CHAPTER 5

MAMMOGRAM MASS SEGMENTATION USING KERNEL BASED FUZZY LEVEL SET

5.1 INTRODUCTION

Computerized mammogram mass segmentation is a challenging problem, because the masses are often poor in contrast, highly connected to the surrounding tissues and possessing ill-defined boundaries. As the region-based segmentation suffers with the problem of boundary leakage when the boundary is blurred and sensitivity to seed position, there is a need for improved segmentation method that can handle this situation and provide a precise boundary of the mass. Based on this motivation, this chapter proposes an automatic mass segmentation technique using kernel based fuzzy level set (KFLS) method that utilize the advantages of fuzzy clustering and level set methods in image segmentation.

Several works used active contour (AC) models for the segmentation of mass boundary (Yuan et al. 2007; Shi et al. 2008; Dubey et al. 2010; Liu et al. 2011; Pang et al. 2012; Xin et al. 2012; Huang et al. 2015). The introduction of active contour model (Osher and Fedkiw 2003) using level set (LS) method has shown proven results in medical image segmentation that uses dynamic implicit interfaces and partial differential equations (PDEs). These models demanded level set function to be re-initialized periodically. There might be some irregularities occurred during the curve evolution process. Many researchers focused on re-initializing the regularity of the level set function (Sethian 1999; Weber et al.
2004). However, this moved the zero level set away from the estimated position. Another limitation was that the conventional LS are computationally intensive.

Li et al. (2005) introduced a fast LS function without re-initialization. This started from an initial contour and reduced the computational expensive for re-initialization. Either manual selection, morphological processing or intensity thresholding were utilized for the evolution of LS function initialization. But this type of initialization was not a reliable choice for an optimum level set segmentation. However, the level set segmentation could be utilized with the initial contour obtained from the fuzzy clustering algorithm in the field of medical image segmentation (Suri 2001; Suri et al. 2002; Li et al. 2009; Lie et al. 2011; Alipour and Shanbehzadeh 2014; Gupta et al. 2015).

The fully automatic segmentation method requires the identification of ROI followed by the segmentation of the mass boundaries. The works referred in Yuan et al. (2007); Shi et al. (2008); Elter et al. (2010); Liu et al. (2011); Pang et al. (2012); Xin et al. (2012) and Berber et al. (2013) showed that the breast region with the suspicious lesions is manually cropped and the segmentation algorithm is applied only to that ROI. The proposed method performs an automatic mammogram mass segmentation using a KFLS method. To eliminate the noises and to extract the breast profile, initially a preprocessing algorithm is applied. Subsequently, a kernel based fuzzy c-means (KFCM) clustering is applied to partition the preprocessed mammogram image into different sub-regions by considering the anatomical structure of the mammogram image. Then the sub-region with the mass is automatically identified as a ROI based on certain criteria. Finally, the initial contour obtained in KFCM is used as the input to the level set segmentation that refines the mass boundary. The LS evolution attracts the contour toward the mass boundaries that eliminates the necessity of re-initialization process. With the output of kernel based fuzzy clustering, the controlling parameters needed to regularize the level set evolution are also derived.
Figure 5.1 Proposed KFLS based CAD system for segmenting suspicious mass lesions in mammogram
The remainder of the paper is organized as follows. Section 5.2 describes the detailed methodology of the proposed mass segmentation algorithm. Section 5.3 reports the experimental results and discussion. Section 5.4 contains the concluding remarks.

5.2 METHODOLOGY

The main objective of the proposed system is to assist the radiologist in detecting and segmenting the mammogram abnormalities which in turn improves the diagnosis accuracy. The segmentation of these abnormalities is needed because certain geometric and texture features are extracted from these regions for further analysis in mammogram masses.

The proposed method performs an automatic segmentation of suspicious mass region in a five step process. The process begins with the noise removal using median filtering, followed by breast profile identification using thresholding and morphological operations. The pectoral muscles present in the mammogram is removed with the single seeded region growing method. A kernel based fuzzy clustering is used to identify the ROI and finally the mass boundary refinement is done using level set method. Figure 5.1 shows the proposed approach for breast cancer segmentation in mammograms.

The proposed methodology is tested on the images from the mini-MIAS and DDSM database. The distribution of input images used in this experiment is already discussed in Section 4.2.1 of Chapter 4.

5.2.1 Mammogram preprocessing

Mammogram preprocessing is the process of simplifying the recognition of abnormalities without leaving the important information. Mammograms may contain noises during acquisition. Also certain portion of breast regions in the
mammograms are superimposed over the background structures which are not necessary for the analysis. The structure of the different regions (Wirth et al. 2005) in the mammogram image is given in Figure 5.2.

![Figure 5.2 Anatomy of the breast region](image)

Figure 5.3 Steps in preprocessing stage a) Original image of C_0249 from DDSM database b) Median filtered image c) Binary thresholded image d) Breast profile identified from thresholded image e) Artifacts suppressed and background separated image f) Pectoral removed image
The initial preprocessing is performed on the input mammogram to remove the noises, to extract the breast profile and to remove the pectoral muscles. The artifacts, labels and markers present in the mammogram also removed during the preprocessing. The preprocessing stages carried out in this work are explained in Section 3.2.2 of Chapter 3. The steps involved in the preprocessing are shown in Figure 5.3.

5.2.2 Kernel based fuzzy clustering

Fuzzy clustering allows the data to be clustered in two or more clusters. This is frequently used in pattern recognition and widely applied in medical image processing. Fuzzy c-means (FCM) is one of the most popular fuzzy clustering methods used in image segmentation due to its robust characteristics for ambiguity (Wang and Bu, 2010).

