Separation of Malignant and Benign Masses using Maximum-Likelihood Modeling and Neural Networks

Lisa Kinnard[a,b,c], Shih-Chung B. Lo [a], Paul Wang[c], Matthew Freedman[a], Mohamed Chouikha[b]

[a]ISIS Center, Department of Radiology, Georgetown University Medical Center, Washington, D.C.
[b]Department of Electrical Engineering, Howard University, Washington, D.C., USA
[c]Biomedical NMR Laboratory, Department of Radiology, Howard University, Washington, D.C.

ABSTRACT

This study attempted to accurately segment the masses and distinguish malignant from benign tumors. The masses were segmented using a technique that combines pixel aggregation with likelihood analysis. We found that the segmentation method can delineate the tumor body as well as tumor peripheral regions covering typical mass boundaries and some spiculation patterns. We have developed a multiple circular path convolution neural network (MCPCNN) to analyze a set of mass intensity, shape, andtexture features for determination of the tumors as malignant or benign. The features were also fed into a conventional neural network for comparison. We also used values obtained from the maximum likelihood values as inputs into a conventional backpropagation neural network. We have tested these methods on 51 mammograms using a grouped Jackknife experiment incorporated with the ROC method. Tumor sizes ranged from 6mm to 3cm. The conventional neural network whose inputs were image features achieved an A_z value of 0.66. However the MCPCNN achieved an A_z value of 0.71. The conventional neural network whose inputs were maximum likelihood values achieved an A_z value of 0.84. In addition, the maximum likelihood segmentation method can identify the mass body and boundary regions, which is essential to the analysis of mammographic masses.

Keywords: Computer-assisted diagnosis, breast cancer, convolution neural networks, feature extraction

1. INTRODUCTION

While many breast cancer diagnostic systems have been developed, fully-automated mass segmentation continues to be a major challenge in this area. Several investigators exploited methods using intensity values to decide if a pixel should be placed in the region of interest (ROI) or background [14,9,5,7]. Petrick[12] et al. developed the density weighted contrast enhancement (DWCE) method which applies a series of filters to the image in an attempt to extract masses. Li[6] et al. developed a competetitive classification strategy, which uses a combined soft and hard classification method for deciding if segmented regions are true or false positives. Li[7] et al. developed a segmentation method that uses probability to determine segmentation contours. Most of these methods are successful at segmenting the tumor body, however, they sometimes do not properly obtain the extended boundaries of the tumor. While conventional region-growing is an excellent pixel-based segmentation method, it may not suitable to use this method alone. It produces many segmentation contours for one tumor image, but does not decide which segmentation contour is the best. Based on the above reasons, we have developed a tumor segmentation method that combines region-growing with probability assessment to determine final segmentation contours for various breast tumor images.

The most recognized obstacles in breast cancer diagnosis are (1) difficulties of diagnostic decision making in calling back patient for further breast examination, (2) the large number of suspected lesions of which only

Further author information: (Send correspondence to Lisa M. Kinnard)
Lisa M. Kinnard: E-mail: kinnard@isis.imac.georgetown.edu, Telephone: 1 202 687 5135
S.C. Ben Lo: E-mail: lo@isis.imac.georgetown.edu, Telephone: 1 202 687 1659,
Address: ISIS, Georgetown University, 2115 Wisconsin Avenue, NW, Washington DC, USA

part of them are malignant lesions; and (3) missed diagnosis of breast cancer. The callback rates vary from 5% to 20% in today's breast cancer screening programs[1,16]. At some medical centers, the positive predictive rate can be 30% to 35%[4,1]while at others this rate can be as low as 10% to 15%.It is well known that effective treatment of breast cancer calls for early detection of cancerous lesions (e.g., clustered microcalcifications and masses associated with malignant cellular processes)[16,11,15] Tumors can be missed because they are obscured by glandular tissue and it is therefore difficult to observe their boundaries. We were motivated by this clinical obstacle and have developed a computer-assisted diagnostic system to serve as a consultant/good second opinion.

