Most developed approaches for the last ten years of breast cancer segmentation for ultrasound-based clustering have been reviewed. Therefore, this paper can help researchers who develop the approaches for breast cancer segmentation or identification using a CAD system for ultrasound images.

II. Breast Cancer Detection Methods

Computer-Assisted Detection (CAD) systems are also made more straightforward by digital mammograms. Such technologies utilize advanced computer programs to identify signs of malignancy in images. If such patterns are detected, the radiologist is alerted by the CAD system, which will then investigate the area of interest with more caution. CAD can be utilized on either modern mammogram images or on conventional films of mammogram images that have been transferred to digital formats. Many works have suggested that CAD may increase the ability of a radiologist to identify and classify mammograms for abnormalities of breast cancer [30]. One work proposed that the CAD system could have reduced by nearly three-quarters of the number of cancers of a breast that were lost in the screening of mammography film. Some of the other works indicate that adding a CAD system to mammogram image capturing does not remarkably increase the breast abnormalities that are detected inaccurately as possible false positive tumors. However, research needs to be done more extensively to guarantee that CAD does not lead to false-negative or positive results and to define the value and appropriate usage of the technology more clearly. A CAD system for breast cancer capturing has recently been approved by the FDA. Early breast cancer detection gives the best survival chance [27]. The faster an abnormality is detected, the more appropriate treatment options are available. Guarantees the most significant result possible. There are several ways in which breast cancer can be diagnosed. These include:

A. Digital Mammography

Several companies have developed mammography machines in a try to enhance the x-ray mammography [38]. Digital mammograms (which still need breast compression) digitally capture the x-ray image, unlike the film mammograms, which generate x-ray images directly onto photographic film. An array of detectors produces a digitized image which can be displayed on a computer screen and manipulated. In theory, this may enable easier identification of tumors that are overcast by the dense breast tissue sometimes seen in younger women [38]. The ability to expand or change the contrast of problematic areas without needing new exposure to x-rays that easy the identification of tumors has been lost by film mammography. The technology could also enhance mammography screening by allowing mammograms to be retrieved, stored, and transmitted electronically. An essential disadvantage to digital mammograms is that the images aren't as highly detailed as the film mammograms. Despite that, digital mammography has been marketed as a significant technological advancement in conventional mammography. More works should be done to evaluate its accuracy.

B. Ultrasound Imaging

Ultrasound imaging is also accompanied by X-ray mammograms to assess if a mass on the breast that appears on mammogram images is a stable or less harm tissue cyst that contains fluid. In the systems of ultrasound image screening, high-frequency sound waves can be produced, which permeate the human body [39]. During the bounce waves off the tissue of the boundary of the human body, distinctive echoes will be created, which a computer utilizes to produce an image called a sonogram. Since the cysts of fluid-filled have a distinguished "sound signature" instead of a solid mass, experts may utilize ultrasound precisely to identify cysts that are existed in the breast. Some palpable lesions (lumps) that can be assessed by ultrasound images, which can assist radiologists.

In contrast, these lumps are not easy to find on mammogram images, particularly within the dense breasts of the women. Research of women with lumps showed that ultrasound images were very efficient in the identification of benign that may have removed, which required more than 50% of the biopsies. More so, some other works showed that the ultrasound images might also help in classifying lesions of non-palpable solid into benign or malignant [40].

Further research suggests that the combination of ultrasound combined with x-ray mammography images might enhance breast cancer screening accuracy and allow early-stage tumor detection in women with dense breasts [41]. More research is required to determine the efficacy of ultrasound as a screening method used in combination with mammography. While ultrasound may be helpful as an alternative to mammography, but when used alone, it will be affected by some limitations on breast cancer identification. Predominantly the ultrasound images are not able to identify small tumors that are smaller than five millimeters is about one-quarter inch—also, abnormalities or microcalcifications associated with proven kinds of breast cancers. However, current progress in the technology if ultrasound images can help in resolving some of the limitations and expand their use in identifying cancer of the breast. But, at this point in development, their ultimate efficiency in the detection of breast cancer cannot be predicted [41].

C. Magnetic Resonance Imaging (MRI)

Physicians use MRI for different medical uses since it was approved in 1985 by the FDA for body imaging [39]. MRI, which is generally considered a safe process, produces an image by evaluating the responses of the tissue component to a magnetic field. The specialization of the system of MRI ameliorated for the screening of breast and authorized by the FDA, as shown in “Fig. 3”, especially for the dense of the breast it has been promised to utilize with mammography as a detection technique. It has been suggested by previous studies, despite MRI sensitive in the abnormality of breast cancer detection, sometimes it cannot show malignancies from other abnormalities of breast tissue that are harmless. An ultrasound, microcalcifications are also not detectable. Despite this, an MRI can detect the presence of breast cancer in patients with sonograms, mammograms, and physical exams that are not final. One more possible utilize of MRI is to identify cancer, which has recurred in the breast that had formerly been subjected to lumpectomy, dislike mammography, MRIs are commonly not limited by scarring that may happen after the surgery. MRI also able to identify lesions in women who have dense breast breasts or breast implants, both of which may interfere with x-ray mammogram interpretation. Therefore, MRI may prove helpful in screening young women at high risk (based on genetic testing or strong family history), who appear to have dense breasts. In this regard, initial results are encouraging, but more research is required to determine the efficacy of MRI breast cancer screening in this population [41].

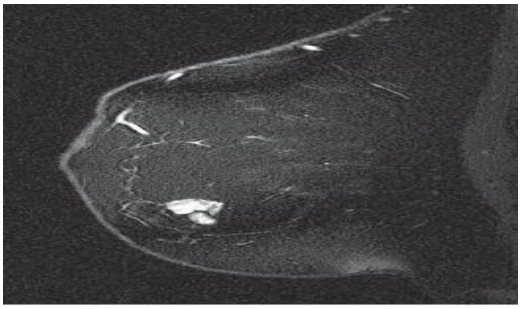

Fig. 3. Example of a magnetic resonance image of the breast [41].

III. Segmentation Techniques for BUS Images

Segmentation of images is a method of dividing a given digital image ***I*** into several non-overlapping regions:

$$\cup I_i = I \quad \text{and} \quad I_i \cap I_j = \emptyset \qquad i \neq j \tag{1}$$

The purpose of separation is to simplify or potentially modify an image's illustration to be more tangible and easier to understand. Separation of pictures is generally utilized to find items and margins (for instance, outlines and curvatures) in pictures. Image separation is the method where all pixels in a picture are given labels that pixels have similar labels to allocate stable features. The result of picture separation is a collection of fragments that includes the whole picture or a collection of shapes separated from the picture. All pixels in an area are like some specific or processed property, for example, shading, power, or quality. Neighboring areas are outstandingly dissimilar regarding the equivalent features. Once operated on a group of images, especially in clinical imaging, the following outlines after image separation could be utilized to produce 3D reproductions through the assistance of introduction calculations such as using 3D shapes. In the chest malignancy CAD framework, the purpose of separation is to divide the ROI, i.e., the vulnerable region, from the background. In this part, we portray the outlook of methods established for BUS image separation by their approved means [2]. “Fig. 4” shows the ultrasound image segmentation techniques we presented in this paper and also shows the clustering methods reviewed. Table I. shows the pros and cons of the reviewed methods.

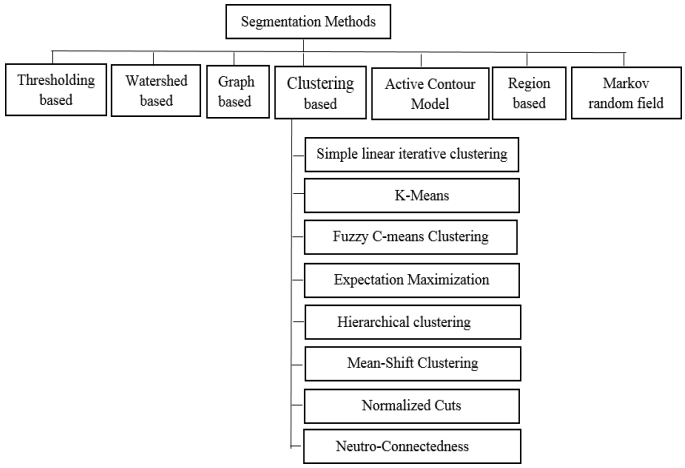

Fig. 4. Most image segmentation techniques and Clustering-based methods used for BUS image segmentation.