K-means algorithm is the base of the standard FCM. The main difference between K-means and FCM is that every object is limited to only one cluster in K-means, but in case of FCM, the membership function \( \mu_{mn} \) indicates the degree of membership of the \( n^{th} \) subject to the \( m^{th} \) cluster. The algorithm attempts to minimize the fitness or objective function \( J \) defined as follows.

\[
J = \sum_{n=1}^{N} \sum_{m=1}^{C} \mu_{mn} \| i_n - v_m \|^2
\]

where \( N \) represents the number of objects to be labelled. (i.e. number of pixels in the image \( N = N_x \times N_y \).) \( C \) represents the number of clusters. \( v_m \) is the centroid of the \( m^{th} \) cluster. \( V = \{ v_1, v_2, \ldots, v_c \} \) is the set of clustering centers. \( l \) represents the fuzziness controlling parameters \( (l > 1) \). \( i_n \) represents the specific image pixel. \( \| \| \) denotes the Euclidean distance metric.
The membership function $\mu_{mn}$ represents the probability that a pixel belongs to a specific cluster. The membership function has the following properties.

$$\sum_{m=1}^{C} \mu_{mn} = 1 \quad 0 \leq \mu_{mn} \leq 1 \quad \sum_{n=1}^{N} \mu_{mn} > 0$$

The objective function ($J$) is optimized by iteratively updating $\mu_{mn}$ and $v_{m}$.

$$\mu_{mn} = \frac{\sum_{k=1}^{C} \| n - v_{m} \|^{-2/(l-1)}}{\sum_{k=1}^{C} \| n - v_{k} \|^{-2/(l-1)}}$$

Eq. (5.2)

$$v_{m} = \frac{\sum_{n=1}^{N} \mu_{mn}^{l} i_{n}}{\sum_{n=1}^{N} \mu_{mn}^{l}}$$

Eq. (5.3)

The standard FCM algorithm is defined as follows:

Step 1: Set values for $l, C$.

Step 2: Initialize the fuzzy membership matrix $\mu_{mn}^{l}$.

Step 3: Calculate the cluster centre using Eq. (5.3).

Step 4: Update the membership matrix $\mu_{mn}^{l}$ using Eq. (5.2).

Step 5: Repeat step 3 and 4 until the termination criteria is satisfied.

$$|v_{new} - v_{old}| < \varepsilon \text{ or maximum number of iteration is reached. Then the final clustering is obtained.}$$

The standard FCM is robust to noise and outlier that is reflected in segmentation accuracy, as it does not include the spatial information. The limitation of the standard FCM is improved with the spatial FCM where the spatial information is included in the objective function. This clustering is suffered with more computation time because in each step the objective function is calculated by considering the neighbour term or mean or median of the image.
Zhang and Chen (2004) proposed a kernel based FCM. The kernel based FCM provides strong noise robustness for image segmentation. The objective function for KFCM is given as:

\[ J = \sum_{n=1}^{N} \sum_{m=1}^{C} \mu_{mn} \| \Phi(i_n) - \Phi(v_m) \|^2 \]  

Eq. (5.4)

where \( \Phi \) is an implicit nonlinear map and other components are same as in Eq. (5.1). Here the Euclidean distance function is replaced with kernel induced distance function. In feature space, a kernel can be a function which is denoted as \( K \), where \( K(x, y) = \langle \Phi(x), \Phi(y) \rangle \).

The most popular kernel function used is Gaussian radial basis function (RBF) kernel

\[ K(x, y) = \exp\left( -\frac{\|x - y\|^2}{\sigma^2} \right) \]

where \( \sigma \) is the width parameter, then the objective function is

\[ J = \sum_{n=1}^{N} \sum_{m=1}^{C} \mu_{mn} \left[ 1 - K(i_n, v_m) \right] \]  

Eq. (5.5)

The other kernel functions (Muller et al. 2001) are polynomial and sigmoid functions, where the equations are little more complex. The Gaussian RBF kernel function is adopted in all applications due to its simplicity. The evolution of objective function over number of iterations using FCM and KFCM is shown in Figure 5.4.

The Figure 5.4 shows that the number of iterations needed to find the optimal cluster center is less in KFCM when compared to FCM. The objective function value is relatively less compared to FCM. The maximum objective function value in FCM is \( 4.5 \times 10^8 \) and the maximum objective function value in KFCM is \( 1.7 \times 10^4 \). Since there is a larger variation in objective function value when compared to FCM, the values are depicted in two different graphs.
5.2.3 Kernel based fuzzy level set (KFLS) algorithm

Currently, lot of research work on image segmentation is concentrated on the AC model. There are several enviable advantages of AC model over classical image segmentation methods, such as edge detection, thresholding and region growing. The AC model achieves sub-pixel accuracy of object boundaries under energy minimization framework. It allows incorporation of prior knowledge
such as shape and intensity distribution for robust image segmentation. It also
provides smooth and closed contours as segmentation results, which are necessary
for further applications, such as shape analysis and recognition.

The level set method innovated by Osher and Sethian (1988) is used to
capture the moving fronts in the image. The contours or surfaces are represented as
the zero level set of a higher dimensional function, usually called a level set
function. The level set method became well known and had far-reaching impact in
various applications, such as computational geometry, fluid dynamics, image
processing, and computer vision. For image processing and computer vision
applications, the level set method has been introduced independently by Caselles et
al. (1993) and Malladi et al. (1995) in the context of active contour (or snake)
models (Kass et al. 1987) for image segmentation. The active contour models are
classified as either parametric or geometric according to their representation and
implementation. The parametric active contours are represented explicitly as
parameterized curves in a Lagrangian framework, while the geometric active
contours are represented implicitly as level sets of a two-dimensional function that
evolves in an Eulerian framework. The evolution of the level set function \( \phi \) for
geometric active contour model is given by

\[
\frac{\partial \phi}{\partial t} = F|\nabla \phi|, \quad \phi^0(x, y) = \phi(0, x, y)
\]

Eq. (5.6)

where \( F = \text{div}(\nabla \phi/|\nabla \phi|) \) represents the speed function that controls the
active contour motion, \( |\nabla \phi| \) represents the normal direction and \( \phi^0 \) represents the
initial contour. A desirable advantage of level set methods is that they can represent
contours of complex topology and are able to handle topological changes, such as
splitting and merging, in a natural and efficient way, which is not allowed in
parametric active contour models.
Level set function may develop some irregularities during its evolution in conventional LS, which cause numerical errors and eventually destroy the stability of LS evolution. To overcome this difficulty, level set function needs to be re-initialized to restore the regularity and maintain stable level set evolution. Re-initialization is performed periodically as a signed distance function by stopping the evolution and reshaping the degraded level set function. Although re-initialization as a numerical remedy is able to maintain the regularity of the LS function, it may incorrectly move the zero level set away from the expected position. In order to solve the problems associated with re-initialization, it is necessary to pursue a level set method that does not require re-initialization.