2. METHODS

Computer-assisted breast cancer diagnosis is divided into three parts, namely, image segmentation, feature calculation, and classification. The next several section will theoretically describe the methods used in the study.

2.1. Segmentation

It is well known that lesion segmentation is one of the most important aspects of computer-assisted diagnosis (CAD_x) because one of the main characteristics of malignant tumors is ill-defined, and/or spiculated borders. Conversely, benign tumors typically have well-defined, rounded borders. Segmentation is therefore extremely important because the diagnosis of a tumor can strongly depend upon image features.

Pixel aggregation is an automated segmentation method in which the region of interest begins as a single pixel and grows based on surrounding pixels with similar properties, e.g., grayscale level or texture.[2] It is a commonly used method[13,14,9]due to its simplicity and accuracy. The computer will use the maximum intensity as the "seed point" -a pixel that is similar to the suspected lesion and is located somewhere inside the suspected lesion. The next 4- or 8-neighboring pixel is checked for similarity so that the region can grow. If pixels in the 4- or 8-neighboring region are similar, they are added to the region. The region continues to grow until there are no remaining similar pixels that are 4- or 8-neighbors of those in the grown region.

Our implementation of this method checks the 4-neighbors of the seed pixel and uses a graylevel threshold as the similarity criterion. If a 4-neighbor of a pixel has an intensity value greater than or equal to a set threshold, it is included in the region of interest. The 4-neighbors were checked instead of the 8-neighbors so that surrounding tissue will not be included. The intensity threshold was used as a similarity criterion due to its simplicity and effectiveness.

By using the same seed point with multiple intensity threshold values we obtained between 150 and 300 of gray level change per lesion; however, the computer did not have the ability to choose the best partition. We added a maximum-likelihood component to the region-growing algorithm. The algorithm can be summarized in five steps. The image was first multiplied by a 2D shadow, whose size was approximately the same size as the ROI. We will henceforth refer to the image to which the 2D shadow has been applied as the "fuzzified" image. We started the threshold value at the maximum intensity in the image and decreased the intensities in successive steps. Consequently, we obtained a sequence of growing contours (S_i), where intensity value was the similarity criterion. There was an inverse relationship between intensity value and contour size, i.e., the lower the intensity value, the larger the contour. Next, we calculated the composite probability (P_i) for each contour (S_i):

$$P_i = p(S_i|pdf_i)xp(_{outside}S_i|ROI). \tag{1}$$

where $p(S_i|pdf_i)$ is the probability density function (pdf) of the ROI subject to the fuzzified image (see Fig. 1). This pdf is calculated $inside$ the contour, S_i, where i is the thresholding step. The quantity $p(outsideS_i|ROI)$ is the pdf of the ROI subject to the original image. This pdf is calculated $outside$ the contour, S_i. Next we find the logarithm of the composite probability, P_i in the following way:

$$log(P_i) = log(p(S_i|pdf_i)) + log(p(outsideS_i|ROI)), \tag{2}$$

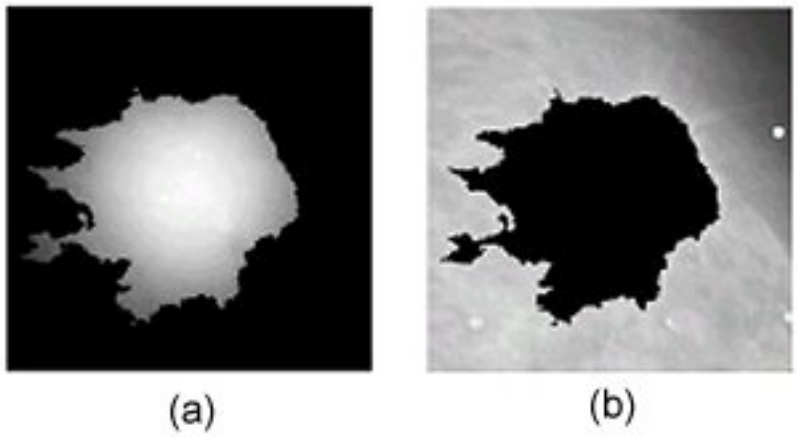


Figure 1: Figure (a) is used to calculate $p(S_i|pdf_i)$. Figure (b) is used to calculate $p(outsideS_i|ROI)$