A. *Thresholding-based*

One of the commonly utilized methods for the segmentation of monochrome images and has been immensely used to segment ultrasound images is thresholding [42,43]. Thresholding is based upon a clip level or the value of the threshold that convert the image from grayscale into binary, as the simplest way of segmenting images. The classic thresholding approach is relatively basic and primitive. It may not work well with a unimodal histogram for the image since it only considers the statistics of gray-level, without taking any information from a spatial location within the account. Besides, specific images have a certain contrast ratio in the distribution of gray-level among the background and object [25].

Liu et al. [44] introduced a computational framework capable of detecting and segmenting lesions autonomously in medical images. They executed a Non-local Low-Rank Filtering (NLLR) technique to eliminate speckle noise in images. In contour initialization, Liu et al. used the Otsu-based adaptive thresholding (OBAT) algorithm coupled with morphologic steps to classify tumor regions and configure tumor contour effectively. Eventually, an improved Chan-Vese model with the derivation of the ratio of exponentially weighted averages (CV-ROEWA) was used after the initialization of tumor contours.

Zebari et al. [45] Focused on the wavelet-based enhancement stage to produce a quality image that can help improve segmentation and extraction features. In literature, various combinations of thresholding-based techniques were pursued to improve segmentation outcomes, including combining optimization of an ant colony and regulation parameters with the k-means algorithm [46].

B. *Watershed-based*

The watershed change reflects the inclination size of a picture as a topographic outward and is a well-recognized picture separation calculation for grayscale pictures [47, 48]. Pixels with the most remarkable powers of angle size relate to the altered lines. Water set on any pixel that is surrounded by a typical change line streams plunging to a rudimentary adjacent force. Demanding pixels to a shared minimum form symbolizes a section. A blurry level picture can be observed as a topographic help, where a pixel's blurred level is decrypted in mitigation. A drop of water dropping geologically moves beside the way to ultimately reach the nearest spot. Spontaneously, an improvement in the watershed relates to the furthest reach of the drops of water in adjoining the bowls. Numerous kinds of watershed lines might be registered in the process of picture handling. Watershed lines on the centers, boundaries, or crossed type lines on the two centers and the ends might be categorized in diagrams. At the same time, watersheds may be displayed in a dependable area. Additionally, watersheds have a few exceptional results. The watershed algorithm is used mainly for segmentation purposes in image processing. The result of segmentation from [47] is shown in “Fig. 5”.

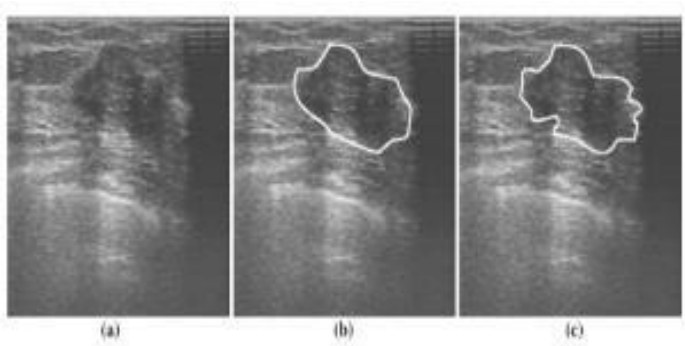

Fig. 5. Transformation based on Watershed for Ultrasound Segmentation: (a) The original Ultrasound, (b) Manual Delineation, and (c) Result of Segmentation[47].

In a recent study, Nugroho et al. [49, 50] introduced the technique of classification of breast nodule characteristics that allow for differentiation into non-circumcised and circumcised categories. The technique employs marker removal using an adaptive median filter, image normalization, and Speckle Reduction Anisotropic Diffusion (SRAD) filter using pre-processing, neutrosophic, and finally segmentation using the watershed technique. Hopeful results have been achieved as the technique has been able to classify breast nodule by exploiting margin characteristics, which is crucial to assist radiologists in the accurate interpretation of US images.

C. Graph-Based Methods

The unsophisticated base and concreate systems, graph-based division have been examined more widely. Recently, the technique of showing pictures has become biased and purposeless. A pixel or a set of pixels is typically linked to bulges and sharp incumbrances, which describes the difference between the pixels within the same area. Then, the image will be separated by a standard design that copies suitable groups. All pixels separated from those calculations are studied as a part of the entity in the image. Several standard calculations of this classification are diagrams, limits, small cuts, less covering tree-based separation, subdivision-based item classification. Typical separation develops from [51, 52] are presented in “Fig 6”.

In 2014, Zhou et al. [51] presented a new semi-automatic segmentation method on ultrasound images utilizing Gaussian filtration, histogram equalization, mean shift, and graph cutting. Only the interaction needed to choose two diagonal points to locate an ROI as image input; after that, by utilizing bicubic interpolation to minimize calculation time, the ROI has been shrunken. After that, a Gaussian filter smoothes the shrunken image, and afterward, the histogram equalization enhances the contrast. Then the image which is enhanced is filtered by the mean pyramid shift to improve homogeneity. On the filtered image, the seeds of background and object are generated automatically for graph cuts. Utilizing these seeds, the filtered image has been converted to the binary that contains image background and the image object by graph cuts. Eventually, the binary image is widened by a factor of two utilizing bicubic interpolation, and to polish the tumor contour, the widened image has been processed based on the opening and closing of morphological operations. The introduced method has been evaluated on 69 ultrasound images from different ultrasound scanners and classified as 38 benign and 31 malignant. Based on the evaluation results, it has been shown that the proposed technique can be used for breast cancer segmentation of ultrasound images.

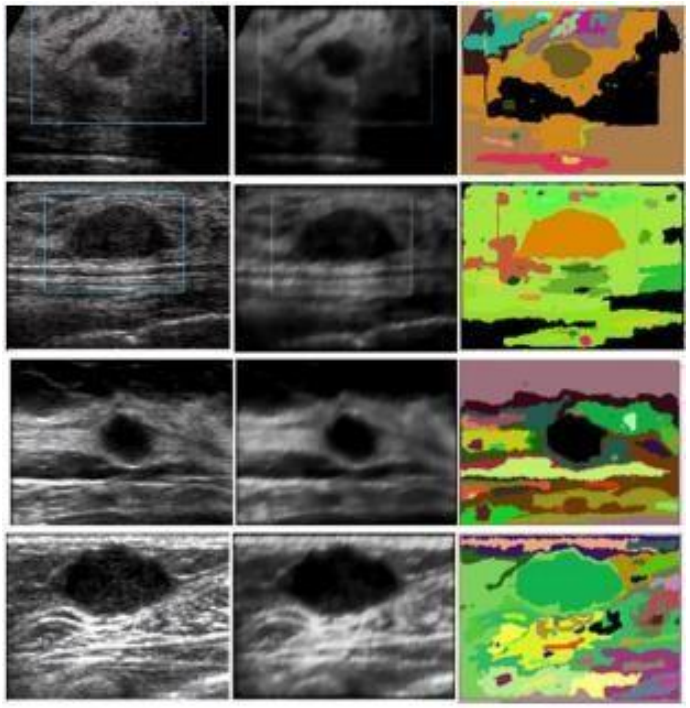

Fig 6: The Result of Segmentation based on graph-based technique. Original ultrasound Showed in the First Column, Filtered Ultrasound Showed in the Second Column, and The Result of Segmentation Showed in the Third Column [2].

D. Active Contour Model

Active contour model is the other very famous technique of separation for ultrasound images, which is usually recognized as the snake [52, 53]. It can be used as an outline for defining an item from a blurry 2D picture and has been enormously applied as a way of separation relying on boundaries. That method searches for decreasing the power linked with the original contour equal to the amount of internal and external energies. The active contour model vigorously changes its structure and comes close to the looked-for contour. In the process of alteration, the energy is computed through using inner and outer energy. The inner energy originated from the contour model is utilized for overwhelming the contour form, and consistency, and for excerpting the contour of the wanted item by using the outer energy brought from the image quality.

