Breast Tumor Detection Via Active Contour Technique

Eman A. Radhi[1]   **Mohammed Y. Kamil[1]***

[1]College of Sciences, Mustansiriyah University, Baghdad, Iraq
* Corresponding author's Email: m80y98@uomustansiriyah.edu.iq

Abstract: The task of segmenting breast tumours in mammograms is very difficult, as its difficulty lies in the lack of contrast between the tumour and the surrounding breast tissue, especially when dealing with small tumours that are not clear boundaries and hidden under the tissues. Segmentation algorithms often lose the path of tumor boundaries in an attempt to determine the position of them. Active contours are used widely for segmentation as a high-level technology for boundary recognition. The main aim to create a clear contrast between the tumour and the normal breast region. In this study, two approaches to active contour are applied: snakes and level sets. The proposed methods were applied to all abnormal mammogram images taken from mini-MIAS database. The first approach showed a weakness in the segmentation of this type of image, while the other approach was able to segment all the mammogram's tumours. The Chan-Vese method was the most superior of all the active contour segmentation methods. The proposed models were tested in two ways, the first is statistical the best result was for the Chan-Vese method and it came as follows (90%, 95% 98%, 97%, 97%) for Jaccard, Dice, PF-Score, Precision, and Sensitivity respectively. And the other is based on the segmented region's characteristics, Chan-Vese was able to accurately determine the location and shape of the tumor. The proposed Chan-Vese approach is appropriate in adopting computer assisted detection systems to predict tumor boundaries and locations in mammography for its reliability and superior performance over other algorithms.

Keywords: Breast tumour, Mammography, Segmentation, Active contour, Level set, Chan-Vese.

1. Introduction

Breast cancer is one of the world's leading causes of death of women. An early breast cancer diagnosis is important to increase the survival rate [1]. Since the similarities between breast masses and microcalcifications with the background, the manual methods which are practiced by radiologists May fail to segment such abnormalities. The need for early diagnosis includes the development of automated technologies that can enable radiologists to reliably identify breast cancer and treat patients. The most powerful image system for the identification and diagnosis of breast cancer was considered mammography [2]. Active Contour Model (ACM) is considered one of the powerful methods of computer-assisted medical image segmentation. The reason for this method's strength is its high potential to segment and trace the boundaries of the object of interest, as it can regulate the shape of the contour through the restrictions it imposes on it where it uses constraints

(top to bottom) inferred from image information and (bottom to top) for the location, shape and size of this object [3]. The active contour is a thin curve drawn on the image inside or outside the object of interest. This curve is deformed towards the salient object with the help of the forces acting on it, the final shape of the curve is formed by repeating the energy reduction associated with the curve as the shape changes with each iteration until it reaches the final shape where the boundaries of the object are at the lowest energy limit. The curve behaves like a rubber band extending outside of a convex planar region would snap to the boundary of the region when released, as the reduction in energy corresponds to a decrease in the elasticity of the elastic band, making it shrink toward the boundary [4]. The active contour can be divided into two basic approaches: (1) snakes which are represented as explicit parameterized segmentation curves (e.g., parametric). (2) geometric active contours which are represented as level sets based on the implicit representation of curves [5].

562

M. A. Duarte, et al. in 2015 proposed a method based on the methodology of "geodesic active contours" (GAC). Radiologists are actively engaged in the final phase of the process. GAC along with anisotropic texture filtering and also the experience of radiologists. The suggested procedure was tested using 1000 ROIs derived from the DDSM Dataset for segmented images, the quantitative evaluation, based on the area overlap measure AOM, yielded a mean of 0.52 ± 0.20 [6]. F. Liu, et al. in 2015 proposed the stopping function in the classical GAC model based on the SPF model, to avoid boundary leaking, where the local image information was used to refine the rough contour. The model was tested on the digital database for screening mammography DDSM. Area overlay metric AOM indicates to 0.94 [7]. K. L. Kashyap, et al. in 2017 The technique of clustering based on RBF was proposed for interpolation of variational level set function, they used method of Subtraction of original and enhanced inverted images to suppress the normal breast tissue and highlight the suspicious mass and then implement the level set function is performed without re-initialization due to the mesh-free method. The algorithm was tested on a sample from MIAS and DDSM database it was the highest sensitivity of 95.12 [8] D. Saraswathi, et al. in 2017 proposed a technique merge the Chan-Vese method with Fuzzy C means to solve initialization, convergence to image segmentation and noise reduction. MIAS database is used and dice indices were 0.9 [9]. S. Soomro, et al. in 2017 proposed a SPF function is characterized by a phase-shifted Heaviside function, which helps to achieve the optimal solution in fewer iterations. The suggested approach is tested on many mini-MIAS Dataset mammogram images. Sensitivity indices were 0.9865 [10]. A. Niaz, et al. in 2020 proposed an active contour approach based on a reformed mixed local and global fitted function. They worked to capture the ambiguous boundaries of areas of interest by included a p-Laplace term with Chan-Vese to reduce the requirements for reinitialization and represses incorrect segmentation contours. The technique was tested on 25% from the MIAS Database, The result signifies an accuracy of 98.4% [11]. N. Badshah, et al. in 2020 proposed a model that combines Laplacian of Gaussian and level set function based on local intensity information for the image in order to overcome the problem of inhomogeneity MRI and mammogram images. Geometrical points were identified near the tumour boundary prevent the contour from leaking. The algorithm was applied on various mammogram images from real data set collected from different hospitals in Peshawar, Pakistan. Accuracy was 99.91% [12]. M. Malathi, et

al. in 2021 used ACM to detect the abnormal image before passing through the classification stage, where they examined the images in two ways, first through the deep belief network which showed an accuracy of 97.51%. The second method is CNN which gave an accuracy of 95.65% [13].