Li et al. (2010) presented a new variational formulation that forces the level set function to be close to a signed distance function, and therefore completely eliminates the need of the costly re-initialization procedure. This variational energy function consists of an internal energy term and an external energy term, respectively. The internal energy term penalizes the deviation of the level set function from a signed distance function, whereas the external energy term drives the motion of the zero level set to the desired image features such as object boundaries. The result of the level set evolution function is the gradient flow that minimizes the overall energy functional. Due to the internal energy, the level set function is naturally and automatically kept as an approximate signed distance function during the evolution.

The performance of LS method depends on proper initialization of contour and selection of appropriate parameters for LS function evolution. Different results may be obtained, due to the improper contour initialization. The LS method introduced by Li et al. (2010) not only eliminates the need for re-initialization, but also allows the use of more general functions as the initial LS functions. The initial contour region can sometimes be obtained by a simple and efficient preliminary segmentation step, such as thresholding, such that the initial
region is close to the region to be segmented. However, the level set segmentation could be utilized with the initial contour obtained from other clustering algorithms. Thus the proposed region-based initialization of level set function, which is not only computationally more efficient than computing signed distance function, but also allows for more flexible applications.

Hence, the proposed kernel based fuzzy level set algorithm initializes the mass boundary obtained from KFCM followed by morphological opening to initiate the LS function evolution. The resulted KFCM is utilized to initiate the level set, to estimate the controlling parameters and also to regularize LS evolution. The approach adopted for the proposed methodology allows the initial contour to move towards the object boundaries that removes the manual initialization and the re-initialization process of LS function.

After KFCM clustering algorithm, a defuzzification process is performed to convert the fuzzy partition matrix \((U \_ MF)\) to a crisp partition. The maximum membership procedure is the most important method to defuzzify the partition matrix \((U \_ MF)\). This procedure assigns the pixel \(i\) to the class \(C\) with the highest membership by

\[
C_{ki} = \{\text{arg}_k(\max(U \_ MF_{ki})) | k = 1,2,\ldots,c\} \tag{5.7}
\]

The component of interest \((R_k)\) in KFCM is selected based on the cluster index which contains ROI with maximum mode value.

\[
R_k = \{\text{arg}_k(\max(MODE(C_k))) | k = 1,2,\ldots,c\} \tag{5.8}
\]

A binary image \(B_k\) is obtained from \(R_k\) by

\[
B_k = R_k \geq b_o \tag{5.9}
\]

where \(b_o \in (0,1)\) is an adjustable threshold. Then initialize the level set function \(\phi^0\) as
\[ \phi^0(x, y) = -4e(0.5 - B_k) \]  
Eq. (5.10)

where \( e \) is a constant to adjust the Dirac function. Dirac function is defined as

\[ \delta_e(x) = \begin{cases} 
0, & |x| > e \\
\frac{1}{2e} \left[ 1 + \cos \left( \frac{\pi x}{e} \right) \right], & |x| \leq e 
\end{cases} \]  
Eq. (5.11)

The fast level set formulation (Lie et al. 2011) is defined as

\[ \frac{d\phi}{dt} = \mu \zeta(\phi) + \xi(g, \phi) \]  
Eq. (5.12)

where the first term \( \zeta(\phi) \) defines the penalty momentum or distance regularize term of \( \phi \) deviating from the signed distance function.

\[ \zeta(\phi) = \Delta \phi - \text{div} \left( \frac{\nabla \phi}{|\nabla \phi|} \right) \]  
Eq. (5.13)

The second term \( \xi(g, \phi) \) incorporates image gradient information responsible of driving the zero level curves towards the object boundaries by

\[ \xi(g, \phi) = \lambda \delta(\phi) \text{div} \left( g \frac{\nabla \phi}{|\nabla \phi|} \right) + v g \delta(\phi) \]  
Eq. (5.14)

where \( \delta(\phi) \) denotes the Dirac function which is defined in Eq. (5.11). The terms \( \mu \) (weighting coefficient of the penalty term \( \zeta(\phi) \)), \( \lambda \) (coefficient of the contour length for smoothness regulation) and \( v \) (artificial balloon force) controls the contribution of these terms.

Generally, in LS algorithm, the value of \( v \) is often set as global constant. A modified balloon force pulls or pushes the dynamic interface adaptively towards the object of interest represented as \( G(R_k) = 1 - 2R_k \), where \( G(R_k) \in [-1, 1] \) is a matrix with a variable force that pull or push each image pixel. So Eq. (5.14) is modified as
\[ \xi(g, \phi) = \lambda \delta(\phi) \text{div} \left( g \frac{\nabla \phi}{|\nabla \phi|} \right) + g R_k \delta(\phi) \]  
\text{Eq. (5.15)}

The coefficient of the contour length \( \lambda \) is calculated as \( \lambda = 0.1 \zeta \) and the penalty coefficient \( \mu \) is calculated as \( \mu = 0.2 / \xi \). The \( \zeta \) is represented as \( \zeta = \alpha / \ell \), where \( \ell \) and \( \alpha \) estimate the length and the area from the initial LS function \( \phi^0 \) obtained from the kernel based fuzzy clustering.

\[ \alpha = \int I H(\phi^0) \text{d}x \text{d}y \]  
\text{Eq. (5.16)}

\[ l = \int \delta(\phi^0) \text{d}x \text{d}y \]  
\text{Eq. (5.17)}

where the Heaviside function \( H(\phi^0) \) is

\[ H(\phi^0) = \begin{cases} 1, & \phi^0 \geq 0 \\ 0, & \phi^0 < 0 \end{cases} \]  
\text{Eq. (5.18)}

An edge indication function \( g \) is used to stop level set evolution near the optimal solution defined as

\[ g = \frac{1}{1 + |\nabla (G_\sigma * I)|^2} \]  
\text{Eq. (5.19)}

where \( G_\sigma * I \) represents the convolution of \( I \) with Gaussian smoothing kernel \( G_\sigma \) and \( \nabla \) represents the image gradient.