Finally, we determine the likelihood that the contour represents the tumor body by assessing the maximum likelihood function:

$$argmax(Log(P_i)), \tag{3}$$

Equation 3 intends to find the maximum value of the aforementioned likelihood values as a function of intensity threshold. We assess (so as other investigators[5]) that the intensity value corresponding to this maximum likelihood value is the optimal intensity for the tumor body contour. We also determine the likelihood that the contour represents the tumor extended borders by assessing the maximum change of the likelihood function:

$$argmax(\frac{dLog(P_i)}{di}), \tag{4}$$

i.e., find the steepest jump on the aforementioned function. An intensity value between this jump and the maximum value on the function produces the best contour of the tumor body and its extended borders.

2.2. Feature Calculation

One extremely important task in the separation of malignant and benign tumors is feature selection and calculation. Benign tumors can be lucent at the center and can have well-defined borders; while malignant tumors can have spiculated and/or fuzzy borders. We used the following features:

Global Features

$$Skewness = \frac{1}{N}\frac{\sum_{i,j=0}^{N-1}[g(i,j) - \overline{g(i,j)}]^3}{\sqrt{\sum_{i,j=0}^{N-1}[g(i,j) - \overline{g(i,j)}]^3}} \tag{5}$$

where $g(i,j)$ is intensity value and $\overline{g(i,j)}$ is average intensity value.

$$Kurtosis = \frac{1}{N}\frac{\sum_{i,j=0}^{N-1}[g(i,j) - \overline{g(i,j)}]^4}{\sqrt{\sum_{i,j=0}^{N-1}[g(i,j) - \overline{g(i,j)}]^4}} \tag{6}$$

$$Circularity = \frac{A_1}{A}, \tag{7}$$

where A is the area of the actual ROI; A_1 is the area of the overlapped region of A and the effective circle A_c, which is defined as the circle whose area is equal to A and is centered at the corresponding centroid of A.

$$Compactness = \frac{p^2}{a}, \tag{8}$$

where, p=tumor perimeter and a=tumor area

$$perimeter = tumor\ perimeter. \tag{9}$$

Local Features

These intensity features were calculated on the 10^o ROI as it was divided into 10^o sectors in the polar coordinate system, therefore each tumor contained 36 sectors.

$$\overline{g(i,j)} = \frac{1}{N}\sum_{i,j=0}^{N-1} g(i,j), \tag{10}$$

where Mean = $\overline{g(i,j)}$, N is the total pixel number inside the ROI

$$Contrast = \frac{P_f - P_b}{P_f}, \tag{11}$$

where P_f is the average gray-level inside the ROI's and P_b is the average gray-level surrounding the ROI.

$$\sigma_f^2 = \frac{1}{N}\sum_{i=1}^{N}(g(i,j) - \overline{g(i,j)})^2, \tag{12}$$

where σ_f^2 = standard deviation.

$$Area = tumor\ area \tag{13}$$

$$\sigma_n = \frac{1}{N_b}\sum_{i=1}^{N_b}(r_i - \bar{r})^2, \tag{14}$$

where σ_n = Deviation of the Normalized Radial Length, N_b is the total number of pixels located on the boundary of the ROI, r_i is the value of the normalized radial length from the boundary coordinate (x_i, y_i) to the centroid of the ROI; $\bar{r}$ is the mean of r_i.

$$Roughness = ([\frac{1}{N_b}\sum_{i=1}^{N_b}(r_i - \bar{r})^4]^{\frac{1}{4}} - [\frac{1}{N_b}\sum_{i=1}^{N_b}(r_i - \bar{r})^2]^{\frac{1}{2}}])/\bar{r}. \tag{15}$$

$$radial\ length = length\ of\ radius, \tag{16}$$

where *length of radius* is the distance from the center of the tumor to its edge.