The early detection of breast cancer is challenging in the medical imaging diagnose because difficult of identifying the boundaries of the tumour. Mammography was used as one of the most effective methods for detection. The main goal of this work is the detection of mass and calcifications in the breast. This work consists of the implementation and modification of active contour technique for detection of tumour boundaries in breast. Wherefore, Gaussian filter was used as a pre-processing as it works to intensify the low-frequency components while reducing the high-frequency components as the breast structure is enhanced and have noise suppression. Then, CLAHE was used to improve the contrast of the image, which increases the contrast between the tumour and the surrounding tissue and sharpens the edges of the tumour. Also, comparison among of parametric active contour segmentation methods that depend on the explicit representation of the segmentation boundary and the geometric segmentation method that relies on the implicit representation of the segmentation boundary is based. The remainder of this paper is organized as follows: The work background and the proposed methodology was described in section 2, and the effect of pre-processing and parameters was described in section 3, while in section 4 the method of obtaining images was clarified and explain characteristics of tumors margins, in section 5 the performance measures are explained, while in section 6 the experience and results were analyzed. Finally, in section 7 the conclusions were presented.

2. Methodology

2.1 Parametric active contour Model / PACM

The contour is a closed curve ($C(s)$) that can be explicitly represented (parametric) in the plane (x, y) as follows:

$$C(s) = \big(x(s), y(s)\big) \qquad (1)$$

s , the normalized arc length. $s \in [0,1]$ is the parametric domain. Contour has total energy ($E(C)$) that is the sum of two internal and external energies that can be expressed [14]:

$$E(C) = E_{INT} + E_{EXT} \qquad (2)$$

Internal Energy Function (E_{INT}) this energy is related to the contour itself, as it controls the length and the degree of smoothness, and thus it is a sum of two terms, as

$$E_{INT} = E_{ELAST} + E_{BIND} \qquad (3)$$

The first term from Eq. (2) represents the elasticity energy (E_{ELAS}) is responsible for controlling the length of the contour, whereby it makes the contour behave as an elastic membrane to resist stretching and thus be able to continue, this term tries to remove the curve concavity and shrink the curve. According to basic mechanics, it is proportional to the square of the first derivative of the contour with respect to s as follows:

$$E_{ELAST} = \frac{1}{2}\int_s \alpha |C'(s)|^2 ds \qquad (4)$$

where the $C'(s)$, indicates the first-order derivative with respect to s, α is the tension coefficient, as it is the measurement factor that controls the contour elasticity energy as it determines the extent to which the contour can extend at points s [15]. The second term of the internal energy equation is E_{BIND} this energy makes the contour behave like a thin plate to resist bending where it is a smooth curve without sharp angles. According to basic mechanics, this energy can be expressed as the square of the second derivative of the contour with respect to s as follows:

$$E_{BIND} = \frac{1}{2}\int_s \beta |C''(s)|^2 ds \qquad (5)$$

where the ($C''(s)$), indicates the second-order derivative with respect to s, β is known as the stiffness factor, and it is the parameter that controls the smoothness of the contour [16], that is, the ability of the contour to bend at points s.
External Energy Function (E_{EXT}) is obtained from the image being segmented and is formulated so that it forces the snake to areas of interest in segmentation, such as lines or edges. It can be expressed as follows:

$$E_{EXT} = \int_s E_{IMAGE}(C(s)) ds \qquad (6)$$

where $E_{IMAGE}(C(s))$ denotes a scalar function defined on the image plane, so the local minimum of $E_{IMAGE}(C(s))$ attracts the contour to edges. It has different formulas, the most famous of which is the image gradient, that is given as follows:

$$E_{IMAGE}(C(s)) = -|\nabla f(x,y)|^2 \qquad (7)$$

$f(x,y)$ is the original image containing the object to be segmented $\nabla f(x,y)$, the gradient image will have high grey levels at pixels located at edges in $f(x,y)$. The negative signal is used to push the contour towards the greatest rate of change (towards the edges), When the negative sign is removed, the contour will point in the opposite direction, that is, away from the edges. Now we can get the final form of the total energy equation for the contour according to the explicit representation of the contour, which we can call the traditional snake as follows [17]:

$$E(C(s)) = \frac{\alpha}{2}\int_s |C'(s)|^2 ds + \frac{\beta}{2}\int_s |C''(s)|^2 ds + \int_s E_{IMAGE}(C(s)) ds \qquad (8)$$

Traditional snakes give an accurate identification of the edges of the object to be segmented only if the contour is drawn near the edges, meaning that we need prior knowledge of the shape of the object because it uses local information for the contour and this is one of the difficult problems in addition to that traditional snakes maintain the same structure during the development of the contour, so it cannot Discover more than one boundary at a time, meaning that it cannot be divided into multiple boundaries or consist of the merging of multiple primary lines [18].