Menon et al. [54] proposed simplistic filters to lessen the noise of the speckle and enhance the images. Compared to the traditional SRAD technique, the technique demonstrated superior capabilities in yielding pre-processed images. Active region-based techniques on the local contour can then be used to identify ROI on the pre-processed images effectively. Menon et al. work performed the classification of BUS images through the implementation of feature metrics based on histograms, morphology, and texture. Incorporation of a combination of feature metrics and active contour local region-based technique managed to efficiently perform image segmentation.

E. Region-Based Technique

An image is segmented into different regions by region-based technique. The regions are distinguished based on various criteria, including object, intensity, or color. A region growing algorithm may be used to segment an image into various regions. The algorithm begins with a seed point, or seed area selected, and gradually evaluates and either discards or adds neighbors to a region depending on their value of similarity. Once a stop condition is satisfied, the algorithm will stop. Two useful techniques may be used in region-based techniques encompassing splitting and merging regions and growing regions [55].

Jiang et al. [56] proposed an adaptive OBTA and morphology steps to choose a seed area. Jiang et al. then used a technique that grew an adaptive region to identify potential tumors in the NS domain. The region's growing direction is governed by several constraints, including the distance between the candidate growing points and the seed region, the texture homogeneity value, and differences in the value of similarity.

F. Markov random field

Ultrasound image separation is still shown as a problematic puzzle because of the presence of other manufactured objects. To give an appropriate solution to that trouble, Markov chose a casual area and determined the posterior MAP – based technique and was applied to estimate a falsified area during the classification of areas that have the same amount of strength inhomogeneity. Moreover, US picture separation may be reflected as a category matter where the only way is to allocate pixels to a group of classes, that can be an ordinary symbol for MRFs. Markov casual area model has been utilized for separating Ultrasound images [56, 57]. On the other hand, the calculation approaches the intensification of the class categories' posterior approximation and guesses the group restrictions. The MRF form responds to spatial correlations gained from the reiterative separation development among the labels. The method of transferring the pixel labels would be achieved through adjusting posterior speculation or post-marginal speculation.

In 2012, Xian et al. [58] recommended a form-based possibility scheme for precise and robust separation of poor-quality clinical pictures, joining 3D priori data with MAP possibility regularity limits with the MRF separation background. The three-dimensional form limits the absolute position, the item's model, the form, though the regular field restricts the equal border through displaying the segment characteristic. A BUS record of 131 situations was examined for the suggested process, and the tentative outcomes presented that the applied way was correct and suitable in separating BUS pictures. Label

G. Clustering-based

Grouping could be the primary and significant unverified learning obstacle; therefore, like all other obstacles, it works with obtaining a formation in a group of categorized information. Grouping could be defined as "the process of arranging items into clusters whose components are alike in some ways." That's why a cluster is a group of items that are "alike" among them is "different" to the items related to the other groups [58-62]. A provided method of N dimensions demonstrates each pixel or a set of pixels (i, e., region) in a picture, a comparison that determines the directions and their grouping within the N-dimensional measuring area indicates the connection between the related pixels or group of pixels. Therefore, grouping in measuring an area could be a sign of a picture region comparison and could be applied for separation. The course direction explains a few characteristics of the picture, and for that reason, it is called a feature vector. A comparison between picture regions or pixels indicates grouping (i, e., small separation distance) in the area aspect. The similarity and dissimilarity of [25, 63] can be shown in "Fig 7".

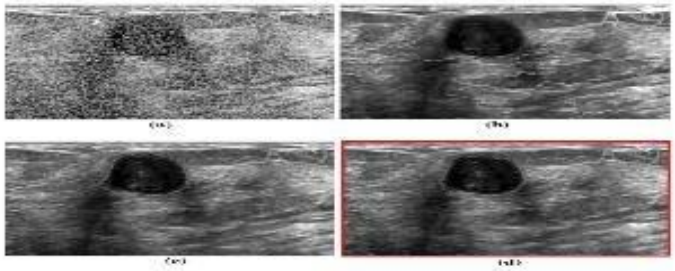

Fig. 7. A Study of the Effects of Segmentation Utilizing Clustering and Transforming the Watershed on the Original Image: *(*a) Applying Watershed Transforms, (b) Applying Marker-Controlled Watershed Transforms, (c) clustering, and (d) Clustering-based Segmentation with the Histogram Equalized Probability Image [64].

1) Simple linear iterative clustering (SLIC)

Numerous clinical imaging technology could gain various kinds of clinical pictures, for example, MRI, CT scan images, ultrasound pictures. Partial domain and signal manufactured objects usually cause an inferior opening. The edge of ROI is ambiguous. The SLIC calculation is an excellent pixel-based separation calculation, better than other pixel-level calculations regarding feature removal and calculable evaluation, keeping data boundaries and the necessity of required data. For that reason, nowadays, the SLIC calculation has been used in clinical picture separation. Fang et al. [65] suggested a new way, a malignant breast categorization with US pictures depended on SLIC (BCCUI). First, he used the Region of Interest (ROI) abstraction depending on Simple Linear Clustering (SLIC) calculation and area expanding calculation to remove ROI at the high-level pixel. Then, the characteristics of the ROI were removed. Moreover, the Support Vector Machine (SVM) organizer was utilized [66]. It verified that the organizer who appeared in work has a specific studying sense and more utilized value.

2) K-Means

The most popular unconfirmed learning calculations that organize the provided information collection into a specified number of groups, the amount set on a priori [66-68]. The pieces of information are grouped utilizing distance metrics like Pearson Correlation, Euclidean, Spearman, Manhattan, and Mahala Nobis. K-Means is a clear and easy grouping technique to increase inter-group or decrease intra-group modification in the process of grouping. However, it might not reach an overall

goal [11,69,70] because of its sensitivity to irregularity and clamor.

Samundeeswari et al. [46] improved K-Means grouping calculation with ACO for group-centric utilization and standardized limitation to enhance and alleviate the groups. Overcoming apparent ambiguity in BUS picture by recovering methods that were used at the beginning of the process. The primary stage also keeps damaged borders and prepares the picture more appropriate for separation. The DE spotted picture was meant for separation by utilizing the suggested Regularized K-Means (ReKM) grouping calculation. In his study, the Ant Colony Optimization (ACO) calculation was applied to filter the primary group centers in the process of grouping. The remote metrics utilized to locate the same pixels in the grouping are improved with the regularization parameter (λ) as a preservative metrical remote. A set of lexical strategies were applied after practicing the method to remove the specific segment of the chest malignant. In the end, the suggested work was correlated to traditional K-Means grouping by utilizing different group metrical authentication.

Çiklaçandir et al. [19] recommended K-Means grouping calculation to find the damage in the pictures. The outcomes of three unlike filters (Median, Laplace, Sobel) were researched in the study. Additionally, various separating effects were taken into consideration. Based on the accurate results, it was clear that the accurateness of the separations amplified. Furthermore, the Median filter was the most excellent in comparison to others.

3) Fuzzy C-means Clustering (FCM)

FCM calculation was initially announced by Dunn [71], and after that, it was made better in quality by Bezdek [72]. It has become the most frequently utilized way for grouping, and it organizes the provided group of information into various comparable groups. Fuzzy C-means grouping calculation is broadly applied in image separation, which contains KFCM, PFC, FCM, HFCM, GFCM, MSFCM, and SL_FCM, etc. [73].

Panigrahi et al. [74] suggested a new grouping technique which is known as Multi-Scale Gaussian Kernel prompted Fuzzy C-Means (MSGKFCM) for separation of cuts mechanically removed Region of Interest (ROI) is an ultrasound to restrict the noise. Moreover, a hybrid method applying Multi-Scale Vector Field Convolution and MsGKFCM was recommended to acquire a perfect cut border in chest ultrasound pictures. Primarily, the pictures were categorized by utilizing Speckle Reducing Anisotropic Diffusion (SRAD) method. Afterward, MsGKFCM was functioned on sorted pictures to separate the noise and discover a suitable group center. The identified group was more utilized by MsVFC to regulate the exact cut of the edge. The experimental outcomes recommended that the suggested method could be applied as a practiced scheme to support clinical specialists by offering factual proofs compared to the state-of-the-art calculations.