Given a second-order joint probability matrix $P_{d,\theta}(i,j)$, where $P_{d,\theta}(i,j)$ is the joint gray level distribution of a pixel pair (i,j) with the distance d and in the direction θ, six texture features are defined as follows:

$$E_{d,\theta}(i,j) = \sum_{i=1}^{L}\sum_{j=1}^{L} P_{d,\theta}(i,j)^2, \tag{17}$$

where $E_{d,\theta}(i,j)$ = energy.

$$I_{d,\theta}(i,j) = \sum_{i=1}^{L}\sum_{j=1}^{L}(i-j)^2 P_{d,\theta}(i,j), \tag{18}$$

where $I_{d,\theta}(i,j)$ = inertia.

$$E = \sum_{i=1}^{L}\sum_{j=1}^{L} P_{d,\theta}(i,j) log_2 P_{d,\theta}(i,j), \tag{19}$$

where E = entropy.

$$IDM_{d,\theta} = \sum_{i=1}^{L}\sum_{j=1}^{L}\frac{1}{1+(i-j)^2} P_{d,\theta}(i,j), \tag{20}$$

where, $IDM_{d,\theta}$ = Inverse Difference Moment.

$$DE_{d,\theta} = -\sum_{k=0}^{n-1} P_{x-y}(k) log_2 P_{x-y}(k), P_{x-y}(k) = \sum_{i=0}^{n-1}\sum_{j=0}^{n-1} P_{d,\theta}(i,j), \tag{21}$$

for $|i-j| = k, k = 0, 1, ..., n-1$ where, $DE_{d,\theta}$ = Difference Entropy.

2.3. Classifiers

We used a conventional backpropagation neural network for two of the three studies described in this paper. It is comprised of an input layer, one hidden layer, and one output. We used a Multiple Circular Path Neural Network[8] for the third study described in this paper. It is comprised of 3 input layers, one hidden layer and one output. The first input layer is fully connected, i.e., all inputs connect to all hidden nodes. The second input layer is called a self correlation path, i.e., each node on the layer connects to a single set of the 18 image features for the fan-in and fully connects to the hidden nodes for fan-out. The third input layer is called a neighborhood correlation path, i.e., each node on the layer connects to the input nodes of adjacent sectors for the fan-in and fully connects to the hidden nodes for fan-out. A more detailed explanation of the MCPCNN can be found in reference.[8] Our study used 18 hidden layer nodes.

3. EXPERIMENT

The image samples were chosen from several databases compiled by the ISIS Center of the Georgetown University (GU) Radiology Department and the University of Florida's Digital Database for Screening Mammography (DDSM).[3] They are a mixture of "obvious" cases and "not obvious" cases. The "obvious" cases contain tumors that are easily identifiable as malignant or benign while the "not obvious" cases are those that radiologists find difficult to observe and/or classify. Forty malignant and forty benign tumors were tested during this experiment. The GU films were digitized at a resolution of 100μm using a Lumiscan digitizer. The DDSM films were digitized at 43 and 50 μm's using both the Lumiscan and Howtek digitizers. We compensated for this difference in resolution by reducing the DDSM images to half their normal sizes. The images were of varying contrasts and the tumors were of varying sizes. There were 28 malignant cases and 23 benign cases.

The experiment was subdivided into three studies as shown in table 1 below.