2.2 Geometric active contour model (GACM)

(GACM) differs from (PACM) in that it develops the contour using geometric measurements, GACM is also called the "level set model", level sets in our context are sets of points of a 2-D curve formed by the intersection of a plane and a 3-D surface. Unlike the parametric representation, level sets are based on implicit representations. An important aspect of this approach is that it can adapt to changing topology (e.g., the discovery of holes within a region, and the emergence of new regions) during curve evolution. Because of the continuity of derivative first and second order, parametric curves do not have this power naturally, Osher and Sethian presented the level-set approach in [19]. The level set model was introduced independently by Caselles et al. [20]. and Malladi et al. [21] in the image processing field, based on the level set method (LSM).

With the time, the topology of the curve changes, meaning that the curve may be divided into separate curves, or more than one curve may merge to form one curve. The best way to deal with easily variable topologies is LSM, where the curve is embedded in a surface of higher dimensions [22]. The contour is precisely the intersection between the function of the level set and the plane of x-y, we say that a 2-D curve C is embedded into a 3-D function $Z = \Phi(x, y)$ by letting C be the zero-level set of this function.

In the discussion that follows, level curves will become segmentation boundaries, and the power of this concept is that the level set approach does not require an explicit representation of these boundaries. A set of levels is described as the collection of all (x, y) pairs that have the same z value (Height). The set of zero levels is the set of all the pairs (x, y) satisfying $\Phi(x, y) = 0$. The zero-level set is, in other words, $\{(x, y)|\Phi(x, y) = 0\}$.

The initial contour is the zero-level set at time t = 0. Were x and y being the coordinates on the image grid. The level set function in 3D appears as a surface that works on the development of the initial contour gradually over time, depending on the speed function instead of the direct evolution. In other words, at any time (t) we extract a value (z) for all pair(x, y)s and the contour at that moment it is the zero-level set. This means that the contour develops with the evolution of the surface and the final contour is the zero-level set of the last evolution of the surface. we can also classify geometric active contours in two essential categories "Edge-based Active Contouring Models" (GEACM) and "Region-based Active Contour models". GEACM have some drawbacks. Since it is developing the contour in just one direction, either inside or outside, an initial contour then the primary contour must be placed entirely inside or outside the "Region of interested" (ROI), and thus an amount of prior information must be provided on the image as well. GEACM depend on the action of the image gradient, active edge-based contours may bypass the fuzzy boundaries and are sensitive to local minima or noise [4]. Most "region-based active contour models" consist of two components: the regularity component, which determines the smooth shape of the contours, and the energy minimization component, which searches within a subset for uniformity of the desired feature. A pleasant advantage of active region-based contours is that the initial contours can be put anywhere in the image as region-based segmentation depends on the minimization of global energy instead of the minimization of local energy. It is not required to provide a lot of prior knowledge about the part to be divided. However, the Chan-Vese method is a region-based model applied in "geometric active contour" system. A literature review has found that it does not do well for images with intensity inhomogeneity, despite having a wider convergence spectrum and effectively handling topological changes using the (CVM). The key explanation is that the (CVM) algorithm is built on the premise that image intensities have stayed constant in each region, leading to the incorrect movement of the contour.

2.3 Proposed algorithm

Input: Mammogram image, draw initial contour at which the evolution of the segmentation begins.
Output: Final snake locked on breast tumor in the mammogram image.
Begin
Step1: Load the first image from database (mini MIAS database).
Step2: Input the value of σ, x, y for removing the noise by Gaussian Mask.
Step3: enhancement image contrast which result from step2 by CLAHE technique.
Step4: Choose the segmentation method PACM or GACM (GEACM OR CVM).
Step5: Input parameters the contour.
Step6: Input the No. of iteration.
Step7: Draw initial contour by freehand tool.
Step8: The snake can be formed by following over a certain number of iterations until a statistical equilibrium is reached.
Step9: Save the result image after applying the step8.
End

3. Effect of pre-processing and parameters

3.1 Pre-processing

Using segmentation algorithms directly on the original images produces unwanted effects due to the inherent noise and poor contrast of the mammogram images. The suggested method uses a Gaussian filter, which is a filtering approach based on peak detection. The peak measurement is carried out on the basis that peaks are to pulse. The significant factor is that this filter corrects the spectral coefficient of interest and all amplitude spectrum coefficients within the filter window. It is a linear low pass filter, and greater significance is given to the pixels near the edge, thus reducing-edge blurring. The degree of smoothening is controllable, and computing efficiency is also in this filter [23]. "contrast limited adaptive histogram equalization" (CLAHE) is a technique that enhances the poor contrast problem of digital images, particularly medical images [24]. Instead of the whole image, CLAHE works on a specific area,

called a tile. The contrast of each tile is improved. The histogram of the output area roughly matches the histogram described by the distribution parameter to remove arbitrarily induced borders. The adjacent tiles are then mixed using bilinear interpolation. The contrast can be reduced, especially in homogeneous areas to avoid amplifying any noise that may be existing in the image.

3.2 Parameters

For image preprocessing, Gaussian Sigma (σ) was defined, a greater sigma value produces a smoother image than a smaller σ value. 0.6 is a reasonable estimate for a mammogram. For PACM segmentation, the practice has proven that the optimal value for the parameter α (elasticity force serves to prevent the curve from expanding) for mammogram images is 0.4. The appropriate value for the parameter β (rigidity force prevents the curve from twisting excessively) is 0.2. The parameter κ (applies to the external force), the appropriate value is 0.15. Values higher than this value led to the dominance of the external force because the internal force (elasticity and Rigidity) cannot maintain the smoothness of the contour. For GACM segmentation, after the tests, the weighted length term parameter $\lambda = 1$. The weighted area term parameter v: The values of v are determined by the quality of the image's boundaries and the status of the objects in the image. The appropriate value for the parameter v is 0. The appropriate value for the parameter μ (the internal (penalizing) energy term parameter, and it must be greater than zero) for such type of images is 3.

