Breast mass contour segmentation algorithm in digital mammograms

Tolga Berbera,*, Adil Alpkocakb, Pinar Balci, Oguz Diclec

a Dokuz Eylul University, Department of Computer Engineering, Graduate School of Natural and Applied Sciences, Izmir, Turkey
b Dokuz Eylul University, Department of Computer Engineering, Engineering Faculty, Izmir, Turkey
c Dokuz Eylul University, Department of Radiodiagnosics, Medical School, Izmir, Turkey

ABSTRACT

Many computer aided diagnosis (CAD) systems help radiologist on difficult task of mass detection in a breast mammogram and, besides, they also provide interpretation about detected mass. One of the most crucial information of a mass is its shape and contour, since it provides valuable information about spread ability of a mass. However, accuracy of shape recognition of a mass highly related with the precision of detected mass contours. In this work, we introduce a new segmentation algorithm, breast mass contour segmentation, based on classical seed region growing algorithm to enhance contour of a mass from a given region of interest with ability to adjust threshold value adaptively. The new approach is evaluated over a dataset with 260 masses whose contours are manually annotated by expert radiologists. The performance of the method is evaluated with respect to a set of different evaluation metrics, such as specificity, sensitivity, balanced accuracy, Yassonf and Hausdorff error distances. The results obtained from experimentations shows that our method outperforms the other compared methods. All the findings and details of approach are presented in detail.

© 2012 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

Breast cancer is the most commonly diagnosed cancer type among women such that a new breast cancer case is diagnosed in every 2.2 min and a woman dies from breast cancer in every 13 min [1]. Early detection of breast cancer is the key factor for increasing recovery rate from disease like other cancer types. Digital mammography systems is the key tool for breast cancer detection in early stages with some exceptions which additional diagnosis methods is required for final decision [2,3].

Many researchers have been working on computer-aided diagnosis system (CADx) to detect and identify breast masses automatically or semi-automatically in digital mammograms over several decades. All these works aim to support radiologists in the difficult task of discriminating benign and malignant breast lesions. Hence, it is not surprising that typically only 15–30% of breast biopsies performed on calcifications will be positive for malignancy [4].

Segmentation of breast masses in mammography from the background tissue is an important problem, as well as detection. Especially, finding correct mass contour is the key factor for successful mass interpretation, since mass interpretation depends on shape, contour and margin characteristics of a breast mass. Although most of the current CADx systems provide no such information, we think that a modern CADx system must help radiologist on not only mass detection task, but also mass identification task. For instance, a CADx system that automatically provides suggestions for masses
According to BI-RADS mass properties could improve level of CADx in medical diagnosis. Such a CADx system requires mass boundary to be detected as accurately as possible. Hence, segmentation algorithm used to detect masses should also detect accurate boundary of a mass also.

Accuracy of breast mass contour segmentation plays an important role in CADx systems for mass classification point of view, since most valuable properties of breast masses that define malignancy is related to its morphology. For instance, malign masses tend to spread other areas, while benign masses remain stable. As a result, malign masses form irregular shapes; conversely benign masses form regular shapes. Kangayyan and Nguyen [5] states importance of shape characteristics while defining mass malignancy and proposed several low-level features to determine boundary characteristics by using region shape. However, it is very hard to find such a segmentation algorithm since there is no way to define exact boundary of a mass. Moreover, finding real mass area is a crucial task even for a radiology expert. On the other hand, American College of Radiology (ACR) defines a standard for mammography reporting named BI-RADS [6]. BI-RADS standard categorizes mass boundary into 5 classes; Circumscribed, Obscured, Microlobulated, Ill-defined and Spiculated. Each of these categories is based on rough boundary characteristics of masses selected by radiologists. Hence, accuracy of a segmentation algorithm could be defined by similarity between segmentation results and manually segmented regions.