The level set function is updated as,

\[ \phi^{t+1}(x, y) = \phi^t(x, y) + \tau \left[ \mu \zeta(\phi^t) + \xi(g, \phi) \right] \]  
\text{Eq. (5.20)}

where \( \tau \) represents the time step of the level set evolution and it is assigned as \( \zeta \), \( \phi^t \) represents the level set function at time \( t \) and \( \phi^{t+1} \) represents the level set function at time \( t + 1 \). The algorithm continues until difference between
\( \phi^{t+1} - \phi^t < \varepsilon \) or maximum number of iteration is reached where \( \varepsilon \) is a termination criterion. The control parameters for level set evolution are adaptively selected from the kernel based fuzzy clustering.

### 5.2.4 The proposed algorithm

The proposed mass segmentation system utilizes the kernel based fuzzy clustering and the level set function without re-initialization to develop an automatic mass segmentation for mammogram images.

#### Algorithm 2: Algorithm of the KFLS methodology

1. Select the input image \( (I) \) after preprocessing.
2. Set the number of clusters \( (C) \) and the termination criteria \( (\varepsilon) \) for the KFCM clustering.
3. Perform the KFCM clustering followed by morphological opening on the input image.
4. Select the cluster that contain the ROI from the output of KFCM clustering for further refinement.
   a) Calculate the maximum membership value for each pixel to identify in which cluster it belongs to.
   b) Calculate mode intensity value for each cluster region.
   c) The cluster with maximum mode value is selected as ROI for level set evolution.
5. Initialize the LS function through the initial contour obtained in Step 4.
6. The controlling parameters for the LS function are evaluated by the result of KFCM clustering method as discussed in Section 5.2.3.

The proposed system starts with KFCM clustering to perform an initial segmentation. Later, the output of the KFCM clustering followed by morphological opening is used to initialize the level set evolution function. The outcome of the
KFCM clustering is utilized to compute the controlling parameters of the level set function that is responsible for final contour generation. The algorithm of the KFLS methodology is given in Algorithm 2.

### 5.3 RESULTS AND DISCUSSIONS

The experiments and performance evaluation are carried out on the mammogram images of mini-MIAS and DDSM database. The mammogram images used in this experiment contain significant mass regions and a radiologist marked ROI to evaluate the performance of the segmentation algorithm.

To show the results of the automatic segmentation, two images such as C_0051 and B_3102 are selected from the DDSM data set. Figure 5.5 shows the sample inputs and its preprocessed output. Figure 5.5 (a) and 5.5 (c) shows the sample images of C_0051 with malignant mass and B_3102 with benign mass respectively. The resultant preprocessed image is further applied as an input to the KFCM.

![Sample inputs and its preprocessed output](image)

Figure 5.5 Sample inputs and its preprocessed output a) Input image C_0051, b) Preprocessed image of (a), c) Input image B_3102, d) Preprocessed image of (c)

It is observed that a mammogram image is divided into 6 different number of sub-regions based on the anatomy structure of breast tissues shown in Figure 5.2. The outcome of KFCM segmentation method obtained for maximum membership of each pixel with the value of $C = 6$ and $l = 2$ is depicted in Figure 5.6 (a). The region of mass corresponding to the cluster is highlighted in red color.
The next important process is to identify the cluster index that contains the ROI. From each cluster region, the mode intensity value is calculated. The cluster with the maximum mode value is selected as the cluster with the ROI. Figure 5.6 (b) depicts the selected cluster that contains the mass region.

**DDSM Image: C_0051**

![DDSM Image: C_0051](image1)

**DDSM Image: B_3102**

![DDSM Image: B_3102](image2)

Figure 5.6 Results of KFCM segmentation after post processing a) KFCM segmentation, b) Selected cluster with ROI, c) Post processed image

It is observed from Figure 5.6 (b) that, some of the normal tissues are also identified as abnormalities along with the original abnormality present, which are characterized as false positives. In order to reduce the number of false positives, a post processing algorithm is employed. Then the features such as area and GLCM based sum average measures are computed for each of the identified abnormal regions. The GLCM based sum average feature is analyzed to provide better result in the selection of abnormal region as discussed in Section 4.2.3 of Chapter 4. From the observations of the area computed from these regions, it is noted that abnormal
areas lie in the range from 1000 to 40,000 squared pixels. Based on these empirical observations obtained through experiments, for each selected regions, the measures are computed. Those with values below the maximum are discarded from being abnormal regions. The result after post processing is shown in Figure 5.6 (c). It is noticed from the Figure 5.6 (c) that, after post processing, the false positive regions are eliminated.

After refining the fuzzy clustering, the initial contour of the mass is obtained, which is fed to the LS algorithm to refine the boundary of the mass regions. The KFCM algorithm provides a better approximation of boundaries of interest. The level set algorithm estimates the controlling parameters from kernel based fuzzy clustering automatically and starts the evolution from a region close to the genuine boundary. Then the level set evolution stabilizes automatically once it approaches the genuine boundary, which suppresses boundary leakage and avoid manual intervention.