4. Image acquisition

A mammographic image from the "Mammographic Image Analysis Society" (MIAS) which used in this research to apply the methods suggested. These images have previously been investigated by a specialist in biopsy and radiology. MIAS is a dataset reference. It is chosen because of the different cases it represents. It is also used extensively in related research ventures. The dataset comprises of 322 left and right mammograms, out of 161 women, Patients aged 50 to 65 years of age. In 51 of which malignancy, 64 benign, 207 normal diagnostics have been collected. The abnormalities are known as microcalcifications (CALC), spiculated masses (SPIC), ill-defined masses (MISC), circumscribed masses (CIRC), architectural distortion (ARCH), and asymmetry (ASYM) [25].

5. Performance measures

There are many evaluation criteria for comparing the proposed model with a segmentation technique. These criteria determine the performance of the proposed model by selecting the compatibility between the Grand Truth (R_{GT}), which represents the region manually segmented by an expert radiologist, and the region resulting from the application of the proposed model (R_{SEG}).

The Jaccard similarity (J_S) defines as the area overlap measure between (R_{GT}) and (R_{SEG}). While Dice coefficients (D_S) is calculated the degree to which two binary images have spatial overlap. The (J_S) and (D_S) values are between 0 (no overlap) and 1 (perfect agreement). The values closer to 1 show an accurate result of segmentation, and the values closer to 0 imply poor segmentation. Precision (PRES) is the pixels number of correctly identified which not belonging to the tumour divided by all pixels. While Sensitivity (SENS) is the pixels number of correctly identified belonging to the tumour divided by the number of all pixels. Also, PF-Score (PFS) is the harmonic mean of precision and Sensitivity [26].

The quantitative results give the extent of congruence and difference between R_{GT} and R_{SEG} in terms of the number of correctly and wrongly diagnosed pixels, but it does not give us an idea of the similarity in terms of shape, so we will use the region's properties to provide us more about the shape of the segmentation results. To describe the congruence and the difference in terms of shape between R_{GT} and R_{SEG}, we will use the following area properties (Distance, Area, Diameter, Axis). Distance is the distance between the center R_{GT} and R_{SEG}, the smaller the distance, the higher the accuracy of the measurement. The area is representing the area difference between the two regions divided by area of R_{GT}. The smaller is the value, the higher is the congruence; Diameter, the ratio of between a diameter of R_{SEG} to diameter of R_{GT} which is a value with three possibilities: 1 perfect match, < 1 this means that $R_{SEG} > R_{GT}$, > 1 means that $R_{SEG} < R_{GT}$. Axis the difference between ratio of major axis length of R_{SEG} to minor axis length of R_{SEG} and ratio of major axis length of R_{GT} to minor axis length of R_{GT}, the congruence is higher when the result approaches zero, i.e., the shape approaches the ball. In all the images included in the study, the colors of the external borders will be blue, green, and red for (ROI , R_{GT}, R_{SEG}) respectively [27].

6. Results and discussion

The PACM is weak in identifying fine details even after the image has gone through the pre-processing step because it depends on the explicit representation of the edges and that most of the tumors have weak or hidden borders by the surrounding tissues of the breast, in addition to the fact that the tumors have margins of different shapes, so it is difficult for PACM to deal with them, as it cannot respond to the topological changes that characterize this type of margins, since it has the property of continuity due to the presence of derivatives within some terms of the Eq. (8). Also, it maintaining the shape of the initial contour due to the constraints of regularity which cannot handle structure changes, so it needs to prepare an ROI close to the shape and boundaries of the tumor. The

segmentation criteria scored low results, as the best outcome was for tumor MISC $J_S = 48$ and $D_S = 65$. The rest of the results were listed in the Table 1. This was visually confirmed in the third row from the Fig. 1. Iteration was fixed at 400 for all tumors until the segmentation contour takes sufficient time by deformation to reach the final segmentation contour. Additionally, no change was observed when the iteration was increased beyond this limit.