Many recent studies focuses on machine-learning methods to measure similarity between automatically segmented regions and manual selected regions. Tao et al. [7] propose a classification system to identify spiculation of a mass with help of automatic segmentation approach. In their study, the ROI of a candidate mass is splitted into subregions and each sub region is labeled using machine learning techniques. Then, the final segmentation result is obtained by using graph-cut algorithm and optimization methods. Song et al. [8] propose a segmentation method which uses plane-fitting method based on dynamic programming optimization approach. Proposed method includes a preprocessing step in which edge information is enhanced by using gradients. Experimentation results of the proposed method on a subset of publicly available dataset give an acceptable level of performance. Meanwhile, in another recent study by Song et al. [9], performance of dynamic programming with combination of template matching is tested and found that the template matching approach performs better than plane fitting approach. A very recent an in-depth review on this topic with respect to mass detection can be found in [10].

Some other studies focus on contour tracing approaches since margin characteristics of masses includes important clues about for malignancy of masses. Elter et al. [11] proposed a contour tracing approach to extract shape of the region. To extract contour information, the ROI of the candidate mass is transformed into polar coordinate system, and then contour of estimated mass is calculated by using a shortest path algorithm. Proposed segmentation method tested on 60 ROI images of DDSM dataset and proposed method obtains the highest performance scores. Hong and Sohn [12] propose different representation using the iso-contour maps for topographic, in which a salient region forms a dense quasi-concentric pattern of contours. The topological and geometrical structures of the image are analyzed using an inclusion tree that is a hierarchical representation of the enclosure relationships between contours. They have evaluated their approaches only for salient masses and show that a topographic representation is largely invariant to brightness and contrast, and it provides a robust and efficient representation for the characterization of mammographic features. Some other researchers use well-known segmentation methods to identify mass region. Domínguez and Nandi [13] propose a new thresholding approach which uses combination of several threshold levels to detect masses from a digital mammography images. Proposed algorithm achieves 80% of sensitivity and 2.3 false positives per image. Wei et al. [14] uses a seeded region growing approach to extract mass region as a part of their retrieval system. After finding mass boundary, they extract Zernike moment features from region. Proposed system achieves varying precision and recall values around 80% and 85%. Dubey et al. [15] measure performance of level-set and watershed segmentation methods on mass segmentation. They found that Watershed segmentation algorithm produces better results than level-set segmentation approach with respect to relative error measure.

In this work, a breast masses contour segmentation (BMCS) method is presented, which is based on region growing algorithm with additional capability of adaptive threshold value to extract optimum contour information. The proposed approach is evaluated using on a reference mammography dataset, where expert radiologists manually selected boundaries. The performance of the proposed approach is directly compared to the performance of the implementations of three state-of-the-art region growing approaches solutions, in terms of a set of different measurement metrics, such as Yassnoff distance, balanced accuracy. We show that BMCS outperformed other three segmentation methods.

The rest of the paper is organized as follows. In Section 2, we introduce basic definitions about breast mass segmentation and provide the state-of-the-art methods on the topic and detailed representation of our proposed method. Next, in Section 3, we first give a comprehensive discussion of evaluation methods for a good comparison of mass segmentation methods. Later, we describe the dataset we used for evaluation of our approaches and provide the results we obtained from our experimentation. The last, the Section 4 concludes the paper and give a projection on further studies on this topic.

2. Breast mass contours segmentation (BMCS)

Breast mass detection using a single mammography image relies on the fact that pixels inside a mass have different characteristics from the other pixels within the breast area. These characteristics can be simply related to gray-level intensity, texture or morphological characteristics.

Breast mass segmentation methods can be roughly classified into three groups: region-based, contour-based and clustering methods. Region-based methods divide the image into spatially connected homogeneous regions while contour-based methods rely on the boundary of regions. On the other
hand, clustering methods group together the pixels having the same properties and might result in non-connected regions. Since we aim to enhance boundary of a mass from a given region, we propose an improvement on region growing approach, which is one of the well-known region based method.

2.1. Reference segmentation methods

We choose three segmentation methods to evaluate our segmentation approach. All of the segmentation approaches are region based, since we focused on mass boundary enhancement instead of mass detection. And some of the works [15,16] used region based methods to enhance initial detection results. Selected segmentation approaches are discussed below.