Alamelumangai et al. [75] offered a novel method in the field of computing science to diagnose malignancy in the breast by using US images. Several new procedures were taken in the received US image to improve it to get better analysis. In the end, the harmful and malignant cells were differentiated by grouping them. In his research, Modified Fuzzy Possibilities C-Means technique was applied with Repulsion feature to group the cancer cells. The suggested computer programs that were utilized in the examining process were assessed by showing factual chest US images. The accurate prediction of cancer in the breast was more significant in comparison to the traditional method. Additionally, the typical difference brought from the applied method was less typical compared to the traditional system. The tentative outcomes appeared that the applied technique detects cancer areas more accurately because it makes it easier to detect the cancer region.

Prabusankarlal et al. [44] recommended a new technique for the separation of breast US images. The suggested technique uses separate wavelet changes to run a multiresolution analysis of the internal US image. When the resolution level intensifies, the noise decreases, at the same time, the information about the image gets reduced. The suitable resolution level, which includes necessary information for the tumor, was selected inevitably by involving a similar structure. The feature vector for all pixels was structured through trying intra-resolution and inter-resolution information of the picture. The capacity of feature vector was decreased by utilizing important analysis of the elements. The decreased group of feature vectors was separated into two different groups by utilizing dimensional normalized Fuzzy C-Means calculation.

4) Expectation Maximization

Expectation-Maximization (EM) is the most popular calculation applied for intensity approximation of information details in an unconfirmed group. The calculation depends on observing the most significant probability assessments of restrictions when the information model relies on specific underlying changes. In EM, irregular moves of Expectation (E) and Maximization (M) are functioned repeatedly until finding outcomes [76, 77].

The constraints that appeared on the M step are utilized to start another E step, and the action is iterated until the meeting at the same point [78]. Expectation-Maximization grouping [79] reckons the possibility of the groups utilizing the Expectation-Maximization (EM) calculation. The outcome is an expected group of K multivariate divisions, each describing a group, through each example a vector is given to the group with the most restricted possibility. Many expectations on the model resemble a variety of constrictions on the covariance conditions of all divisions. The model will become more stable when the constriction is less, but it would be essential to have more samples to gain positive values of the extra constraints. At all times, the operator of the calculator is I charge of determining the level of constrictions to utilize depending on the available amount of information. The EM calculation is a repeated numerical method to maximize the multipart possibility that pixel fits into the division [80]. The EM calculation involves two primary levels. The first level is guessing the possible allocation of the variable and, on the other level, is amplifying the log-likelihood operation. This repeated process iterates two stages and moves on until meeting at the local optimum [80].

5) Hierarchical clustering

produces a categorized diagram that has similar points between the vectors known as a dendrogram [4]. The typical

performance depends on agglomerative grouping, which prepares the calculation by transfer vector to it in the unattached group and describing the space between each group relying on metrical distances (e.g., Euclidean) or correspondence (e.g., correlation). Then, the algorithm combines the two neighboring groups and renews all the spaces to the newly created group through a related technique, and that is emphasized until one group remains, which includes all vectors. The most common three methods for updating the spaces are complete, single, or average links. That procedure does not explain the division in the system, but a series of organized parts, where all parts include fewer clusters than the former part. To receive a part with K groups, the operation must be ended K-1 stages before its conclusion. Dissimilar links cause different parts; therefore, the kind of utilized link should be chosen based on the type of information for grouping. For example, comprehensive and standard links manage to create compressed clusters; meanwhile, a single link can create groups with more sophisticated forms, but it can be disturbed by unauthentic information [81].

6) Mean-Shift Clustering

It embodies an overall non-parametric method outcome/grouping process in comparison to the conventional method of grouping [82]. There are no fixed expectations on how it is distributed nor how many groups/types are there. The mean shift was firstly recommended by Fukunaga et al. [83]. It was modified by Cheng et al. [84] so that it can analyses images. After that, it was more developed by Comaniciu et al. [80] to inferior vision troubles, especially separation smoothing changes and tracing.

The initial plan in using the mean shift is to deal with the points in the D-dimensional view space as an experimental possibility intensity function where compressed areas match the highest limit or styles of the necessary allocation. For every information in the component area, one plays out a slope increase methodology on the neighborhood calculated intensity until grouping. The fixed purposes of this methodology embody the methods of allocation. Moreover, the information related to the corresponding fixed points is considered as individuals from a similar group.

7) Normalized Cuts (NCut)

It is an unverified separation method established by [84] that doesn't need prior values and has three main features. It manages the separation problem as if it is a picture-part problem, and it is built on a general principle. Moreover, it increases the entire contrast among the different clusters and the entire connection within the clusters. In NCut, all voxels are measured as nodules. A degree of difference between the two nodules may be recognized when relied on illumination, space, color, etc. [85].

8) Neutro-Connectedness

The connectivity of units is a significant topological feature in picture processing. The fuzzy connectivity expressed on the fuzzy subclass has drawn too much focus and has been broadly used in image separation [85-88]. Still, the fuzzy connectivity did not show the intrinsic impossibility of the features of an image, which complicates the separation of objects from low-quality pictures. The next connectedness is an expansion of the fuzzy connectivity: it explains the power of connectivity among pixels of an image, as well as the impossibility of the equivalent connectivity [89, 90].

TABLE I. THE PROS AND CONS OF THE REVIEWED METHODS

Ref.	Segmentation methods	Description	Pros	Cons
[91, 92]	Threshold	Threshold values are founded based on the histogram peaks of an image.	• It is an easy way, no need for prior information. • Good performance when images have prominent edges.	• It cannot act well with images having unimodal histograms. • Inappropriate for images having many or smooth edges.
[92, 33]	Watershed	It is relay on topographical representations.	• Stable results can be shown no matter how the contrast is poor. • Fast, easy and intuitive way, simple, and fast.	• Over- segmentation. • It has noise sensitivity. • It requires complex calculations and Gradients.
[93, 94]	Graph-Based	The original image is mapped to a weighted undirected graph	• Graph-based techniques are popular because of their applicability on global optimization	• Computational complexity is extremely increased when the image size increased.
[33, 2]	Active contour	Snake deformation mode is used.	• A variety of shapes could be used in the extraction of a lesion with maintaining the correct, accurate boundaries	• With images having noise and weak boundaries, low accuracy is shown. • Slow iteration speed.
[92, 95]	Region-based	The image is divided into homogenous regions.	• It shows good immunity to noise where it is hard to identify the edges • Acts well with the easy definition of the similarity. criteria.	• Different results are provided when different seeds selected. • It consumes memory space and time.

[2]	Markov random field	Statistical model that can be utilized within segmentation techniques. MRFs model spatial interactions between neighboring or nearby pixels.	• A mechanism for modeling a diversity of image properties is provided by the local correlations. • It typically used in medical imaging because most pixels belong to the same class as their neighboring pixels.	• Parameters selection can control the strength of spatial interactions. The too-high setting can result in an excessively smooth segmentation and a loss of important structural details. • Computationally intensive algorithms are needed.
[96]	Simple linear iterative	It utilizes K-mean clustering for superpixels generation,	• fast computational speed. • Superior boundary adherence characteristic. • Works well in medical image segmentation	• It does not work well with images having too much.
[91]	K-means	It categorizes the input points of data into many groups produced on their inherent distance from one another	• Simple and fast • It is a systematic algorithm. • It offers the best result when the data set is discrete.	• It is sensitive to outliers, noises, and initial values. The initial values are selected randomly from the data vector.
[97]	Fuzzy C-means Clustering	It depends on the mean of each cluster and gathering similar data values in the same clusters.	• It is a proper clustering technique for medical image segmentation.	• It is challenging to tackle data sets with high dimensional and a massive number of prototypes. • It has initialization sensitivity and is readily fall in local optima • It is sensitive to noises.
[92, 98]	Expectation Maximization	It finds those values of the parameters of the overall model, which maximize the likelihood of the data.	• Avoid merging between object and background. • It has prevailed in many other segmentation techniques, because of its simplicity and performance.	• It fails to utilize the strong spatial correlation between neighboring pixels. • It is sensitive to the noise level.
[99]	Mean-Shift Clustering	Relay on the general notion that locally averaging data results in moving to higher density, and therefore more typical, regions. Iterating this can be done in two distinct ways, depending on whether the dataset itself is updated: mode finding (MS), or smoothing (BMS), both of which can be used for clustering,	• The number of basic image entities can significantly minimize, and due to the significant discontinuity retaining filtering characteristic. • For feature space analysis, it is an enormously versatile tool. • Act well with the arbitrary feature spaces.	• It requires a long computation time. • The only aspect that can control the output is the kernel bandwidth.
[100]	Hierarchical Clustering	Build a dendrogram representing the nested grouping of patterns and the similarity levels at which groupings change.	• A high correlation could be provided with the characteristics of the original database, • These techniques are more versatile. • The distances between each pattern need to be computed, rather than clusters centroid calculations.	• It does not scale up well with the number of observations. • It is hard to handle variant sized clusters and convex shapes. • No objective function is directly minimized
[91, 92]	Normalized Cuts	It depends on graph theory. It was aiming at optimal splitting by minimizing the number of regions.	• After splitting, it does not require merge regions. • It defines edges well.	• Large and complex computation is needed.
[101,102]	Neutro-Connectedness	It generalizes the fuzzy subset and fuzzy connectedness.	• Precise and robust in tumor segmentation • It tackles the problem of a weak boundary in BUS images • Shows indeterminacy of the corresponding connectedness	• Human interaction is required