Experiments 1 and 2 used 6 global and 12x36 sector features to yield a total of 438 image features per tumor. There were 18 hidden nodes and 1 output for both the BP and MCPCNN classifiers. The training and testing method used was the jackknife method. Experiment 3 used 19 likelihood feature values per tumor. There were 15 hidden nodes and 1 output for the BP classifier. The training and testing method used was the jackknife method. The results were analyzed using the LABROC4 program.[10]

Experiment	Features	Neural Network
1	Image Features	Conventional NN
2	Image Features	MCPCNN
3	ML-curve as features	Conventional NN

Table 1: This table summarizes the studies presented in this paper.

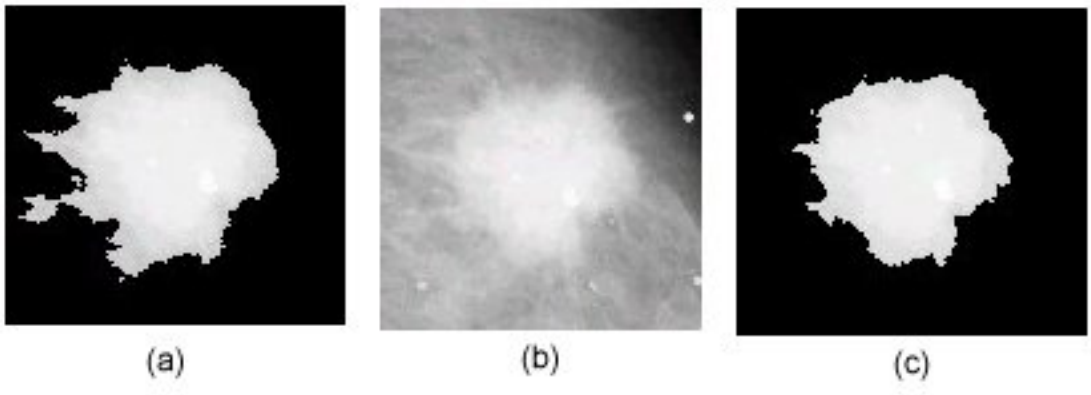


Figure 2. The segmentation results for a malignant tumor. Part (a) shows the segmentation result produced by the maximum likelihood change intensity choice, part (b) shows the original image, and part (c) shows the segmentation result produced by the maximum likelihood intensity choice.

4. RESULTS

Here are two examples of segmentation results for both malignant (see Fig. 2) and benign (see Fig. 4) cases. Each example gives the segmentation result produced by the maximum likelihood value on the curves described in section 2.1.

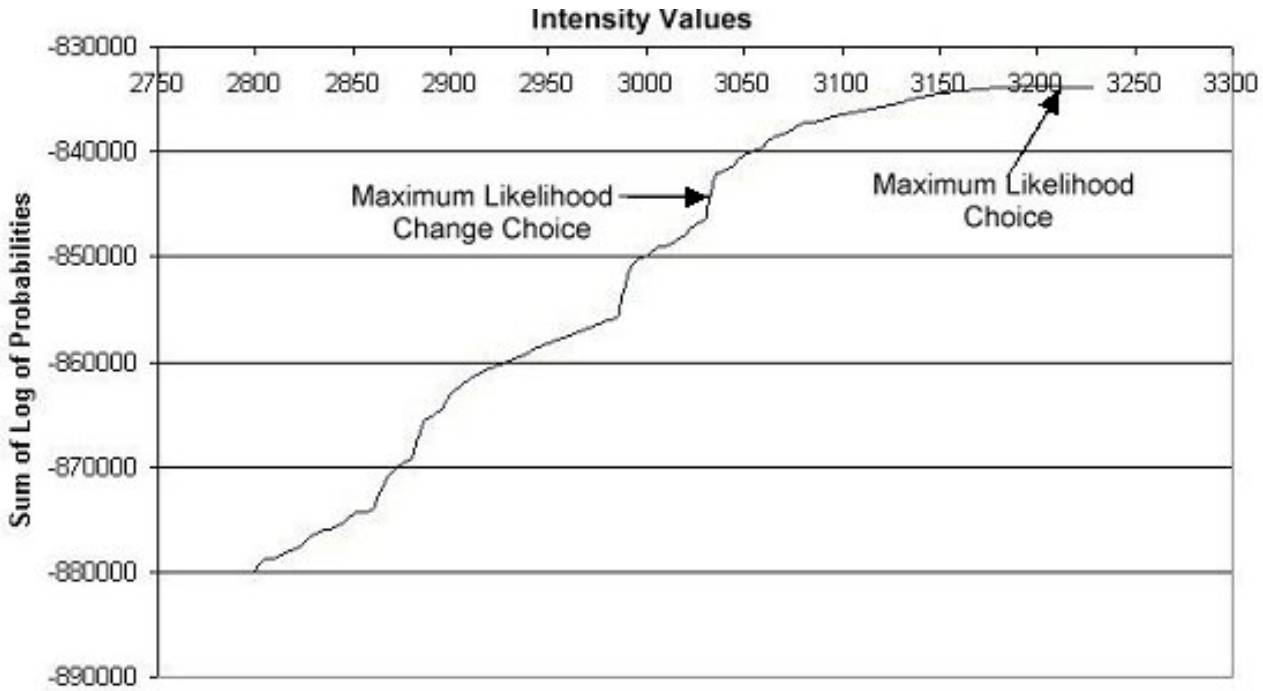
The following is a table, which gives the A_z values produced by the neural network.

