Pixel-Based Morphological Technique for Breast Tumour Detection

Adepoju T. M., Ojo J. A., Omidiora E. O., and Olabiyisi O. S.

Abstract—Breast region segmentation is the process of splitting mammogram image into breast region and background to focus and limit the search for abnormality on the breast region without the effect of the background on the results. In addition, performance of existing Computer Aided Detection (CAD) systems for detection of malignant tumours in breast tissue have been limited by the methods of segmentation. Image segmentation is a multi-objective problem where multiple criteria must be considered for extraction of breast region. The developed segmentation technique in this paper considered intensity of pixel for image binarization (using Otsu thresholding) and shape for image boundary refinement (using mathematical morphological processes), to detect exact location of tumour in breast tissue. The developed technique was evaluated using Kappa agreement scale (Hit, Miss and Over-hit). A moderate value of 0.59 in Kappa agreement scale was achieved for the segmentation. The Two-stage Segmentation Technique is efficient to extract the locations of breast tumour with low level of false positive.

Key words—Segmentation, pixel, morphology, tumour, Hit, Kappa, benign and malignant.

1 INTRODUCTION

The World Health Organization’s International Agency for Research on Cancer estimated that more than 150,000 women worldwide die of breast cancer each year [15]. Mammography screening is the golden standard method for early detection of tumour in breast, but 10-30% of the tumour are missed by mammography because tumour are obscured by radiographically dense breast tissue [12]. Mammography is one of the commonly used methods to detect breast tumour but radiologist show variation in interpreting mammograms [6]. Mammography is a non-invasive screening tool recommended for young women who have symptoms of breast cancer or have a high risk of breast cancer, as well as for women older than 40 years even if there is no sign of the disease [2].

Current research works conducted in the area of breast cancer detection and classification focused on segmentation because of the fact that breast cancer is becoming the most common form of cancer disease of today’s female population.

Segmentation refers to the process of partitioning a digital image into multiple segments [13], to simplify and change the representation of an image into something that is more meaningful and easier to analyze. Accuracy of breast cancer in mammograms depend on the segmentation of images [10]. Since there is no general solution to the image segmentation problem; different algorithms often have to be combined in order to effectively solve an image segmentation problem[7].

Image segmentation is a multi-objective problem (multiple criteria must be considered for image segmentation such as intensity of pixels, texture, shape and colour) [7]. The most significant feature that indicates whether a mass is benign or malignant is its shape[9]. The shape can be round, oval, lobular, or irregular as indicated in Figure 1. Masses that are sharply defined (circumscribed oval and round masses) are usually benign. Masses with irregular shapes (faint jagged edge) and indistinct or spiculated margins have a higher likelihood of malignancy.

![Morphology of masses](http://www.jesr.org)

Fig. 1: Morphology of masses [9]
2 RELATED WORKS

Recently, several researchers have developed breast cancer segmentation model that involved combination of technique to locate breast cancer. [5] developed an algorithm for segmenting speculated masses based on pulse coupled neural networks (PCNN) in conjunction with fuzzy set theory. [4] used active contour model (ACM) based on self-organizing network (SON) to segment the ROI. This model explores the principle of isomorphism and self-organization to create flexible contours that characterizes the shapes in the image.

[1] segmented suspicious masses in polar domain. They used adaptive level set segmentation method (ALSSM) to adaptively adjust the border threshold at each angle in order to provide high-quality segmentation results. They extended their work in [1] using speculation segmentation with level sets (SSLS) to detect and segment speculated masses. In conjunction with level set segmentation they used Dixon and Taylor line operator (DTLO) and a generalized version of DTLO (GDTLO).

[16] employed a dual-stage method to extract masses from the surrounding tissues. Radial gradient index (RGI) based segmentation was used to yield an initial contour close to the lesion boundary location and a region-based active contour model was utilized to evolve the contour further to the lesion boundary.

[3] applied active contours to segmenting the pectoral muscle and localized dense tissues by using the maxima method. The textures of the located zones are analyzed through the co-occurrence matrix and the Haralick features to classify them in normal or abnormal tissues using the SVM.

[10] developed a multistage segmentation method to segment the mammograms based on watershed algorithm and level set method. They used watershed transform to provide a coarse and fast pre-segmentation, and used the resultant segmentation as the initial contour for the level set segmentation. In the combined algorithm, the segmentation results from the watershed were used as the input of the level set segmentation and the level set algorithm is used to refine the boundary of the segmented image.

3 TWO-STAGE SEGMENTATION TECHNIQUE

The developed algorithm considered intensity of pixels and shape as criteria for the two-stage segmentation. The acquired images from Mammographic Image Analysis Society (MIAS) were pre-processed using polygon approximation method. The pre-processed mammograms were pre-sorted into low density (fatty) and high density breast tissue images, following medical procedural approach. The flow chart in Figure 2 presents the two-stage segmentation technique based on Otsu thresholding for the image binarization and morphological erosion to refine the image shape.