Watershed segmentation method is based on watershed transform [17,18]. This method aims to find catchment basins, which define border between two objects. If water falls into these basins, level of the water rises until neighbor basins share the same level. So output of the algorithm is a hierarchy of catchment basins. The key point is to find most discriminative basins, since most discriminative basins are the basins that separate two different objects.

Level-set segmentation approach is region enhancement approach which evolves an initial region according to an energy field [19]. In other words, level-set methods start with an initial region and evolve this region while minimizing region energy. Energy of a region is calculated using the level-set equation, which is a partial differential equation, in each step. Finally algorithm stops when difference between energy of region in consecutive steps falls under a threshold value.

Our method, BMCS is based on seeded region growing (SRG) method [20]. SRG segmentation approach expands an initial region or point by similar neighbor pixels. Similar term means a pixel whose intensity is in predefined range with seed region/point in this scope. So, we include seeded region growing method in performance comparison.

2.2. Breast mass contour segmentation algorithm

In this study, we propose a breast mass contour segmentation (BMCS) approach for a given ROI in an image, which is based on seeded region growing (SRG) algorithm and it is a variant of region based segmentation method. SRG algorithm starts with a seed point and extends it by adding neighbor pixels that intensity value exceeds a predefined threshold value. But this approach is not appropriate for mass contour segmentation, because it is hard to choose a general threshold value for all kind of masses. Thus, SRG segmentation method needs to be tuned for each kind of masses. Otherwise, SRG segmentation method produces smaller or larger mass segments than reference mass segment, and these results are commonly referenced as under- and over-segmentation, respectively. We proposed a new thresholding approach for SRG, which adaptively adjusts the threshold value based on mass size estimation to prevent under- and over-segmentation.

BMCS contains four steps; (1) ROI detection, (2) preprocessing, (3) mass size estimation and (4) segmentation. At first step, BMCS needs a rough bounding rectangle of a mass. Since we focus on mass boundary enhancement, we omitted how bounding rectangle is selected. We assume that bounding rectangle of a mass region are drawn by user and given to system as input. We do not prefer to use result of a machine segmentation method, since false positives may occur in machine segmentation techniques [21].

In preprocessing step, we aim to increase contrast level between mass and background tissue; as well as emphasizing contour information. Initially, we trimmed out the pixels with extreme intensity values, which are highest and lowest 5% of intensity histogram. Although we eliminate extreme pixel intensities from ROI, there are still exceptional circumstances must be solved, like calcifications appearing in front of the masses. So, we applied a median filter with window size of % of the bigger ROI dimension (width or height) to ROI for eliminating undesirable situations. However, applying a median filter to a ROI may produce artificial contours. Hence, we applied an averaging filter with the same size of median filter to overcome false contour problem. Finally, a Laplacian edge enhancement filter with same size of previously applied filters is used to enhance edges of ROI to emphasize edge information.

In mass size estimation step, we estimate mass size empirically by using an intensity histogram based segmentation approach called OTSU. OTSU segmentation algorithm aims to find one or more split points on intensity histogram which separates whole intensity histogram into two or more groups whose intra-class variances are minimum [22]. In this step, we initially applied OTSU histogram segmentation approach to divide ROI into two groups namely large and small segment. Therefore, we could say that size of a segment in ROI should be larger than small segment and smaller than large segment. As a result, we obtained an estimation of size interval for real mass segment. Moreover, we used division point of intensity histogram found in this step as our initial threshold value in forth step.

BMCS is based on seeded region growing (SRG) segmentation algorithm [20], which requires a predefined seed points or seed area in an image and expands initial seed by adding neighbor pixels while they satisfy following condition.

\[ |\mu - I_{x,y}| < \theta \]

where \( \mu \) is the average intensity level of segmented region, \( I_{x,y} \) is the intensity level at point \((x, y)\) which is a neighbor of current segmented region and \( \theta \) is the intensity threshold. \( \mu \) value is recalculated when a new pixel is added to evolving region. We proposed some changes in SRG to apply breast mass segmentation problem. First, we select the brighter pixels in center of ROI as initial seed points, since we know that breast masses contain the brighter pixels than their neighbor pixels and there is a real mass in center of ROI. We also expand initial seed with neighbor pixels satisfying following condition.