Experiment	Features	Neural Network	Az
1	Image Features	Conventional NN	0.66
2	Image Features	MCPCNN	0.71
3	ML-curve as features	Conventional NN	0.84

Table 2: Results from Experiments 1-3.

5. CONCLUSION AND DISCUSSION

In analyzing the segmentation results we drew several conclusions. We discovered that there was a marked difference between the likelihood functions in malignant cases and the likelihood functions in benign cases. The likelihood function in the benign case often experiences a sharp drop, while the likelihood function in the malignant case is often smoother. In the image, a sharp drop value in the likelihood function represents an abrupt change in the area as well as likelihood value. We observed thatin benign cases, the likelihood function sharp changes are much more evident because benign tumors usually have well-defined borders. Conversely, in many malignant cases, the likelihood functions are smoother because many of their the borders are ill-defined. In



..

Figure 3. A likelihood function with respect to threshold values for all segmentation steps (malignant case) shown in Fig. 2.

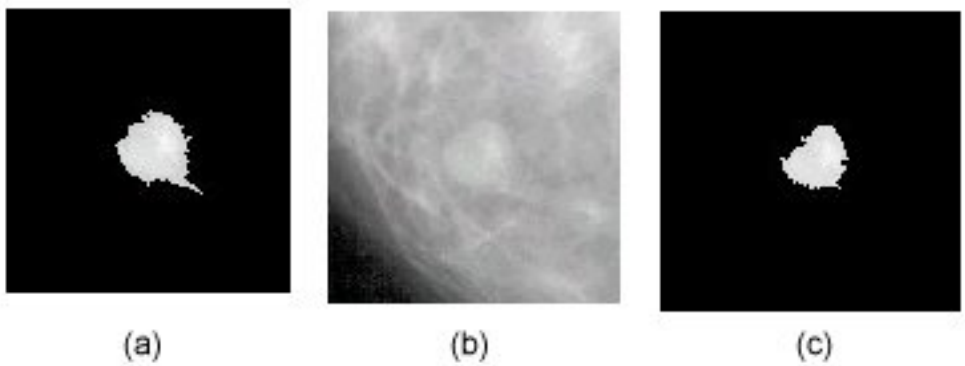
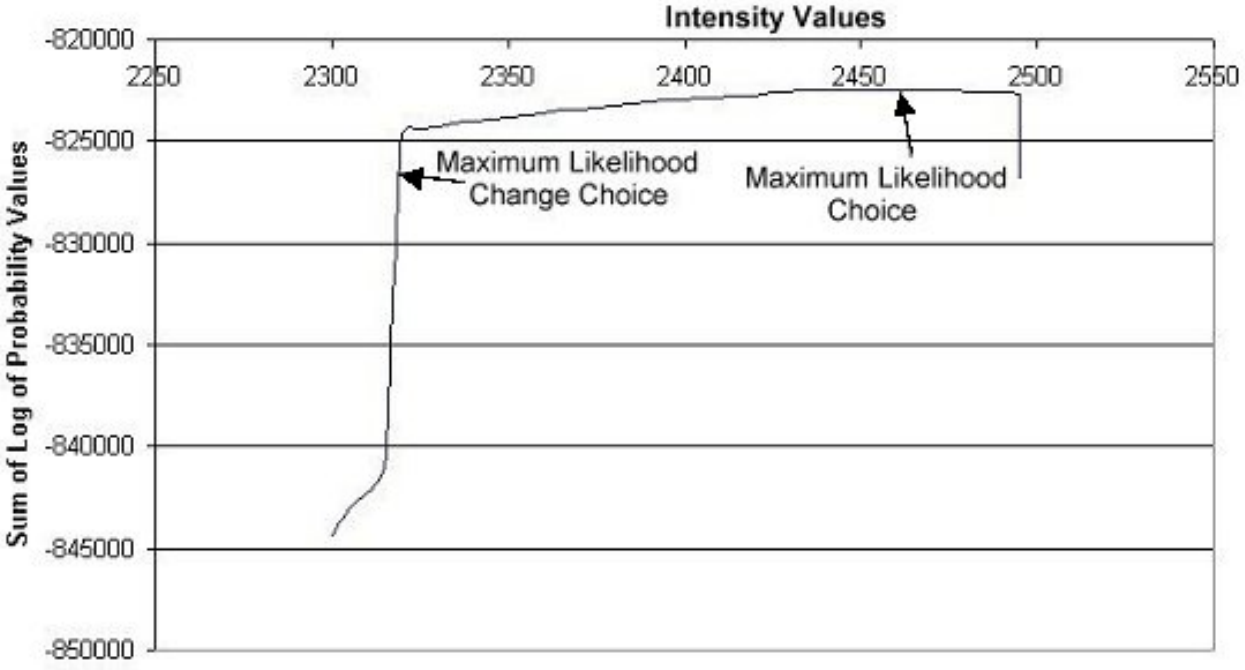


Figure 4. The segmentation results for a benign tumor. Part (a) shows the segmentation result produced by the maximum likelihood change intensity choice, part (b) shows the original image, and part (c) shows the segmentation result produced by the maximum likelihood intensity choice.