The KFLS algorithm allows a flexible initialization in medical image segmentation. Three different initializing methods are evaluated and compared as shown in Figure 5.7. The manual initialization and thresholding are used for level set initialization. Due to weak and indistinct tissues, image inhomogenity and boundary leakage, manual initialization is considered not a good choice for optimal LS segmentation. Prior knowledge is required by manual initialization for initial point incorporation into deformable models. However the problem of local minima exists with this approach. Since the initial point is away from the ROI boundaries, the evolution time is very slow.

In case of thresholding initialization, an optimal threshold need to be calculated, this is a challenging process. Fuzzy clustering initialization adaptively obtains the approximate boundary of the ROI, and it is suitable to initiate the LS segmentation. There are various merits with initialization of contour in LS using
KFCM. Since the initial contours are seen close to the boundaries of the ROI, the curve evolution speed is high. The process of initial contour extraction is achieved without human intervention. The prior knowledge is not necessary for the initial contour selection. KFCM provides an adaptive process for the potential component of interest for extraction. Thus the presence of KFCM is significant for the LS evolution.

Figure 5.7 Initialization methods for level set segmentation of C_0051 (Column 1 and Column 2) and B_3102 (Column 3 and Column 4) images a) and c) Manual initialization, b) and d) Final segmentation after 100 iterations, e) and g) Initialization by thresholding (0.75-0.85), f) and h) Final segmentation after 100 iterations, i) and k) Initialization by KFCM, j) and l) Final segmentation after 100 iterations
Figure 5.8 shows the performance of the level set algorithm on the contour extracted from the kernel based fuzzy clustering. The output shows the contour initialization using KFCM and the refined process of kernel based fuzzy level set segmentation after 50 iterations, 75 iterations and 100 iterations.

To stop the evolution of a contour, a predetermined threshold is often used. Additionally, a maximum number of iterations have been set. The convergence of evolution is determined with various metrics, such as the change of level set function, the change of length of contour and the change of area inside the contour. Yuan et al. (2007) developed a dynamic method to terminate contour evolution automatically. During dynamic contour evolution, the weighted difference between the mean slope of foreground and that of background is monitored. The contour evolution is terminated when the weighted slope difference converges to zero. This method provides a way to terminate contour evolution free of predefined threshold. Recently, Kuo et al. (2014) has also found that the dynamic stopping criterion performs better for the termination of contour evolution.
It is observed from exhaustive experimentations conducted on the proposed method that, almost for all images the contour evolution is terminated in and around 100 iterations. Therefore, the maximum number of iterations for level set evolution is set to be 100 in this experiment.

The mass border is one of the important features to classify the abnormality as malignant or benign in a more accurate manner. The proposed mass segmentation with the precise boundary is evaluated against the manually delineated ROI obtained from the experienced radiologist. The significance of this segmentation methodology is evaluated with the performance measures such as sensitivity, precision, missing, over hitting, relative missing, relative error (RE), area overlap measure (AOM) and accuracy as given in Section 4.3 of Chapter 4.

The results of quantitative analysis to prove the effectiveness of the proposed algorithm is shown in Table 5.1. The measures such as sensitivity, specificity, missing, over hitting, relative missing are calculated from the mass region for 10 sample mammogram images from the DDSM data set.

Table 5.1 Performance validation measures of the segmented mass using KFLS for 10 sample mammogram images on DDSM database

<table>
<thead>
<tr>
<th>Case No.</th>
<th>Sensitivity (%)</th>
<th>Precision (%)</th>
<th>Missing (%)</th>
<th>Over Hitting (%)</th>
<th>Relative Missing (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C_0046</td>
<td>93.69</td>
<td>97.21</td>
<td>6.31</td>
<td>2.69</td>
<td>6.54</td>
</tr>
<tr>
<td>C_0051</td>
<td>96.73</td>
<td>98.33</td>
<td>3.27</td>
<td>1.64</td>
<td>3.32</td>
</tr>
<tr>
<td>C_0069</td>
<td>91.17</td>
<td>92.48</td>
<td>8.82</td>
<td>7.42</td>
<td>8.96</td>
</tr>
<tr>
<td>C_0074</td>
<td>86.50</td>
<td>90.26</td>
<td>13.50</td>
<td>9.33</td>
<td>14.09</td>
</tr>
<tr>
<td>C_0226</td>
<td>89.87</td>
<td>94.14</td>
<td>10.13</td>
<td>5.59</td>
<td>10.61</td>
</tr>
<tr>
<td>A_1347</td>
<td>85.69</td>
<td>93.72</td>
<td>14.31</td>
<td>5.73</td>
<td>15.65</td>
</tr>
<tr>
<td>C_0249</td>
<td>91.56</td>
<td>93.55</td>
<td>8.44</td>
<td>6.31</td>
<td>8.62</td>
</tr>
<tr>
<td>B_3091</td>
<td>93.33</td>
<td>96.13</td>
<td>6.66</td>
<td>3.76</td>
<td>6.86</td>
</tr>
<tr>
<td>B_3102</td>
<td>97.90</td>
<td>98.49</td>
<td>2.09</td>
<td>1.49</td>
<td>2.11</td>
</tr>
<tr>
<td>B_3114</td>
<td>94.94</td>
<td>96.43</td>
<td>5.05</td>
<td>3.51</td>
<td>5.13</td>
</tr>
</tbody>
</table>

From the Table 5.1, it is found that the sensitivity achieves a minimum of 80.66% and maximum of 97.90%. Also the algorithm attains the missing mass
ratio from 2.09% to 19.34%. The results show the effectiveness of the segmentation algorithm in segmenting the masses from the mammograms.

The relative error is calculated between the area segmented by the proposed algorithm and the area marked by the experienced radiologist (i.e Ground Truth). Table 5.2 shows the results of the relative error and area overlap measure calculated in the mass area for 10 sample mammogram images from the data set. The algorithm achieves area overlap measure of 70.18% to 96.45%. The minimum relative error achieved in the algorithm is 0.59%.