\[ I_{x,y} > \theta \]

where \( I_{x,y} \) is the intensity of neighbor pixel to be added to current evolving region and \( \theta \) is the intensity threshold. Another crucial problem of SRG is finding a proper threshold value, because it is hard to find a universal threshold value for each
Therefore, we proposed an adaptive threshold modification approach to solve this problem. Initially, we select split point calculated in mass size estimation step as our preliminary threshold value. When no neighbor pixels left to satisfy expansion condition above, we check the size of the current evolving mass region whether its size is greater than size of small segment. If size of current region is shorter than small segment, we decrease threshold value and continue to expand region. Similarly, we stop segmentation process if size of the evolving region exceeds the size of the large segment. Consequently, result of our segmentation algorithm produces segments whose size stays in predefined interval. Furthermore, our approach proposes a solution for over- and under-segmentation problems. Finally, we dilate final segment using a sphere-shaped morphological filter to eliminate final edge artifacts. In Fig. 1 and Algorithm 1 details of BMCS is given.

Algorithm 1 Breast mass contour segmentation algorithm

// Input: ROI
// Output: Set Segmented Points
function BMCS(ROI)
  Width, Height ← Size of ROI
  H ← Intensity Histogram of ROI
  Eliminate Pixels of ROI with highest and lowest 5% of H
  Apply Median Filter To ROI with 5% of MAX(Width, Height)
  Apply Smoothing Filter To ROI with 5% of MAX(Width, Height)
  Apply Laplacian Filter to ROI with 5% of MAX(Width, Height)
  T ← OTSU division point of ROI
MinSize ← Count of Pixels whose intensity value above T
MaxSize ← Count of Pixels whose intensity value below T

if MaxSize<MinSize then
    Exchange Values of MaxSize and MinSize
End

MassRegion ← {} NeighbourPixels ← {}
CurrentPixel ← Pixel whose value is Maximum

while Intensity (CurrentPixel) > T && SizeOf (MassRegion)<MaxSize
    NeighbourPixels ← NeighbourPixels ∪ Neighbors of CurrentPixel
    MassRegion ← MassRegion ∪ CurrentPixel
    CurrentPixel ← Best of NeighbourPixels
    if Intensity (CurrentPixel) ≤ T
        if SizeOf (MassRegion) < MinSize
            T ← T - 1
        end
    end
end

Dilate ROI using a sphere-filter with size of 5% of MAX (Width, Height)

return MassRegion

3. Experiments

We run a set of experiments on reference dataset and compared with different region-based segmentation algorithms to evaluate the performance of BMCS. Each algorithm we tested is implemented in C++ by using QT UI Framework and OpenCV image processing library. However, for a successful quantitative comparison, we need to select suitable evaluation metrics at first. Thus, we first provide a survey on segmentation evaluation metrics.

3.1. Segmentation evaluation metrics

For a successful comparison, evaluation metrics are crucial for researchers to choose the right segmentation algorithm for their application needs or to compare their segmentation methods with other ones in the literature [23]. Segmentation evaluation methods for this purpose can be classified into two major groups similarly to the machine-learning algorithm. The first group is supervised methods, which needs previous region boundaries, and the other one is unsupervised methods that do not need previous region boundaries. Since we already have selected region boundaries, supervised evaluation metrics are more useful in our case.

3.2. Supervised evaluation measures

Supervised evaluation metrics measures accuracy of segmentation method by using conformance of machine segmented region (MSR) and reference region (RR) [24]. Conformance between MSR and RR is determined in two ways; area conformance and edge conformance. Similarly, we can classify supervised segmentation evaluation metrics into two groups; area-based metrics (ABM) and edge-based metrics (EBM).