GEACM was able to identify large tumors with clear boundaries with a great degree of accuracy, as the best result was obtained when the tumor was (CIRC) with results $J_S = 88\%$ and $D_S = 94\%$ for being one of the tumors with relatively clear boundaries, this is due to the fact that GEACM is an edge-based segmentation method. As for tumors with weak margins and different topologies, it was able to
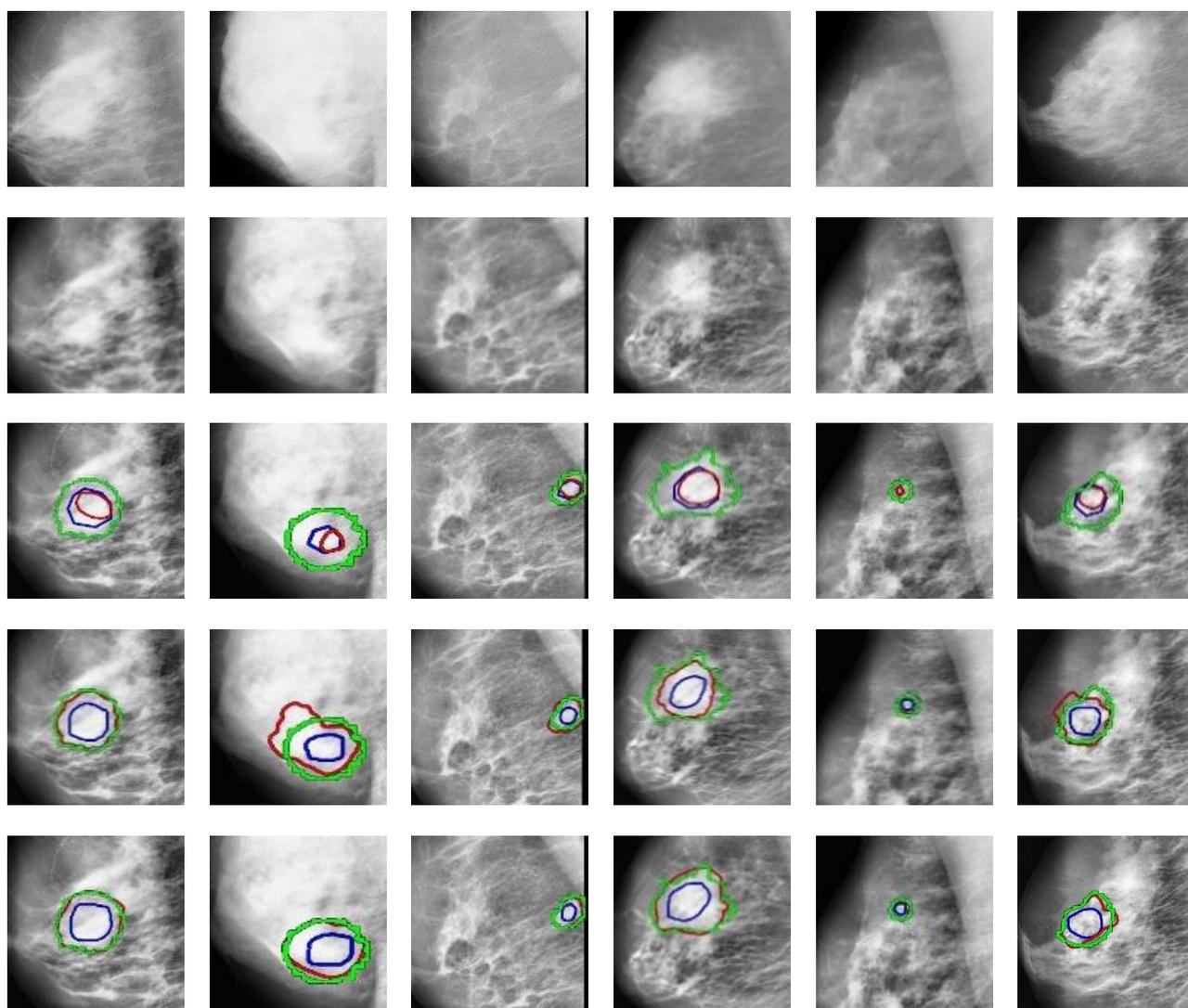
Figure. 1 Segmentation result of the images, first row returns to the original images, second row returns to the images after Pre-processing, third row returns to the final result by PACM, fourth row returns to the final result by GEACM, last row returns to the result by CVM.

Table 1. Measurements (J_S, D_S, PF_S, PRE_S, SEN_S, and Iteration) for PACM segmentation

PACM						
Similarity	CIRC	SPIC	MISC	ASYM	CALC	ARCH
J_S %	10	12	48	31	25	24
D_S %	17	22	65	47	40	39
PF_S %	4	17	66	20	44	21
PRE_S %	7	20	45	35	42	9
SEN_S %	5	18	53	26	43	13
Iteration	400					

Table 2. Measurements (J_S, D_S, PF_S, PRE_S, SEN_S, and Iteration) for GEACM segmentation

GEACM						
Similarity	CIRC	SPIC	MISC	ASYM	CALC	ARCH
J_S %	88	60	68	68	73	66
D_S %	94	75	81	81	85	79
PF_S %	98	49	74	60	95	53
PRE_S %	98	63	79	39	97	57
SEN_S %	98	56	76	47	96	55
Iteration	500	700	450	700	200	700

Table 3. Measurements (J_S, D_S, PF_S, PRE_S, SEN_S, and Iteration) for CVM segmentation

CVM						
Similarity	CIRC	SPIC	MISC	ASYM	CALC	ARCH
J_S %	90	84	84	82	78	78
D_S %	95	91	91	90	88	88
PF_S %	98	95	97	75	94	79
PRE_S %	97	98	98	63	98	86
SEN_S %	97	96	98	69	96	83
Iteration	260	190	190	440	110	320

determine them to an acceptable degree, because it is a segmentation method that depends on the implicit representation based on the level set, as it has the advantage of changing the initial contour with changing the topology of the boundaries of the tumor. The leakage that occurred in the contour of some tumors where a number of false positives were recorded came as a result of the algorithm's inability to distinguish weak boundaries of tumors, where the weakest result was when the tumor (SPIC) as follows $J_S = 60\%$ and $D_S = 75\%$. All results are detailed in Table 2 and the GEACM segmentation contour can be traced in the fourth line of Fig. 1.