Otsu threshold is a pixel based algorithm in which a threshold T was set to maximum intensity value of initial point of interest. The gray scale images were converted to binary images (background and object) by estimating the class probability as shown in Figure 2. The class mean of both the background and object is computed according to expression in Figure 2. The average of the class mean as shown in the Figure determined the calculated threshold \( T_c \) for the images. The binary images were obtained by comparing the calculated threshold with set threshold as presented in the flow chart. The shape of the object in the binary images were obtained by image erosion according to the expression in the flow chart. The irregularities or false lines connected to the boundary of the segmented images are eliminated by erosion process of mathematical morphology as stated in equation 1.

\[
\beta(A) = (A \oplus B) - ((A \oplus B) \ominus B)
\]

(1)

The segmentation algorithm stepped through different threshold to ascertain the threshold value with better binary image. The performance of the segmentation algorithm was evaluated by calculating and analysing, Kappa scale of agreements (Hit, Miss and Overhit) while confusion matrix is used to determine the accuracy of classifying the segmented mammograms into normal benign and malignant tumour. These are defined as follows:

i. Hit: denotes the ratio of correct segmentation.

\[
Hit = \frac{TP}{TP + FN}
\]

ii. Miss: denotes the ratio of missing segmentation.

\[
Miss = \frac{FN}{TP + FN}
\]

iii. Over Hit: denotes the ratio of false segmentation.

\[
Over\ Hit = \frac{FP}{TP + FN}
\]

iv. Kappa: is a measure of inter-observer agreement.

\[
Kappa = \frac{2 \times Hit}{2 \times Hit + Miss + OverHit}
\]

v. Accuracy: it is the fraction of correctly classified image with regard to all images of that ground truth class.

\[
Accuracy = \frac{\text{no of diagonal mammogram in the matrix}}{\text{Total number of mammogram}}
\]
value 0.900. The segmentation accuracy (0.59) at threshold
value 0.900 is preferred because when a threshold value is
reduced, more abnormality are usually located (sample,
2003). The moderate segmentation accuracy (0.590) was
considered as better Kappa agreement result when
compare with other result in the table, since the objective at
this stage is to locate abnormality.

Some of the segmented abnormal and normal images are
shown in Figure 3 and Figure 4 respectively. It was
observed that the images in Figure 3 contain abnormality,
showing that the mammograms are affected while the
images in Figure 4 appeared blank, showing that algorithm
for segmentation could not find any abnormality in the
mammograms. It was observed that the shape of the
segmented objects in Figure 3(a) and 3(b) is the exact shape
of the tumour in the acquired gray scale image when
benchmark with golden standard in the MIAS images.

Also, the images in the Figure 3(a) and 3(b) were classified
correctly into benign and malignant tumour base on the
shape appearance.

4 RESULTS AND DISCUSSION

The images obtained from pre-processing stage aids
identification of low density (fatty) breast tissue images that
were used in the two-stage segmentation analysis. The results
obtained at the segmentation stage is presented in Table 1. The
metrics used to obtain these results are True Positive (TP), True
Negative (TN), False Positive (FP) and False Negative (FN). The
two-stage segmentation algorithm stepped through different
threshold values starting from 0.900. It was observed that at
0.900, 0.517 of the mammograms were segmented correctly,
0.429 of the images missed the exact location of abnormality
while 0.381 of the images include other pixels besides the
abnormality pixels. It was observed that the segmentation
accuracy of the mammogram at threshold value of 0.900 with
false positive (8), was 0.590 which fall into moderate
segmentation in kappa agreement scale. The segmentation
accuracy (0.500, 0.470 and 0.390) obtained at the other threshold
values are lower than the segmentation accuracy at threshold

Fig. 2. Flow chart of the two-stage segmentation technique

![Flow chart of the two-stage segmentation technique](image)

![Matlab GUI for Two-stage segmentation of a normal mammogram](image)

![Matlab GUI for Two-stage segmentation of a benign mammogram](image)
<table>
<thead>
<tr>
<th>Metrics</th>
<th>Results</th>
<th>Threshold values</th>
<th>Hit</th>
<th>Miss</th>
<th>Overhit</th>
<th>Kappa Value</th>
<th>Actual Class</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>TP</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F P</td>
<td>8</td>
<td>0.900</td>
<td>0.571</td>
<td>0.429</td>
<td>0.381</td>
<td>0.500</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>F N</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TP</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F P</td>
<td>6</td>
<td>0.915</td>
<td>0.571</td>
<td>0.286</td>
<td>0.500</td>
<td></td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>F N</td>
<td>22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TP</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F P</td>
<td>6</td>
<td>0.925</td>
<td>0.571</td>
<td>0.286</td>
<td>0.500</td>
<td></td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>F N</td>
<td>22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TP</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F P</td>
<td>5</td>
<td>0.935</td>
<td>0.619</td>
<td>0.238</td>
<td>0.470</td>
<td></td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>F N</td>
<td>23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TP</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F P</td>
<td>4</td>
<td>0.945</td>
<td>0.714</td>
<td>0.390</td>
<td></td>
<td></td>
<td>B</td>
<td></td>
</tr>
<tr>
<td>F N</td>
<td>24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TP</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F P</td>
<td>4</td>
<td>0.950</td>
<td>0.714</td>
<td>0.390</td>
<td></td>
<td></td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>F N</td>
<td>24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Two-stage segmentation results

5 CONCLUSION

In conclusion, The inter-observer agreement between the radiologist and the proposed segmentation algorithm, at a given threshold of 0.9 was at moderate level (0.59) of the kappa scale, has in effect improved the detection of abnormality at exact location.

Fig. 4. Matlab GUI for Two-stage segmentation and classification of a normal mammogram
REFERENCES