IV. Performance Evaluation of Segmentation Methods

To quantitatively measure the experiment results, some criteria are adopted in the research. To evaluate the performance of segmentation techniques, the most commonly used evaluation methods are expressed as follows:

$$TP\ Rate = TP / (TP + FN) \quad (2)$$

$$\text{Accuracy} = (TP + TN)/ (TP + TN + FP + FN) \quad (3)$$

$$\text{Sensitivity} = TP/ (TP + FN) \quad (4)$$

To evaluate the performance of the segmentation techniques usually utilized True Positive (TP), False Positive (FP), True Negative (TN), False Negative (FN), and Similarity (SI). One of the standard metrics which has been utilized in the evaluation is TP. Denotes all fraction tissues in the right region that the segmented region overlaps. FP indicates the amount of tissue incorrectly detected by the segmentation technique as a fraction of all amount of tissue in the real tumor region. FN indicates the fraction of tissue known in the real tumor region that is lost by the segmentation technique. More so, the SI, which is also called as overlap or Jaccard index by some researcher assets, the overlap ratio, and (Dice), indicates the dice coefficient. As a result, TP, TN, FP, FN, and SI denote better performance in segmentation methods [2].

$$\text{SI} = \frac{A_s \cap A_g}{A_s \cup A_g} \quad (5)$$

$$\text{FP} = \frac{|A_g \cup A_s - A_g|}{|A_g|} \quad (6)$$

$$\text{FN} = \frac{|A_g \cup A_s - A_s|}{|A_g|} \quad (7)$$

$$Dice = \frac{2\,(A_g \cap A_s)}{|A_g| + |A_s|} \quad (8)$$

A_s is the pixel set of the automatically generated nodule region by the proposed method while A_g is the pixel set of manually outlined nodule regions by the radiologist.

The PRI refers to how exactly the human and computer segmented image is identical; PRI counts the fraction of pairs of pixels whose labelling is consistent between the computed segmentation and the ground truth image. The resultant value ranges between 0-1, and the segments are said to be as identical if it is 1 [2].

$$PR(S_{test,}\{S_k\}) = \frac{1}{\binom{N}{2}} \sum_{\substack{i,j \\ i<j}} C_{ij} + (1 - C_{ij}) * (1 - P_{ij}) \quad (9)$$

where C_{ij} denotes the event and P_{ij} denotes the probability of a pair of pixels *i* and *j* having the same label in the test image. S_k is the set of segments.

V. Experimental Studies of Breast Cancer Ultrasound Image Segmentation Using Clustering

In the past 10 years, several clustering algorithms have been reported in the literature for breast ultrasound image segmentation using clustering; Table II shows the clustering techniques that have been reviewed.

TABLE II. Summary Of Review Of Some Clustering Segmentation Techniques

Ref.	Year	Author(s)	No. Of Images	Segmentation Technique	Metrics
[19]	2019	Çiklaçandir et al.	---------	K-means clustering	Accuracy Rate = 0.7708 With Median Filter
[29]	2018	Lal et al.	60 Total	Neutrosophic ℓ-means (NLM) clustering	TP = 96.04% FP = 2.89% SI = 95.03%
[44]	2016	Prabusankarlal et al.	150 Total 90 Benign 60 Malignant	Spatial regularized fuzzy c-means	Dice = 0.8595
[46]	2016	Samundeeswari et al.	4 Total	Regularized K-Means (ReKM) with λ=0.007	PRI = 96%
[65]	2019	Fang et al.	150 Total 88 malignant 62 benign	Simple Linear Iterative Clustering (SLIC) and region growing	Accuracy = 88.00% Sensitivity = 92.05%
[74]	2019	Panigrahi et al.	127 Total 75 malignant mass 52 are solid benign mass	Multi-scale Gaussian Kernel induced Fuzzy C -means (MsGKFCM)	Accuracy= 97.3158%

[75]	2012	Alamelumangai et al.	10 Total	Modified Fuzzy Possibilistic C-Means with Repulsion	Accuracy = 98.762%
[89]	2014	Xian et al.	131 Total 71 Malignant 60 Benign	Neutro-connectedness	SI= 79.65%
[103]	2012	Shan et al.	—	Neutrosophic ℓ-means (NLM)	TP = 92.4%, FP = 7.2%
[104]	2011	Su et al.	132 Total 67 benign 65 malignant	Normalized Cut	TP Ratio = 95.08%
[105]	2017	Nugroho et al.	62 Total	Neutrosophic FCM with normalisation	Dice = 87.0% FN = 4.0% FP = 25.4%
[106]	2011	Shan el at.	60 Total 29 Benign 31 Malignant	Neutrosophic ℓ-means (NLM) clustering	TP = 94.36% FP = 8.08% SI = 87.39%
[107]	2018	Rajasri et al.	50 Total	Fuzzy Clustering	Accuracy = 73% With Regularization
[108]	2017	Feng et al.	30 Total	Fuzzy c-means (FCM)	Sensitivity = 93.60 ± 5.33%
[109]	2018	Panigrahi et al.	127 Total 75 Malignant 52 Solid Benign	*MsGKFCM_S*	Accuracy = 97.513%

This part is a discussion about the segmentation of breast cancer ultrasound images. Nowadays, breast cancer is the most common form of cancer among the woman worldwide. One key strategy for minimizing breast cancer mortality is the creation of effective early detection techniques. Breast ultrasound (BUS) image segmentation in Computer-Aided Diagnosis (CAD) applications is essential and demanding. Image segmentation attempts to represent the image in a more straightforward and more meaningful manner while maintaining essential features to make interpretation easier. As we reviewed most of the clustering techniques for ultrasound image segmentation, we realize that most of them are having good tumor segmentation accuracy. However, some research is having wrong accuracy segmentation of breast tumor in ultrasound images, those techniques which are having wicked accuracy segmentation the reason is not reducing the amount of noise in the ultrasound image which case segmentation accuracy to decrease, in contrast, if noise removal and image enhancement are performed it will help in better segmentation result.

VI. Conclusion

Ultrasound image segmentation is one of the most challenging and active research areas in the field of medical image processing. The segmentation of ultrasound images is a challenging task, and itis necessary to make future research work more accurate, precise, and speedy. Thus, there is no single method which can be considered suitable for neither all type of images nor all methods equally suitable for a particular type of image. Due to these factors, the segmentation of images remains a challenge in the processing of images and computer vision and remains a pending problem in the world. Further works may be conducted to develop efficient segmentation methods. This review presents an overview of ultrasound imaging segmentation techniques in general and focuses on eight clustering methods for breast cancer detection. These techniques are essential in the diagnosis of benign or malignant breast lesions and can further improve the early detection of breast cancer. This review may help as a guide for residents of the medical sector, junior radiologists and researchers are interested in imaging and identifying breast cancer.

References

[1] D. M. Abdulqader, A. M. Abdulazeez, and D. Q. Zeebaree, 'Machine Learning Supervised Algorithms of Gene Selection: A Review', *Mach. Learn.*, vol. 62, no. 03, 2020.

[2] Q. Huang, Y. Luo, and Q. Zhang, 'Breast ultrasound image segmentation: a survey', *Int. J. Comput. Assist. Radiol. Surg.*, vol. 12, no. 3, pp. 493–507, 2017.