The CVM outperformed both PACM and GEACM in capturing the boundaries of all types of tumors in the dataset, whether with clear or weak margins because it is a segmentation method based on region and not boundaries as in the previous two algorithms, that is, it is able to collect tumor pixels according to the characteristics of similarity between them without dependence on edges with the ability to define edges.

CVM shares with GEACM the ability to track tumor topology because it relies on implicit representation based on the level set. CVM differs from the previous two algorithms by not having to configure a good ROI because it is looking to reduce the global energy of the image instead of the local. The best segmentation result for the tumor was CIRC, with results $J_S = 90\%$ and $D_S = 95\%$, and the weakest was for CALC and ARCH with $J_S = 78\%$ and $D_S =$

88%. The rest of the results are detailed in the Table 3, and the results were visually confirmed in the fifth line of the Fig. 1.

Tables 2 and 3 note that the number of iterations of GEACM is greater than that of CVM, this is due to the GEACM have a bias in one direction only either inside or outside the region of interest. So, it is limited in movement and needs more iterations to reach the segmentation required. The segmentation method is very slow and needs more time. while, CVM has a bi-directional bias, which means the initial contour can deform both inside and outside the region of interest at the same time, reducing the number of iterations and thus reducing the time.

Furthermore, from discussing and interpreting the images visually, and statistical results assessment for using active contour algorithms. There is confusion in the visual and statistical evaluation of the results. May be led to an inappropriate medical decision. Wherefore, statistical properties are proposed based on the calculation of the values of the area properties of the images. These characteristics (Distance, Area, Diameter, Axis) aid in identifying breast tumors and decision making with the highest accuracy.

From Fig. 2 it can be seen the performance of the CVM & PACM algorithms is identical in CIRC & MISC tumors. So, the two algorithms can identify the affected region with high accuracy. Despite the
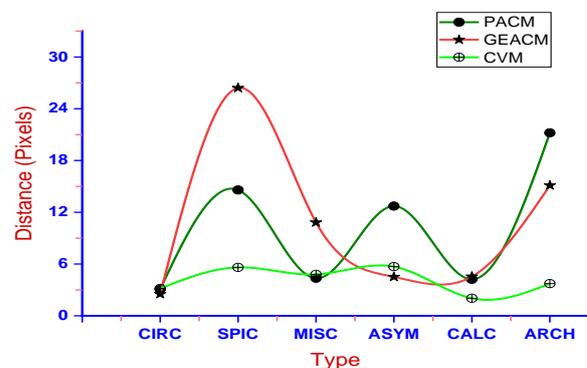
Figure. 2 Relationship between Distance and tumor type for the three ACM algorithms
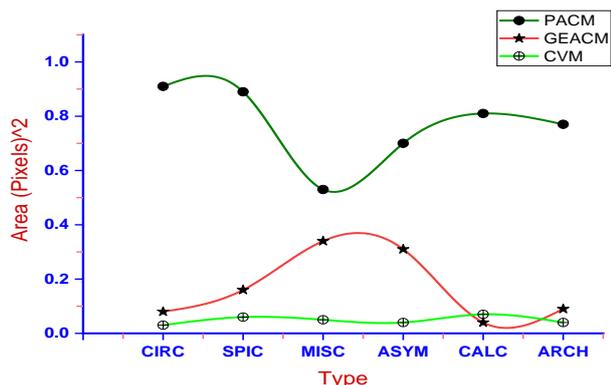
Figure. 3 Relationship between Area and tumor type for the three ACM algorithms
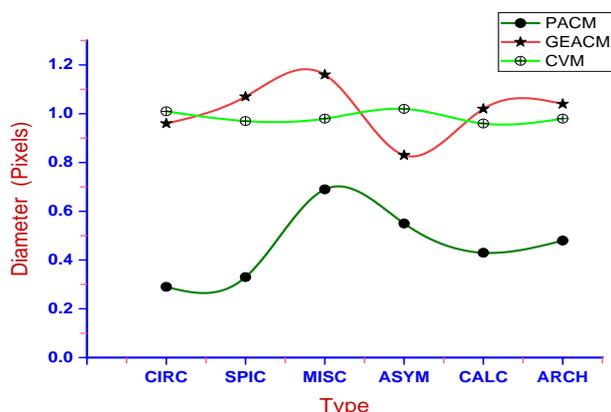
Figure. 4 Relationship between Diameter and tumor type for the three ACM algorithms
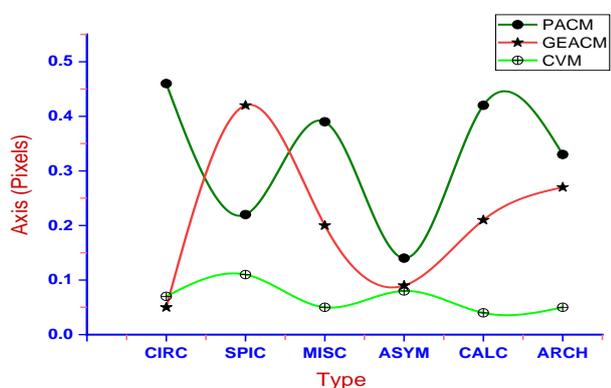
Figure. 5 Relationship between Axis and tumor type for the three ACM algorithms

superiority of the CVM algorithm in the rest of the tumors.

In the area feature, it can be noted the superiority of the GEACM and CVM algorithms in the CALC tumor as shown in Fig. 3. This is due to the calcifications extend over a larger area and GEACM and CVM has been able to determine the boundaries of the area better than PACM. There is a high close between GEACM & CVM in the CIRC & ARCH areas.