Table 5.2 Comparison of segmented mass area (pixels) with the ground truth for 10 sample mammogram images on DDSM database

<table>
<thead>
<tr>
<th>Case No.</th>
<th>Ground Truth Area (A)</th>
<th>Estimated Area (A)</th>
<th>TP</th>
<th>FP</th>
<th>FN</th>
<th>RE (%)</th>
<th>AOM (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C_0046</td>
<td>1411</td>
<td>1360</td>
<td>1322</td>
<td>38</td>
<td>89</td>
<td>3.61</td>
<td>91.24</td>
</tr>
<tr>
<td>C_0051</td>
<td>8655</td>
<td>8514</td>
<td>8372</td>
<td>142</td>
<td>283</td>
<td>1.63</td>
<td>95.17</td>
</tr>
<tr>
<td>C_0069</td>
<td>1631</td>
<td>1608</td>
<td>1487</td>
<td>121</td>
<td>144</td>
<td>1.41</td>
<td>84.87</td>
</tr>
<tr>
<td>C_0074</td>
<td>6569</td>
<td>6295</td>
<td>5682</td>
<td>613</td>
<td>887</td>
<td>4.17</td>
<td>79.11</td>
</tr>
<tr>
<td>C_0226</td>
<td>2054</td>
<td>1961</td>
<td>1846</td>
<td>115</td>
<td>208</td>
<td>4.53</td>
<td>85.11</td>
</tr>
<tr>
<td>A_1347</td>
<td>2180</td>
<td>1993</td>
<td>1868</td>
<td>125</td>
<td>312</td>
<td>6.58</td>
<td>81.04</td>
</tr>
<tr>
<td>C_0249</td>
<td>1600</td>
<td>1566</td>
<td>1465</td>
<td>101</td>
<td>135</td>
<td>2.13</td>
<td>86.13</td>
</tr>
<tr>
<td>B_3091</td>
<td>1968</td>
<td>1911</td>
<td>1837</td>
<td>74</td>
<td>131</td>
<td>2.89</td>
<td>89.96</td>
</tr>
<tr>
<td>B_3102</td>
<td>2001</td>
<td>1989</td>
<td>1959</td>
<td>30</td>
<td>42</td>
<td>0.59</td>
<td>96.45</td>
</tr>
<tr>
<td>B_3114</td>
<td>3501</td>
<td>3447</td>
<td>3324</td>
<td>123</td>
<td>177</td>
<td>1.54</td>
<td>91.72</td>
</tr>
</tbody>
</table>

Table 5.3 gives percentage of sensitivity, precision, RE, AOM, missing, over hitting and relative missing for 10 sample mammogram images from mini-MIAS database.

The proposed algorithm provides improved performance on the images of the mini-MIAS database as shown in Table 5.3. It provides RE ranges between 1.59% and 27.82% for 10 sample mammogram images. The maximum AOM achieved in this method is 97.53%. As the CIRC and MISC type images provides improved AOM compared to other categories due to its initial contour obtained near to the boundary.
Table 5.3 Results of the proposed KFLS based mass segmentation for 10 sample mammogram images on mini-MIAS database

<table>
<thead>
<tr>
<th>Case No.</th>
<th>Sensitivity (%)</th>
<th>Precision (%)</th>
<th>Missing (%)</th>
<th>Over Hitting (%)</th>
<th>Relative Missing (%)</th>
<th>RE (%)</th>
<th>AOM (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mdb023</td>
<td>97.05</td>
<td>98.61</td>
<td>2.96</td>
<td>1.36</td>
<td>3.01</td>
<td>1.59</td>
<td>95.72</td>
</tr>
<tr>
<td>Mdb028</td>
<td>99.81</td>
<td>97.71</td>
<td>0.18</td>
<td>2.33</td>
<td>0.18</td>
<td>2.15</td>
<td>97.53</td>
</tr>
<tr>
<td>Mdb032</td>
<td>93.79</td>
<td>98.98</td>
<td>6.20</td>
<td>0.95</td>
<td>6.55</td>
<td>5.25</td>
<td>92.90</td>
</tr>
<tr>
<td>Mdb134</td>
<td>98.34</td>
<td>94.85</td>
<td>1.65</td>
<td>5.33</td>
<td>1.59</td>
<td>3.68</td>
<td>93.36</td>
</tr>
<tr>
<td>Mdb092</td>
<td>88.58</td>
<td>99.17</td>
<td>11.41</td>
<td>0.73</td>
<td>12.77</td>
<td>10.67</td>
<td>87.93</td>
</tr>
<tr>
<td>Mdb102</td>
<td>98.96</td>
<td>77.42</td>
<td>1.03</td>
<td>28.85</td>
<td>0.80</td>
<td>27.82</td>
<td>76.80</td>
</tr>
<tr>
<td>Mdb181</td>
<td>92.92</td>
<td>98.65</td>
<td>7.07</td>
<td>1.27</td>
<td>7.51</td>
<td>5.80</td>
<td>91.75</td>
</tr>
<tr>
<td>Mdb202</td>
<td>76.82</td>
<td>97.08</td>
<td>23.17</td>
<td>2.30</td>
<td>29.29</td>
<td>20.87</td>
<td>75.09</td>
</tr>
<tr>
<td>Mdb117</td>
<td>76.25</td>
<td>100</td>
<td>23.74</td>
<td>0</td>
<td>31.13</td>
<td>23.74</td>
<td>76.25</td>
</tr>
<tr>
<td>Mdb152</td>
<td>94.21</td>
<td>83.05</td>
<td>5.78</td>
<td>19.22</td>
<td>5.09</td>
<td>13.44</td>
<td>79.02</td>
</tr>
</tbody>
</table>

Figure 5.9 depicts sample results of segmented mass region in mini-MIAS database obtained from conventional FCM with LS and KFLS segmentation. The FCM with LS algorithm failed to segment the ROI properly (or some pixels are incorrectly segmented) and takes more time to converge as compared to KFCM.