[3] H.-D. Cheng, X. Shi, R. Min, L. Hu, X. Cai, and H. Du, 'Approaches for automated detection and classification of masses in mammograms', *Pattern Recognit.*, vol. 39, no. 4, pp. 646–668, 2006.

[4] H.-D. Cheng, X. Cai, X. Chen, L. Hu, and X. Lou, 'Computer-aided detection and classification of microcalcifications in mammograms: a survey', *Pattern Recognit.*, vol. 36, no. 12, pp. 2967–2991, 2003.

[5] Lee, Carol H. "Screening mammography: proven benefit, continued controversy." Radiologic Clinics of North America 40.3 (2002): 395-407.

[6] Yaseen, D. S., Batumalai, S. A., Ahmed, F. Y., & Fong, S. L. (2019, August). Improved Disabled Mobile Aid Application for Android: Health and Fitness Helper for Disabled People. In *2019 IEEE 10th Control and System Graduate Research Colloquium (ICSGRC)* (pp. 207-212).

[7] D. Q. Zeebaree, H. Haron, and A. M. Abdulazeez, 'Gene selection and classification of microarray data using convolutional neural network', presented at the 2018 International Conference on Advanced Science and Engineering (ICOASE), 2018, pp. 145–150.

[8] J. L. Jesneck, J. Y. Lo, and J. A. Baker, 'Breast mass lesions: computer-aided diagnosis models with mammographic and sonographic descriptors', *Radiology*, vol. 244, no. 2, pp. 390–398, 2007.

[9] R. E. Bird, T. W. Wallace, and B. C. Yankaskas, 'Analysis of cancers missed at screening mammography.', *Radiology*, vol. 184, no. 3, pp. 613–617, 1992.

[10] M. L. Giger, 'Computer-aided diagnosis in radiology', *Acad. Radiol.*, vol. 9, no. 1, pp. 1–3, 2002.

[11] B. Sahiner *et al.*, 'Malignant and benign breast masses on 3D US volumetric images: effect of computer-aided diagnosis on radiologist accuracy', Radiology, vol. 242, no. 3, pp. 716–724, 2007.

[12] Zebari, D. A., Haron, H., Zeebaree, D. Q., & Zain, A. M. (2019, August). A Simultaneous Approach for Compression and Encryption Techniques Using Deoxyribonucleic Acid. In 2019 13th International Conference on

Software, Knowledge, Information Management and Applications (SKIMA) (pp. 1-6). IEEE.

[13] B. O. Anderson et al., 'Breast cancer in limited-resource countries: an overview of the Breast Health Global Initiative 2005 guidelines', Breast J., vol. 12, pp. S3–S15, 2006.

[14] P. Shankar, C. Piccoli, J. Reid, F. Forsberg, and B. Goldberg, 'Application of the compound probability density function for characterization of breast masses in ultrasound B scans', Phys. Med. Biol., vol. 50, no. 10, p. 2241, 2005.

[15] K. Drukker, M. Giger, L. A. Meinel, A. Starkey, J. Janardanan, and H. Abe, 'Quantitative ultrasound image analysis of axillary lymph node status in breast cancer patients', Int. J. Comput. Assist. Radiol. Surg., vol. 8, no. 6, pp. 895–903, 2013.

[16] R.-F. Chang, W.-J. Wu, W. K. Moon, and D.-R. Chen, 'Improvement in breast tumor discrimination by support vector machines and speckle-emphasis texture analysis', Ultrasound Med. Biol., vol. 29, no. 5, pp. 679–686, 2003.

[17] C.-M. Chen et al., 'Breast lesions on sonograms: computer-aided diagnosis with nearly setting-independent features and artificial neural networks', Radiology, vol. 226, no. 2, pp. 504–514, 2003.

[18] M. P. André, M. Galperin, L. K. Olson, K. Richman, S. Payrovi, and P. Phan, 'Improving the accuracy of diagnostic breast ultrasound', in Acoustical Imaging, Springer, 2002, pp. 453–460.

[19] F. G. Y. Çiklaçandir, A. Ertaylan, U. Bınzat, and A. Kut, 'Lesion Detection from the Ultrasound Images Using K-Means Algorithm', presented at the 2019 Medical Technologies Congress (TIPTEKNO), 2019, pp. 1–4.

[20] L. Zhang, Y. Ren, C. Huang, and F. Liu, 'A novel automatic tumor detection for breast cancer ultrasound images', presented at the 2011 Eighth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), 2011, vol. 1, pp. 401–404.

[21] K. S. Camilus and V. Govindan, 'A Review on Graph Based Segmentation.', Int. J. Image Graph. Signal Process., vol. 4, no. 5, 2012.

[22] C.-C. Kang, W.-J. Wang, and C.-H. Kang, 'Image segmentation with complicated background by using seeded region growing', AEU-Int. J. Electron. Commun., vol. 66, no. 9, pp. 767–771, 2012.

[23] K. Prabusankarlal, P. Thirumoorthy, and R. Manavalan, 'Computer aided breast cancer diagnosis techniques in ultrasound: a survey', J. Med. Imaging Health Inform., vol. 4, no. 3, pp. 331–349, 2014.

[24] Mazaheri, S., Sulaiman, P. S. B., Wirza, R., Khalid, F., Kadiman, S., Dimon, M. Z., & Tayebi, R. M. (2013, December). Echocardiography image segmentation: A survey. In 2013 International Conference on Advanced Computer Science Applications and Technologies (pp. 327-332). IEEE.

[25] P. S. Rodrigues and G. A. Giraldi, 'Improving the non-extensive medical image segmentation based on Tsallis entropy', Pattern Anal. Appl., vol. 14, no. 4, pp. 369–379, 2011.s

[26] Huang, Yu-Len, Kao-Lun Wang, and Dar-Ren Chen. "Diagnosis of breast tumors with ultrasonic texture analysis using support vector machines." Neural Computing & Applications 15.2 (2006): 164-169.

[27] D. Q. Zeebaree, H. Haron, A. M. Abdulazeez, and D. A. Zebari, 'Machine learning and Region Growing for Breast Cancer Segmentation', presented at the 2019 International Conference on Advanced Science and Engineering (ICOASE), 2019, pp. 88–93.

[28] M. Xian et al., A benchmark for breast ultrasound image segmentation (BUSIS). Infinite Study, 2018. https://arxiv.org/ftp/arxiv/papers/1801/1801.03182.pdf

[29] M. Lal, L. Kaur, and S. Gupta, 'Modified spatial neutrosophic clustering technique for boundary extraction of tumours in B-mode BUS images', IET Image Process., vol. 12, no. 8, pp. 1338–1344, 2018.

[30] A. Ibrahim, S. Mohammed, and H. A. Ali, 'Breast cancer detection and classification using thermography: a review', presented at the International Conference on Advanced Machine Learning Technologies and Applications, 2018, pp. 496–505.

[31] R. Zebari, A. Abdulazeez, D. Zeebaree, D. Zebari, and J. Saeed, 'A Comprehensive Review of Dimensionality Reduction Techniques for Feature Selection and Feature Extraction', J. Appl. Sci. Technol. Trends, vol. 1, no. 2, pp. 56–70, 2020.

[32] O. Ahmed and A. Brifcani, 'Gene Expression Classification Based on Deep Learning', presented at the 2019 4th Scientific International Conference Najaf (SICN), 2019, pp. 145–149.

[33] J. N. Saeed, 'A SURVEY OF ULTRASONOGRAPHY BREAST CANCER IMAGE SEGMENTATION TECHNIQUES', Acad. J. Nawroz Univ., vol. 9, no. 1, pp. 1–14, 2020.

[34] D. Q. Zeebaree, H. Haron, A. M. Abdulazeez, and S. R. Zeebaree, 'Combination of K-means clustering with Genetic Algorithm: A review', Int. J. Appl. Eng. Res., vol. 12, no. 24, pp. 14238–14245, 2017.

[35] Ahmed, F. Y., Shamsuddin, S. M., & Hashim, S. Z. M. (2013). Improved SpikeProp for using particle swarm optimization. Mathematical Problems in Engineering, 2013., V 2013, http://dx.doi.org/10.1155/2013/257085.