Table 4. Comparison with other works on MIAS dataset

Authors & Year.	Methods Used	J_S	D_S	PF_S	PRE_S	SEN_S
Kashyap et al, 2017[8]	level set	87	89	—	—	95.12
Hmida et al, 2018[28]	Fuzzy contours	—		—	88.08	91.12
Azlan et al, 2020[29]		—		94.34	94.66	93.94
Proposed Method	PACM	48	65	66	45	53
	GEACM	88	94	**98**	**98**	98
	CVM	**90**	**95**	**98**	97	97

Fig. 4 illustrates the optimal CVM method for all tumors with close to GEACM results in CIRC, CALC & ARCH tumors in determining the diameter of abnormal regions.

From Fig. 5 it can be seen the GEACM algorithm outperforms CVM algorithm in CIRC tumor. There is great convergence between them in ASYM tumor through the identification of tumor axes.

There is great difficulty in comparing current work with other techniques. The authors use a different number of mammogram samples from different datasets, public and private, and applying measurement criteria that differ from one study to another. Comparison is made with studies that used the MIAS dataset, as shown in Table 4.

We compared the proposed algorithm with the work of K. L. Kashyap et al.[8], developed a level set function for segmentation of mammogram images using a Mesh-free based radial basis function (RBF) collocation approach. A comparison is also made with the algorithm of M. Hmida et al. [28], where the algorithm was a hybrid between CVM and fuzzy c-means, they dealt with 57 mammogram images of the class of masses only. N. A. N. Azlan et al. [29], used normalization and filtration as a pre-processing before segmenting the image by the active contour method. The algorithms were implemented on all dataset images. Our algorithm with the GACM approach showed superiority over all the previously mentioned algorithms with CVM progression on GEACM according to the statistical results. In addition to our algorithm being unique to successfully deal with all tumours in the dataset, we took an example of difficult segmented tumours calcifications. Whereas, in the previous studies ACMs did not deal with all dataset images if this type of tumour was not segmented or the segmentation results were weaker with this type of tumour. This is due to the appropriate combination of pre-processing methods and the appropriate selection of parameters.

569

7. Conclusions

In this work, the tumor segmentation challenge in mammograms was developed by applying ACM algorithms using both PACM and GACM approaches. Pre-Processing with a mixture of Gaussian filter and CLAHE is important which has proven to work well. The PACM algorithm was not able to efficiently in mammograms. So, it miscalculated boundary locations in most tumors according to segmentation results criteria. The GACM algorithm was able to record efficiency with mammogram tumors as indicated by the segmentation results. As it has the advantage of dealing with different topologies of tumors. But it needs a large number of iterations, because its bias is limited in one direction only, either inside or outside the region of interest. The CVM algorithm outperformed all previous algorithms, as the results were perfect in determining the boundaries and location of all types of breast tumors. The statistical results returned on the highest segmented image are as follows Jaccard=90 and Dace=95. The main contribution of this works was the ability to segment very small tumors with high accuracy. Although the proposed method yields strong and promising results, there are still some aspects that could be improved upon. The initial contour is determined manually in most studies.

In future works, we suggest determining the initial contour automatically based on the size and location of the tumor image. And some parameters must be set experimentally when performing ACM segmentation by image data. Furthermore, we suggest determining the parameters automatically in a certain range taking into account the statistical information of the image.

Conflicts of interest

The authors declare no conflict of interest.

Author contributions

Methodology, formal analysis, data curation, writing—original draft preparation, and editing, Eman A. Radhi; supervision, software, validation, Mohammed Y. Kamil; all authors read and approved the final manuscript.

References

[1] M. Lotfollahi, M. Gity, J. Y. Ye, and A. M. Far, "Segmentation of breast ultrasound images based on active contours using neutrosophic theory", *Journal of Medical Ultrasonics*, Vol. 45, No. 2, pp. 205-212, 2018.

[2] M. Y. Kamil and A. M. Salih, "Mammography Images Segmentation via Fuzzy C-mean and K-mean", *International Journal of Intelligent Engineering Systems*, Vol. 12, No. 1, pp. 22-29, 2019.

[3] T. McInerney and D. Terzopoulos, "Deformable models in medical image analysis: a survey," *Medical image analysis*, Vol. 1, No. 2, pp. 91-108, 1996.

[4] P. P. Rebouças Filho, P. C. Cortez, A. C. da Silva Barros, and V. H. C. De Albuquerque, "Novel adaptive balloon active contour method based on internal force for image segmentation–a systematic evaluation on synthetic and real images", *Expert Systems with Applications*, Vol. 41, No. 17, pp. 7707-7721, 2014.

[5] X. Yang, X. Gao, D. Tao, X. Li, and J. Li, "An efficient MRF embedded level set method for image segmentation", *IEEE transactions on image processing*, Vol. 24, No. 1, pp. 9-21, 2014.

[6] M. A. Duarte, A. V. Alvarenga, C. M. Azevedo, M. J. G. Calas, A. F. Infantosi, and W. C. Pereira, "Evaluating geodesic active contours in microcalcifications segmentation on mammograms", *Computer methods programs in biomedicine*, Vol. 122, No. 3, pp. 304-315, 2015.

[7] F. Liu, Z. Gong, Y. Chen, and Y. Gu, "Segmentation of mass in mammograms by a novel integrated active contour method", *International Journal of Computational Science Engineering*, Vol. 11, No. 2, pp. 207-215, 2015.