![Figure 5.9 Segmentation results of sample ROI of mdb023 from mini-MIAS database](image)

(a) FCM segmentation  
(b) FCM with LS segmentation  
(c) KFCM segmentation  
(d) KFLS segmentation

The comparison of sensitivity, AOM and convergence speed in terms of number of iterations between the proposed KFLS and the FCM with LS segmentation methods are depicted in Figure 5.10.
Figure 5.10 Performance comparisons between FCM with LS and KFLS
It is observed from the Figure 5.10 that the proposed KFLS segmentation method precisely segments the mass boundaries with an improved sensitivity and accuracy compared to conventional FCM with LS. It also shows that the number of iterations needed to converge the clustering is less in KFLS when compared to FCM with LS.

5.3.1 Comparative analysis of KFLS segmentation with MCSU in CA-based segmentation

Mammograms typically have low contrast and boundaries of masses that are vague or obscured by soft tissue or even not visible. Segmentation leak is a challenge for mammographic mass segmentation. The MCSU in CA-based segmentation described in Chapter 4 is prone to boundary leakage, if there is a gap in the boundary. The proposed KFLS segmentation solves this problem by utilizing the KFCM initialization with LS evolution.

DDSM Image: B_3114

Mini-MIAS Image: mdb090

Figure 5.11 Comparison of MCSU in CA-based segmentation and KFLS segmentation a) Sample ROI of B_3114 and mdb090 respectively b) MCSU in CA-based segmentation c) KFCM-based contour initialization d) Boundary refinement by KFLS segmentation
Figure 5.11 (a) shows an example ROI which contains a mass. The mass edge is blurred and partially overlapped by other surrounding tissues, therefore MCSU in CA-based segmentation method fails to segment the boundary as shown in Figure 5.11 (b) due to the weakness of the boundary. Figure 5.11 (c) shows the boundary resulting from KFCM clustering and morphological opening. The boundary is then refined by the LS segmentation as shown in Figure 5.11 (d). The LS segmentation covers most mass edges visually as compared to MCSU in CA-based segmentation.

The average performance comparison of the proposed KFLS mass segmentation algorithm compared to MCSU in CA-based segmentation for various abnormality categories in mini-MIAS database is shown in Table 5.4.

<table>
<thead>
<tr>
<th>Method</th>
<th>Metrics</th>
<th>CIRC</th>
<th>MISC</th>
<th>ASYM</th>
<th>SPIC</th>
<th>ARCH</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCSU in CA</td>
<td>Sensitivity (%)</td>
<td>97.07</td>
<td>93.61</td>
<td>90.95</td>
<td>87.48</td>
<td>86.66</td>
<td>91.16</td>
</tr>
<tr>
<td>KFLS</td>
<td></td>
<td>98.42</td>
<td>96.06</td>
<td>93.77</td>
<td>84.87</td>
<td>85.23</td>
<td></td>
</tr>
<tr>
<td>MCSU in CA</td>
<td>Precision (%)</td>
<td>93.84</td>
<td>90.12</td>
<td>89.73</td>
<td>87.95</td>
<td>80.66</td>
<td>88.46</td>
</tr>
<tr>
<td>KFLS</td>
<td></td>
<td>98.14</td>
<td>96.92</td>
<td>88.29</td>
<td>97.86</td>
<td>91.52</td>
<td></td>
</tr>
<tr>
<td>MCSU in CA</td>
<td>RE (%)</td>
<td>2.02</td>
<td>4.65</td>
<td>11.72</td>
<td>18.91</td>
<td>20.21</td>
<td>11.50</td>
</tr>
<tr>
<td>KFLS</td>
<td></td>
<td>1.87</td>
<td>4.45</td>
<td>11.24</td>
<td>13.33</td>
<td>18.59</td>
<td></td>
</tr>
<tr>
<td>MCSU in CA</td>
<td>AOM (%)</td>
<td>93.02</td>
<td>90.14</td>
<td>80.97</td>
<td>75.97</td>
<td>73.79</td>
<td>82.77</td>
</tr>
<tr>
<td>KFLS</td>
<td></td>
<td>96.63</td>
<td>93.13</td>
<td>82.37</td>
<td>83.42</td>
<td>77.64</td>
<td></td>
</tr>
<tr>
<td>MCSU in CA</td>
<td>Accuracy (%)</td>
<td>98.02</td>
<td>95.14</td>
<td>92.97</td>
<td>87.97</td>
<td>85.79</td>
<td>91.97</td>
</tr>
<tr>
<td>KFLS</td>
<td></td>
<td>99.63</td>
<td>96.13</td>
<td>94.37</td>
<td>85.42</td>
<td>84.64</td>
<td>92.04</td>
</tr>
</tbody>
</table>

In CIRC mass type, margin might have an abrupt transition between the lesion and the surrounding tissue. The MISC masses provide higher intensity values in the center of mass than background regions. The KFCM clustering provides initial contour near to the boundary for these categories. So the proposed algorithm provides high sensitivity of 98.42% for CIRC masses and 96.06% for MISC masses. The MISC masses contain higher gray values in the center of mass than
background regions. For ASYM masses, the proposed KFLS method achieves sensitivity of 93.77%.

In some cases of ARCH and SPIC mass type, no definite mass is visible and display great inconsistency in their shapes. These masses do not have prominent central core regions and possess poor contrast with respect to their background. As the LS segmentation is sensitive to contour initialization, the fuzzy clustering fails to initialize the contour properly. Hence, KFLS provides less sensitivity compared to other mass categories and compared to MCSU in CA-based segmentation. The sensitivities for SPIC and ARCH are 84.87% and 85.23% respectively. Table 5.4 reveals that the proposed algorithm yields less RE of 9.09% and improved AOM of 86.64% with an accuracy of 92.04% compared to the MCSU in CA-based segmentation.