[36] N. Adeen, M. Abdulazeez, and D. Zeebaree, 'Systematic Review of Unsupervised Genomic Clustering Algorithms Techniques for High Dimensional Datasets', vol. 62, no. 3, 2020.

[37] Najat, N., & Abdulazeez, A. M. (2017, November). Gene clustering with partition around mediods algorithm based on weighted and normalized Mahalanobis distance. In 2017 International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS) (pp. 140-145). IEEE.

[38] National Research Council, Mammography and beyond: developing technologies for the early detection of breast cancer. National Academies Press, 2001.Book

[39] P. Thompson, 'Ultrasound in Obstetrics and Gynaecology: European Practice in Gynaecology and Obstetrics', 2010. 30(8), pp. 887–888.

[40] K. Johnson, D. Sarma, and E. S. Hwang, 'Lobular breast cancer series: imaging', Breast Cancer Res., vol. 17, no. 1, p. 94, 2015.

[41] Adeen, I. M. N., Abdulazeez, A. M., & Zeebaree, D. Q. Systematic Review of Unsupervised Genomic Clustering Algorithms Techniques for High Dimensional Datasets., vol. 62, no. 03, 2020.

[42] K. Drukker, M. L. Giger, K. Horsch, M. A. Kupinski, C. J. Vyborny, and E. B. Mendelson, 'Computerized lesion detection on breast ultrasound', Med. Phys., vol. 29, no. 7, pp. 1438–1446, 2002.

[43] K. Horsch, M. L. Giger, L. A. Venta, and C. J. Vyborny, 'Automatic segmentation of breast lesions on ultrasound', Med. Phys., vol. 28, no. 8, pp. 1652–1659, 2001.

[44] K. Prabusankarlal, P. Thirumoorthy, and R. Manavalan, 'Segmentation of breast lesions in ultrasound images through multiresolution analysis using undecimated discrete wavelet transform', Ultrason. Imaging, vol. 38, no. 6, pp. 384–402, 2016.

[45] D. A. Zebari, H. Haron, S. R. Zeebaree, and D. Q. Zeebaree, 'Enhance the Mammogram Images for Both Segmentation and Feature Extraction Using Wavelet Transform', presented at the 2019 International Conference on Advanced Science and Engineering (ICOASE), 2019, pp. 100–105.

[46] E. Samundeeswari, P. Saranya, and R. Manavalan, 'Segmentation of breast ultrasound image using regularized K-means (ReKM) clustering', presented at the 2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), 2016, pp. 1379–1383.

[47] Y.-L. Huang and D.-R. Chen, 'Watershed segmentation for breast tumor in 2-D sonography', Ultrasound Med. Biol., vol. 30, no. 5, pp. 625–632, 2004.

[48] C.-M. Lo et al., 'Multi-dimensional tumor detection in automated whole breast ultrasound using topographic watershed', IEEE Trans. Med. Imaging, vol. 33, no. 7, pp. 1503–1511, 2014.

[49] H. A. Nugroho, Y. Triyani, M. Rahmawaty, and I. Ardiyanto, 'Breast ultrasound image segmentation based on neutrosophic set and watershed method for classifying margin characteristics', presented at the 2017 7th IEEE International Conference on System Engineering and Technology (ICSET), 2017, pp. 43–47.

[50] H. A. Nugroho, Y. Triyani, M. Rahmawaty, and I. Ardiyanto, 'Analysis of margin sharpness for breast nodule classification on ultrasound images', presented at the 2017 9th International Conference on Information Technology and Electrical Engineering (ICITEE), 2017, pp. 1–5.

[51] Q.-H. Huang, S.-Y. Lee, L.-Z. Liu, M.-H. Lu, L.-W. Jin, and A.-H. Li, 'A robust graph-based segmentation method for breast tumors in ultrasound images', Ultrasonics, vol. 52, no. 2, pp. 266–275, 2012.

[52] N. M. A. Ahmed and A. M. A. Brifcani, 'A new modified embedded zerotree wavelet approach for image coding (nmezw)', Int. J. Sci. Eng. Res., vol. 4, 2013.

[53] Y.-L. Huang, Y.-R. Jiang, D.-R. Chen, and W. K. Moon, 'Level set contouring for breast tumor in sonography', J. Digit. Imaging, vol. 20, no. 3, pp. 238–247, 2007.

[54] R. V. Menon, P. Raha, S. Kothari, S. Chakraborty, I. Chakrabarti, and R. Karim, 'Automated detection and classification of mass from breast ultrasound images', presented at the 2015 Fifth National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG), 2015, pp. 1–4.

[55] W. Khan, 'Image segmentation techniques: A survey', J. Image Graph., vol. 1, no. 4, pp. 166–170, 2013.

[56] X. Jiang, Y. Guo, H. Chen, Y. Zhang, and Y. Lu, 'An adaptive region growing based on neutrosophic set in ultrasound domain for image segmentation', IEEE Access, vol. 7, pp. 60584–60593, 2019.

[57] A. Takemura, A. Shimizu, and K. Hamamoto, 'A cost-sensitive extension of AdaBoost with markov random field priors for automated segmentation of breast tumors in ultrasonic images', Int. J. Comput. Assist. Radiol. Surg., vol. 5, no. 5, pp. 537–547, 2010.

[58] M. Xian, J. Huang, Y. Zhang, and X. Tang, 'Multiple-domain knowledge based MRF model for tumor segmentation in breast ultrasound images', presented at the 2012 19th IEEE International Conference on Image Processing, 2012, pp. 2021–2024.

[59] P. F. Felzenszwalb and D. P. Huttenlocher, 'Efficient graph-based image segmentation', Int. J. Comput. Vis., vol. 59, no. 2, pp. 167–181, 2004.

[60] A. Vedaldi and S. Soatto, 'Quick shift and kernel methods for mode seeking', presented at the European conference on computer vision, 2008, pp. 705–718.

[61] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, 'SLIC superpixels compared to state-of-the-art superpixel methods', IEEE Trans. Pattern Anal. Mach. Intell., vol. 34, no. 11, pp. 2274–2282, 2012.

[62] J.-S. Lee, 'Digital image smoothing and the sigma filter', Comput. Vis. Graph. Image Process., vol. 24, no. 2, pp. 255–269, 1983.

[63] N. N. Mohammed and A. M. Abdulazeez, 'Evaluation of partitioning around medoids algorithm with various distances on microarray data', presented at the 2017 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), 2017, pp. 1011–1016.

[64] Kekre, H. B., and Pravin Shrinath. "Tumour Delineation using Statistical Properties of The Breast US Images and Vector Quantization based Clustering Algorithms." International Journal of Image, Graphics & Signal Processing 5.11 (2013).

[65] Z. Fang, W. Zhang, and H. Ma, 'Breast Cancer Classification with Ultrasound Images Based on SLIC', ArXiv Prepr. ArXiv190411322, 2019.

[66] A. Al-Zebari and A. Sengur, 'Performance Comparison of Machine Learning Techniques on Diabetes Disease Detection', presented at the 2019 1st International Informatics and Software Engineering Conference (UBMYK), 2019, pp. 1–4.

[67] S. S. Khan and A. Ahmad, 'Cluster center initialization algorithm for K-means clustering', Pattern Recognit. Lett., vol. 25, no. 11, pp. 1293–1302, 2004.

[68] E. Samundeeswari, P. Saranya, and R. Manavalan, 'M2 Filter For Speckle Noise Suppression In Breast Ultrasound Images', ICTACT J. Image Video Process., vol. 6, no. 2, pp. 1137–1144, 2015.

[69] J. Guo, E. Levina, G. Michailidis, and J. Zhu, 'Pairwise variable selection for high-dimensional model-based clustering', Biometrics, vol. 66, no. 3, pp. 793–804, 2010.

[70] D. M. Sulaiman, A. M. Abdulazeez, H. Haron, and S. S. Sadiq, 'Unsupervised Learning Approach-Based New Optimization K-Means Clustering for Finger Vein Image Localization', presented at the 2019 International Conference on Advanced Science and Engineering (ICOASE), 2019, pp. 82–87.

[71] J. C. Dunn, 'A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters', V3-1973, issue 3.

[72] J. C. Bezdek, Pattern recognition with fuzzy objective function algorithms. Springer Science & Business Media, 2013, Book.