[8] K. L. Kashyap, M. K. Bajpai, and P. Khanna, "Globally supported radial basis function based collocation method for evolution of level set in mass segmentation using mammograms", *Computers in biology medicine*, Vol. 87, pp. 22-37, 2017.

[9] D. Saraswathi, E. Srinivasan, P. Ranjitha, and Technology, "An Efficient Level Set Mammographic Image Segmentation using Fuzzy C Means Clustering", *Asian Journal of Applied Science*, Vol. 1, No. 4, pp. 7-11, 2017.

[10] S. Soomro and K. N. Choi, "Robust active contours for mammogram image segmentation", In: *Proc. of 2017 IEEE International Conference on Image Processing (ICIP)*, pp. 2149-2153, 2017, doi: 10.1109/ICIP.2017.8296662.

[11] A. Niaz, A. A. Memon, K. Rana, A. Joshi, S. Soomro, J. S. Kang, and K. N. Choi , "Inhomogeneous Image Segmentation Using Hybrid Active Contours Model With Application to Breast Tumor Detection", *IEEE Access*, Vol. 8, pp. 186851-186861, 2020.

[12] N. Badshah, H. Rabbani, and H. Atta, "On local active contour model for automatic detection of

tumor in MRI and mammogram images", *Biomedical Signal Processing Control*, Vol. 60, p. 101993, 2020.

[13] M. Malathi, P. Sinthia, F. Farzana, and G. Aloy Anuja Mary, "Breast cancer detection using active contour and classification by deep belief network", *Materials Today: Proceedings*, Vol. 45, pp. 2721-2724, 2021.

[14] P. Moallem, H. Tahvilian, and S. A. Monadjemi, "Parametric active contour model using Gabor balloon energy for texture segmentation", *Signal, Image and Video Processing*, Vol. 10, No. 2, pp. 351-358, 2016, doi: 10.1007/s11760-015-0748-6.

[15] M. Guo, M. Dong, Z. Wang, Y. Ma, and Y. Guo, "A new method for mammographic mass segmentation based on parametric active contour model", In: *Proc. of 2015 International Conference on Wavelet Analysis and Pattern Recognition (ICWAPR)*, pp. 27-33, 2015.

[16] P. Kumar, P. Lewis, and T. McCarthy, "The potential of active contour models in extracting road edges from mobile laser scanning data", *Infrastructures*, Vol. 2, No. 3, p. 9, 2017.

[17] Z. Zhang, C. Duan, T. Lin, S. Zhou, Y. Wang, and X. Gao, "GVFOM: a novel external force for active contour based image segmentation", *Information Sciences*, Vol. 506, pp. 1-18, 2020.

[18] N. P. Tiilikainen, *A comparative study of active contour snakes*, Copenhagen University, Denmark, pp. 21-26, 2007.

[19] J. A. Sethian and S. Osher, "Fronts propagating with curvature dependent speed: Algorithms based on Hamilton-Jacobi formulations," *Journal of Computational Physics*, Vol. 79, No. 1, pp. 12-49, 1988.

[20] V. Caselles, F. Catté, T. Coll, and F. Dibos, "A geometric model for active contours in image processing", *Numerische Mathematik*, Vol. 66, No. 1, pp. 1-31, 1993.

[21] R. Malladi, J. A. Sethian, and B. C. Vemuri, "Shape modeling with front propagation: A level set approach", *IEEE Transactions on Pattern Analysis Machine Intelligence*, Vol. 17, No. 2, pp. 158-175, 1995.

[22] X. Jiang, R. Zhang, and S. Nie, "Image segmentation based on level set method", *Physics Procedia*, Vol. 33, pp. 840-845, 2012.

[23] A. M. Joseph, M. G. John, and A. S. Dhas, "Mammogram image denoising filters: A comparative study", In: *Proc. of 2017 Conference on Emerging Devices and Smart Systems (ICEDSS)*, pp. 184-189, 2017.

[24] A. H. Farhan and M. Y. Kamil, "Texture Analysis of Breast Cancer via LBP, HOG, and GLCM techniques", *IOP Conference Series: Materials Science and Engineering*, Vol. 928, No. 7, p. 072098, 2020.

[25] A. M. Omer, *Detection of breast cancer in mammogram images using texture analysis methods*, Sudan University of Science and Technology, 2019.

[26] M. Y. Kamil, "A deep learning framework to detect Covid-19 disease via chest X-ray and CT scan images", *International Journal of Electrical & Computer Engineering*, Vol. 11, No. 1, 2021.

[27] M. Y. Kamil, "Computer-aided diagnosis system for breast cancer based on the Gabor filter technique", *International Journal of Electrical and Computer Engineering (IJECE)*, Vol. 10, No. 5, pp. 5235-5242, 2020.

[28] M. Hmida, K. Hamrouni, B. Solaiman, and S. Boussetta, "Mammographic mass segmentation using fuzzy contours", *Computer Methods Programs in Biomedicine*, Vol. 164, pp. 131-142, 2018.

[29] N. A. N. Azlan, C. K. Lu, I. Elamvazuthi, and T. B. Tang, "Automatic Detection of Masses From Mammographic Images via Artificial Intelligence Techniques", *IEEE Sensors Journal*, Vol. 20, No. 21, pp. 13094-13102, 2020, doi: 10.1109/JSEN.2020.3002559.