ABM are based on number of pixels that how MSR and reference region intersects. The one of the oldest and most commonly used ABM has been proposed by Yasnoff et al. [25]. Yasnoff proposes segmentation error metric, Ev, which represents the ratio of misclassified pixels to the whole ROI. Ev is defined as follows:

\[
Ev = 1 - \frac{|B_s \cap M_s| + |M_s \cap B_s|}{|B_s + M_s|}
\]

where B and M denotes background and mass pixel sets, subscript S and R denotes segmentation and reference, respectively. By the way, the symbol |·| represents the element count operator over a set. As we know from classification evaluation metrics, second term in previous formula denotes accuracy. Hence formula could also be rewritten as follows:

\[
Ev = 1 - \frac{TP + TN}{Total} = 1 - \text{accuracy}
\]

where TP and TN denote successfully segmented mass and background pixels respectively. However, quality of accuracy measure is open to discuss, because of its weak predictive power. For instance, consider a ROI containing a very small mass only, a segmentation algorithm can achieve high accuracy rate if it assigns all pixels as the background. As a result, Ev of the segmentation algorithm indicates very small error rate, even though algorithm could find no region in the image. To overcome the weak predictive power problem of accuracy measure, Yasnoff proposed distance error (Dv) measure, which considers the spatial distributions of wrong pixels to regions.

The distance error measure, Dv, is defined as follows:

\[
D_v = \frac{100}{A} \sqrt{\frac{1}{k} \sum_{i=1}^{k} d_i^2}
\]

where A is the number of pixels in reference region and d_i is the distance between ith pixel of F_S ∩ B_R and its nearest neighbor in F_S. Rosa et al. [26] creates a confusion matrix of segmented pixels and calculates precision, recall and F-measure instead of accuracy to measure their calculi segmentation algorithm. As we know from classification evaluation, proposed metrics have been defined as follows.

\[
P = \frac{|F_s \cap F_G|}{|F_s \cap F_G| + |F_s \cap B_G|}
\]

\[
R = \text{sensitivity} = \frac{|F_s \cap F_G|}{|F_s \cap F_G| + |B_s \cap F_G|}
\]

\[
F = \frac{2 \cdot P \cdot R}{P + R}
\]

Additionally, following classification metrics could be used for segmentation evaluation.

\[
\text{Specificity} = \frac{|F_s \cap F_G|}{|B_s \cap F_G| + |B_s \cap B_G|}
\]

\[
\text{Balanced accuracy} = \frac{\text{Sensitivity} + \text{Specificity}}{2}
\]
Table 1 – Distribution of mass properties in DEMS.

<table>
<thead>
<tr>
<th>Description</th>
<th>Sample count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mass shape</td>
<td></td>
</tr>
<tr>
<td>Round</td>
<td>50</td>
</tr>
<tr>
<td>Lobular</td>
<td>57</td>
</tr>
<tr>
<td>Irregular</td>
<td>113</td>
</tr>
<tr>
<td>Oval</td>
<td>40</td>
</tr>
<tr>
<td>Mass contour</td>
<td></td>
</tr>
<tr>
<td>Circumscribed</td>
<td>98</td>
</tr>
<tr>
<td>Microlobulated</td>
<td>9</td>
</tr>
<tr>
<td>Obscured</td>
<td>31</td>
</tr>
<tr>
<td>Indistinct</td>
<td>49</td>
</tr>
<tr>
<td>Spiculated</td>
<td>73</td>
</tr>
<tr>
<td>Mass density</td>
<td></td>
</tr>
<tr>
<td>High-density</td>
<td>119</td>
</tr>
<tr>
<td>Equal density/isodense</td>
<td>111</td>
</tr>
<tr>
<td>Low density, but not</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6</td>
</tr>
<tr>
<td>fat-containing</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Fat-containing</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>radiolucent</td>
<td>24</td>
</tr>
</tbody>
</table>

Edge-based metrics (EBM) measures the segmentation error between two edges. Error metric between two edge point sets $A = \{a_1, a_2, \ldots, a_n\}$ and $B = \{b_1, b_2, \ldots, b_n\}$, where $a_i$ and $b_i$ is the ordered pair of $x$ and $y$ points, is named distance closest point (DCP) and defined as follows:

$$d(a_i, B) = \min_j ||b_j - a_i||$$

The Hausdorff distance metric [27] between two curves is maximum DCP between two curves. Formal definition of Hausdorff distance is given below.