Table 5.5 shows the performance of the proposed algorithm with other algorithms in terms of average sensitivity, precision, RE, AOM, accuracy and computation time. In order to prove the efficiency of the proposed algorithm against the state-of-art methods such as watershed segmentation (Dubey et. al., 2010), region growing segmentation (Berber et. al., 2013), snake-based active contour segmentation (Xu et. al., 2006), chan-vese (CV) levelset segmentation (Liu et. al., 2011) and MCSU in CA-based segmentation, these methods are implemented, tested and compared with the same set of image samples from the mini-MIAS and DDSM database.

From Table 5.5, it is understood that the proposed method outperforms watershed segmentation by 17.92%, region growing segmentation by 13.61%, snake-based active contour segmentation by 9.49%, CV level set segmentation by 6.85%, MCSU in CA-based segmentation by 1.18% in terms of average sensitivity. This improvement is due to the initialization of KFCM in the computation of the initial curve that speeds up the evolution of level set function.
Table 5.5 Average performance comparison of various state-of-art approaches with KFLS based mass segmentation

<table>
<thead>
<tr>
<th>Approach</th>
<th>Sensitivity (%)</th>
<th>Precision (%)</th>
<th>RE (%)</th>
<th>AOM (%)</th>
<th>Accuracy (%)</th>
<th>Computation Time (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Watershed Segmentation (Dubey et al. 2010)</td>
<td>79.34</td>
<td>80.58</td>
<td>15.37</td>
<td>73.56</td>
<td>83.34</td>
<td>3.56</td>
</tr>
<tr>
<td>Region Growing Segmentation (Berber et al. 2013)</td>
<td>82.35</td>
<td>82.94</td>
<td>14.72</td>
<td>75.45</td>
<td>85.3</td>
<td>4.78</td>
</tr>
<tr>
<td>Snake-based Active Contour Segmentation (Xu et al. 2006)</td>
<td>85.45</td>
<td>85.29</td>
<td>14.22</td>
<td>75.78</td>
<td>86.45</td>
<td>10.83</td>
</tr>
<tr>
<td>Chan-vese Levelset Segmentation (Liu et al. 2011)</td>
<td>87.56</td>
<td>85.54</td>
<td>13.45</td>
<td>78.43</td>
<td>89.21</td>
<td>9.14</td>
</tr>
<tr>
<td>MCSU in CA-based Mass Segmentation</td>
<td>92.46</td>
<td>90.83</td>
<td>11.53</td>
<td>82.49</td>
<td>93.23</td>
<td>4.82</td>
</tr>
<tr>
<td>KFLS Segmentation (proposed)</td>
<td>93.56</td>
<td>92.67</td>
<td>10.12</td>
<td>85.27</td>
<td>94.89</td>
<td>8.29</td>
</tr>
</tbody>
</table>
Similarly, the analysis of the proposed method in terms of AOM yields an improvement of 15.91% compared to watershed segmentation, 13.01% compared to region growing segmentation, 12.52% compared to snake-based active contour segmentation, 8.72% compared to CV level set segmentation and 3.37% compared to MCSU in CA-based segmentation. The proposed method is able to cover considerable mass region than any other method with average relative error of 10.12%.

According to the result analysis, it is found that the problem of contour initialization in LS function is addressed automatically by a KFCM based rough segmentation. The initial contours selected by KFCM results in the improvement of contour evolution, since the level set evolution starts from a region which is closed to the genuine boundaries, and is also selected automatically.

5.4 CONCLUSION

A new kernel based fuzzy level set algorithm has been proposed for automatic mass segmentation in mammograms. A preprocessing method based on thresholding and morphological operations are used to eliminate the artifacts from the mammogram. The median filter reduces the noise in the input mammogram. A kernel based FCM followed by morphological opening is applied to the preprocessed image to identify the region with ROI, followed by the post processing step to remove the FPs. The KFLS algorithm utilizes the output of the fuzzy clustering as the initial contour for the level set evolution.

The reasons for using fuzzy technique based image segmentation algorithm for contour initialization are as follows: Since the contrast in mammograms is very low and the boundary between normal tissue and mass is unclear, the traditional segmentation methods might not work well. The classical region growing based segmentation technique tries to precisely define ROIs, but to find a criterion for segmentation is difficult as most of the malignant tumors with
fuzzy boundaries extend from a dense core region to the surrounding tissues. Similarly, the classical global or local thresholding techniques try to segment ROIs, but the techniques are only effective for the objects with clear boundaries. The fuzzy logic based approaches are useful for segmenting suspicious regions and are capable of addressing above issues.

The level set algorithm starts its evolution from the boundary close to the region of interest. Automatic selection of the starting point from fuzzy clustering reduces the user intervention. The combination of kernel based fuzzy clustering and level set segmentation speeds up the entire segmentation process and improves the segmentation. Experimental results show that the proposed KFLS-based segmentation method obtains good segmentation results in terms of several metrics.

When compared to MCSU in CA-based segmentation, the proposed KFLS-based segmentation provides statistically significant (at 5% significant level with p-value 0.2775 i.e. less than t-value 0.6130) improvement in terms of sensitivity and avoid boundary leakages in case of weak boundaries. The penalty term and the energy term present in the LS function evolution prevents the contour from leaking the object boundary at areas of poor edges or if there are gaps in the boundary. The performance of the LS algorithm is mainly depends on the initial contour which provides improved results for the regular shape masses. If the fuzzy clustering fails to initialize the contour properly, then the KFLS algorithm achieves less performance than MCSU in CA-based segmentation. The fuzzy clustering is also sensitive to selection of the initial cluster centre, so sometimes it may be easily trapped into local minima. The computation time of the KFLS method is high compared to MCSU in CA-based segmentation method due to the convergence of fuzzy clustering and the level set function evolution.
Once the suspicious regions are identified, these regions may be analyzed by extracting various features that can be used in the CADx system for efficient diagnosis. The next chapter introduces a computer aided breast cancer classification system that utilizes the multiresolution analysis for efficient feature extraction to characterize the suspicious mass regions as normal or abnormal and the abnormalities into benign or malignant.