[73] J. Miao, X. Zhou, and T.-Z. Huang, 'Local segmentation of images using an improved fuzzy C-means clustering algorithm based on self-adaptive dictionary learning', Appl. Soft Comput., p. 106200, 2020.

[74] L. Panigrahi, K. Verma, and B. K. Singh, 'Ultrasound image segmentation using a novel multi-scale Gaussian kernel fuzzy clustering and multi-scale vector field convolution', Expert Syst. Appl., vol. 115, pp. 486–498, 2019.

[75] N. Alamelumangai and J. Devishree, 'Automated clustering of cancer cells using fuzzy C means with repulsions in ultrasound images', J. Artif. Intell., vol. 5, pp. 14–25, 2012.

[76] D. Naik and P. Shah, 'A review on image segmentation clustering algorithms', Int J Comput Sci Inf. Technol, vol. 5, no. 3, pp. 3289–93, 2014.

[77] Zeebaree, D. Q., Haron, H., Abdulazeez, A. M., & Zebari, D. A. (2019, April). Trainable Model Based on New Uniform LBP Feature to Identify the Risk of the Breast Cancer. In 2019 International Conference on Advanced Science and Engineering (ICOASE) (pp. 106-111). IEEE.

[78] S. Tatiraju and A. Mehta, 'Image Segmentation using k-means clustering, EM and Normalized Cuts', Dep. EECS, vol. 1, pp. 1–7, 2008.

[79] K. Y. Yeung, C. Fraley, A. Murua, A. E. Raftery, and W. L. Ruzzo, 'Model-based clustering and data transformations for gene expression data', Bioinformatics, vol. 17, no. 10, pp. 977–987, 2001.

[80] P. Lukac, R. Hudec, M. Benco, P. Kamencay, Z. Dubcova, and M. Zachariasova, 'Simple comparison of image segmentation algorithms based on evaluation criterion', presented at the Proceedings of 21st International Conference Radioelektronika 2011, 2011, pp. 1–4.

[81] E. R. Dougherty et al., 'Inference from clustering with application to gene-expression microarrays', J. Comput. Biol., vol. 9, no. 1, pp. 105–126, 2002.

[82] D. Comaniciu, V. Ramesh, and P. Meer, 'The variable bandwidth mean shift and data-driven scale selection', presented at the Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001, 2001, vol. 1, pp. 438–445.

[83] K. Fukunaga and L. Hostetler, 'The estimation of the gradient of a density function, with applications in pattern recognition', IEEE Trans. Inf. Theory, vol. 21, no. 1, pp. 32–40, 1975.

[84] Y. Cheng, 'Mean shift, mode seeking, and clustering', IEEE Trans. Pattern Anal. Mach. Intell., vol. 17, no. 8, pp. 790–799, 1995.

[85] D. Comaniciu and P. Meer, 'Mean shift: A robust approach toward feature space analysis', IEEE Trans. Pattern Anal. Mach. Intell., vol. 24, no. 5, pp. 603–619, 2002.

[86] J. Carballido-Gamio, S. J. Belongie, and S. Majumdar, 'Normalized cuts in 3-D for spinal MRI segmentation', IEEE Trans. Med. Imaging, vol. 23, no. 1, pp. 36–44, 2004.

[87] A. Rosenfeld, 'Fuzzy digital topology', Inf. Control, vol. 40, no. 1, pp. 76–87, 1979.

[88] L. Chen and J. Zhang, 'Fuzzy subfiber and its application to seismic lithology classification', Inf. Sci.-Appl., vol. 1, no. 2, pp. 77–95, 1994.

[89] M. Xian, H.-D. Cheng, and Y. Zhang, 'A fully automatic breast ultrasound image segmentation approach based on neutro-connectedness', presented at the 2014 22nd International Conference on Pattern Recognition, 2014, pp. 2495–2500.

[90] O. M. S. Hassan, A. M. Abdulazeez, and V. M. TİRYAKİ, 'Gait-based human gender classification using lifting 5/3 wavelet and principal component analysis', presented at the 2018 International Conference on Advanced Science and Engineering (ICOASE), 2018, pp. 173–178.

[91] Norouzi, A., Rahim, M. S. M., Altameem, A., Saba, T., Rad, A. E., Rehman, A., & Uddin, M. (2014). Medical image segmentation methods, algorithms, and applications. IETE Technical Review, 31(3), 199-213.

[92] Xian, M., Zhang, Y., Cheng, H. D., Xu, F., Zhang, B., & Ding, J. (2018). Automatic breast ultrasound image segmentation: A survey. Pattern Recognition, 79, 340-355.

[93] Feng, W., Xiang, H., & Zhu, Y. (2011, May). An improved graph-based image segmentation algorithm and its gpu acceleration. In 2011

Workshop on Digital Media and Digital Content Management (pp. 237-241). IEEE.

[94] Çiğla, C., & Alatan, A. A. (2010, September). Efficient graph-based image segmentation via speeded-up turbo pixels. In 2010 IEEE International Conference on Image Processing (pp. 3013-3016). IEEE.

[95] Sharma, P., & Suji, J. (2016). A review on image segmentation with its clustering techniques. International Journal of Signal Processing, Image Processing and Pattern Recognition, 9(5), 209-218.

[96] Qin, F., Guo, J., & Lang, F. (2014). Superpixel segmentation for polarimetric SAR imagery using local iterative clustering. IEEE Geoscience and Remote Sensing Letters, 12(1), 13-17.

[97] Grover, N. (2014). A study of various fuzzy clustering algorithms. International Journal of Engineering Research, 3(3), 177-181.

[98] Tolba, M. F., Mostafa, M. G., Gharib, T. F., & Salem, M. A. M. (2003). MR-Brain Image Segmentation Using Gaussian Multiresolution Analysis and the EM Algorithm. In ICEIS (2) (pp. 165-170).

[99] Carreira-Perpinán, M. A. (2015). A review of mean-shift algorithms for clustering. arXiv preprint arXiv:1503.00687.

[100] Islam, S., & Ahmed, M. (2013). Implementation of image segmentation for natural images using clustering methods. Int J Emerg Technol Adv Eng, 3(3), 175-80.

[101] Xian, M., Zhang, Y., Cheng, H. D., Xu, F., & Ding, J. (2016). Neutro-connectedness cut. IEEE Transactions on Image Processing, 25(10), 4691-4703.

[102] Koundal, D., & Sharma, B. (2019). Challenges and future directions in neutrosophic set-based medical image analysis. In Neutrosophic Set in Medical Image Analysis (pp. 313-343). Academic Press.

[103] J. Shan, H. Cheng, and Y. Wang, 'A novel segmentation method for breast ultrasound images based on neutrosophic l-means clustering', Med. Phys., vol. 39, no. 9, pp. 5669–5682, 2012.

[104] Y. Su, Y. Wang, J. Jiao, and Y. Guo, 'Automatic detection and classification of breast tumors in ultrasonic images using texture and morphological features', Open Med. Inform. J., vol. 5, no. 1, 2011.

[105] H. A. Nugroho, M. Rahmawaty, Y. Triyani, and I. Ardiyanto, 'Neutrosophic and fuzzy C-means clustering for breast ultrasound image segmentation', presented at the 2017 9th International Conference on Information Technology and Electrical Engineering (ICITEE), 2017, pp. 1–5.

[106] J. Shan, 'A fully automatic segmentation method for breast ultrasound images', 2011, theses and Dissertations. 905

[107] S. Rajasri and D. R. Gandh, 'Rule Based Fuzzy Image Segmentation for the Detection of Breast Cancer from Ultrasound Image', presented at the 2018 International Conference on Intelligent Computing and Communication for Smart World (I2C2SW), 2018, pp. 254–259.

[108] Y. Feng et al., 'An adaptive fuzzy C-means method utilizing neighboring information for breast tumor segmentation in ultrasound images', Med. Phys., vol. 44, no. 7, pp. 3752–3760, 2017.

[109] L. Panigrahi, K. Verma, and B. K. Singh, 'Hybrid segmentation method based on multi-scale Gaussian kernel fuzzy clustering with spatial bias correction and region-scalable fitting for breast US images', IET Comput. Vis., vol. 12, no. 8, pp. 1067–1077, 2018.