$$h(A, B) = \max(\max_i d(a_i, B), \max_j d(b_j, A))$$

Hausdorff distance metric measures general edge conformance between two edge curves. Odet et al. [28] proposes two metrics to measure error of each over- and under-segmented edge pixels. These metrics are defined below.

$$UDI = \frac{1}{N_0} \sum_{i=0}^{N_0} \left( \frac{d_u(i)}{d_{TH}} \right)^n$$

where $N_0$ and $N_U$ are the number of over- and under-segmented edge pixels, $d_u$ and $d_i$ are the distances of over- and under-segmented pixels, $d_{TH}$ is the maximum allowed distance threshold, and $n$ is the scale parameter for distances below $d_{TH}$.

4. Reference dataset

In this work, we use DEMS dataset which is developed in Dokuz Eylul University Radio Diagnostics Department [29]. Dataset contains totally 485 cases obtained from University Hospital’s PACS system and, 265 of the cases include an abnormality. Each case has four images in standard modalities (i.e., one CC and one MLO view of left and right breast). Expert radiologist annotated each image in three phases. In the first phase, radiology expert selects contour of an abnormality roughly and then annotates the selection with the help of Mammography Annotation Ontology [30]. In second phase, expert reviews and fine-tunes the previously selected regions to provide more precise contour information. As a result, our evaluation dataset contains 260 masses whose distribution is given in Table 1.

We develop two annotation tools to help dataset preparation phase. The first tool is a general annotation application aiming to annotate each case by using a predefined ontology. Some additional capabilities like automatic report generation, finding similar cases, etc. is also embedded to the application. Second tool is for reviewing process. This tool shows only regions previously selected by first application in three views, original ROI view, original selection view, and review view. In Fig. 2 both annotation and review tools are shown. Modifications made by second tool are separately stored. So, expert can make any modification without losing original region data. A synchronization mechanism is also embedded to second tool to combine reviewed region data to original one. Both tools

Table 2 – Performance metrics of samples given in Fig. 3.

<table>
<thead>
<tr>
<th></th>
<th>Accuracy (%)</th>
<th>Balanced accuracy (%)</th>
<th>$E_p$ (%)</th>
<th>$D_p$ (%)</th>
<th>FPR (%)</th>
<th>TPR (%)</th>
<th>H</th>
<th>(UDI, ODI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BMCS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample 1</td>
<td>95.09</td>
<td>94.98</td>
<td>4.91</td>
<td>0.58</td>
<td>6.34</td>
<td>96.30</td>
<td>55</td>
<td>(0,0,0)</td>
</tr>
<tr>
<td>Sample 2</td>
<td>85.93</td>
<td>89.23</td>
<td>14.07</td>
<td>0.00</td>
<td>0.00</td>
<td>78.46</td>
<td>11.18</td>
<td>(0,0,0)</td>
</tr>
<tr>
<td>Sample 3</td>
<td>91.11</td>
<td>91.80</td>
<td>8.89</td>
<td>0.19</td>
<td>3.07</td>
<td>86.66</td>
<td>154.65</td>
<td>(0,0,0)</td>
</tr>
<tr>
<td>Level set segmentation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample 1</td>
<td>89.96</td>
<td>90.49</td>
<td>10.04</td>
<td>0.70</td>
<td>2.91</td>
<td>83.90</td>
<td>124.17</td>
<td>(0,0,0.21)</td>
</tr>
<tr>
<td>Sample 2</td>
<td>74.16</td>
<td>80.22</td>
<td>25.84</td>
<td>0.00</td>
<td>0.00</td>
<td>60.44</td>
<td>20.13</td>
<td>(0,0,0)</td>
</tr>
<tr>
<td>Sample 3</td>
<td>82.39</td>
<td>81.48</td>
<td>17.61</td>
<td>0.63</td>
<td>25.38</td>
<td>88.34</td>
<td>317.29</td>
<td>(0,0,0.10)</td>
</tr>
<tr>
<td>Wei’s segmentation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample 1</td>
<td>86.95</td>
<td>87.91</td>
<td>13.05</td>
<td>0.02</td>
<td>0.18</td>
<td>76.00</td>
<td>114.24</td>
<td>(0,0,0)</td>
</tr>
<tr>
<td>Sample 2</td>
<td>65.43</td>
<td>73.54</td>
<td>34.57</td>
<td>0.00</td>
<td>0.00</td>
<td>47.07</td>
<td>27.30</td>
<td>(0,0,0)</td>
</tr>
<tr>
<td>Sample 3</td>
<td>90.56</td>
<td>90.96</td>
<td>9.44</td>
<td>1.00</td>
<td>5.99</td>
<td>87.92</td>
<td>142.39</td>
<td>(0,0,0)</td>
</tr>
<tr>
<td>Watershed segmentation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample 1</td>
<td>70.11</td>
<td>72.35</td>
<td>29.89</td>
<td>0.03</td>
<td>0.17</td>
<td>44.86</td>
<td>154.56</td>
<td>(0,0.18,0)</td>
</tr>
<tr>
<td>Sample 2</td>
<td>72.26</td>
<td>78.76</td>
<td>27.74</td>
<td>0.00</td>
<td>0.00</td>
<td>57.53</td>
<td>36.80</td>
<td>(0.22,0)</td>
</tr>
<tr>
<td>Sample 3</td>
<td>52.00</td>
<td>57.65</td>
<td>48.00</td>
<td>0.00</td>
<td>0.00</td>
<td>15.30</td>
<td>633.54</td>
<td>(0.70,0)</td>
</tr>
</tbody>
</table>

Bold values indicate the best segmentation evaluation scores among all samples.
are built using QT Cross Platform UI Toolkit and can run on
Windows, Linux, and Mac OSX platforms.

4.1. Evaluation

We evaluated the performance of our segmentation approach,
BMCS, against to the three other region based segmentation
methods. The first method is proposed by Dominguez and
Nandi [13], which uses level-set segmentation algorithm with
Chan-Vese energy fields. It uses an image contrast enhance-
ment method and applies a median filter, since median
filtering is used in several mammographic mass detection
algorithms for reducing noise while preserving edge informa-
tion [15,21,31]. In this algorithm, we used manually segmented
region instead of using a predefined zero level. In other words,
level set segmentation algorithm is used to enhance or fine-
tune a roughly selected region. Finally, region that has the
largest area is labeled as mass.

The second method is proposed by Wei et al. [16], and uses
seeded region growing approach to enhance mass boundary.
Proposed method includes a preprocessing step including a
linear contrast enhancement method and median filtering to

Fig. 2 – Snapshots of the annotation tool (above) and review tool (below) used for marking mass boundaries.
enhance distinction between mass region and background. Finally, multi-seeded region growing method with empirically defined constant area threshold is used to find mass boundary.

The third method uses linear mapping and histogram equalization method for image enhancement. Then, watershed segmentation algorithm is applied on enhanced image using edge pixels for background region marker and centroid of highest valued pixels for foreground region marker. Fig. 3 depicts sample segmentation results of four methods and Table 2 has corresponding error metrics of each sample.

BMCS algorithm has the highest true positive rate, which means that segmentation result cover maximum area of manually selected region among other methods for the first sample. On the other hand, FPR rate of proposed segmentation method is the highest score for the same sample. But produced false pixels distance to the manual selections has no significant difference according to the Yasnooff error ($E_Y$) metric comparing to the other methods. Similarly, BMCS produces the closest segment to the reference segment. BMCS outperforms all the other methods in all metrics for second sample segmentation result. For last sample, level-set segmentation method and Wei’s segmentation method produces higher true positive values than our segmentation algorithm but they produce false positives at least two times higher than our method. Additionally, false positives produced by BMCS method have no statistical significance with level set and Wei’s segmentation methods. Moreover, balanced accuracy of our method outperforms other methods. This means that BMCS is good at finding true pixels adequately, while eliminating false pixels.

Overall performance of all methods is given in Table 3. Watershed segmentation method has the lowest false positive rate and $D_Y$ values. So we can infer that watershed method produces very accurate and consistent with manual selection. But watershed segmentation method has the lowest true positive rate and accuracy values, so watershed segmentation could not cover mass area as accurate as other methods. Additionally difference between BMCS and watershed in false positive rate is statistically insignificant ($p = 0.83$). Moreover, watershed segmentation approach obtains the highest Hausdorff distance value. On the other hand difference between BMCS and watershed segmentation in accuracy and $D_Y$ is statistically significant ($p=0$ for both cases). This means that BMCS and watershed segmentation algorithms produces same amount of false positive pixels but results obtained from Watershed segmentation method fits better to reference data. On the contrary, BMCS produces more accurate results that Watershed segmentation algorithm. These results indicate that our segmentation approach reduces false positive rate while producing high accuracy values comparing to watershed segmentation approach.

Level-set segmentation approach has the highest accuracy and true positive rates. But its false positive rate, $D_Y$, $H$ and (UDI, ODI) metrics are higher than BMCS. Statistically, difference between BMCS and Level-set segmentation is significant.

| Table 3 – Performance results of different segmentation methods on DEMS. |
|---------------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Method                          | Accuracy (%)    | Balanced accuracy (%) | $E_Y$ (%) | $D_Y$ | FPR (%) | TPR (%) | $H$ | (UDI, ODI) |
| BMCS                            | 80.04           | 83.15             | 19.96       | 1.73       | 10.86 | 77.16 | 95.99 | (0.0051, 0.0019) |
| Chanese levelset                | 81.15           | 82.47             | **18.85**   | **2.27**   | **14.45** | **79.38** | **100.07** | (0.0076, 0.0045) |
| Wei’s method                    | 78.22           | 78.51             | 21.78       | 2.60       | 20.97 | 78.00 | 118.76 | (0.0363, 0.0023) |
| Watershed segmentation          | 59.15           | 67.30             | 40.85       | **1.01**   | 10.51 | 45.10 | 188.48 | (0.1418, 0.0020) |

Bold values indicate the best segmentation evaluation scores.
in all metrics for 10% confidence interval. Hence, we can say that BMCS produces less false positive pixels with more suitable to reference data than level-set method, and produces more optimum segmentation result compared with level-set segmentation method.

Finally, BMCS outperforms other methods in terms of balanced accuracy, Hausdorff distance and scalable discrepancy metrics. So, we can say that our segmentation method produces acceptable mass segments with less possibility of producing over- and under-segmented masses.

5. Conclusion

In this work, we present a new algorithm, called breast mass contour segmentation, for breast mass boundary enhancement for a given ROI in mammograms. The approach is an extended version of classical seeded region growing algorithm with additional capability to dynamically adjust threshold value and proper stopping conditions for the size of segments to compensate for under and over segmentation.

In order to evaluate our method, we first provide an extensive summary of segmentation evaluation methods available in literature, discussing descriptive powers of each evaluation metric individually, since segmentation accuracy or simple overlapping rate is not enough to express segmentation performance. The performance of BMCS is tested against the leading three region-based segmentation approaches, namely, level-set segmentation, seeded region growing segmentation and watershed segmentation. In order to perform a comparison we have developed a data set containing 260 masses. We have showed that our BMCS mostly outperforms all the other methods tested. Evaluation results shows that proposed method minimizes false positive pixels while obtaining high accuracy rates and least distance values to reference region. In conclusion, our suggestion to classical region growing algorithm, that is using adaptive threshold value and two stopping condition to prevent under and over segmentation, works well and improves the segmentation results in terms of several metrics.

The shape, contour and margins of a breast mass involve valuable information for determining the severity of a breast mass. So it is important to find mass area as precisely and accurately as possible. Once breast mass area is segmented accurately, properties of segmented mass can be extracted and analyzed to determine mass malignancy. It is sure that the quality of the extracted mass properties directly depends on the success rate of the mass segmentation algorithms used. In future, we plan to further analyze the segmented mass area to extract more informative and more descriptive low-level features that can be used in CAD systems.

Acknowledgements

This work is supported by The Scientific and Technological Research Council of Turkey (TUBITAK) under project number 107E217.

REFERENCES