analyzing the likelihood functions for malignant cases we recognized that those curves with very sharp changes were produced from tumors with well-defined borders and vice versa; i.e., there were malignant tumors that could be mistaken as benign and vice versa.

The maximum likelihood curves used as inputs to the BP neural network produced the best performance overall. The image features used as inputs to the MCPCNN produced the second best performance. The image features used as inputs to the BP produced the worst performance. Since we received the best results by using the likelihood functions as features, we expect that the MCPCNN may improve the overall results by giving



..

Figure 5. HA likelihood function with respect to threshold values for all segmentation steps (benign case) shown in Fig. 4.

the likelihood functions in every sector.

ACKNOWLEDGMENTS

This work has been supported by the following grants: DAMD17-00-1-0291, DAAG55-98-1-0187, and DAMD17-00-1-0267.

REFERENCES

1. Frankel SD, Sickel EA, Curpen BN, Sollito RA, Ominsky SH, Galvin HB, *Initial versus subsequent screening mammography: Comparison of findings and their prognostics significance.* AJR, 1995, vol. 164, pp. 1107-1109.
2. Gonzalez RC, Woods RE. Digital Image Processing Reading, MA: Addison Wesley, 1992.
3. Heath M, Bowyer KW, Kopans D et al, *Current status of the Digital Database for Screening Mammography*, Digital Mammography, Kluwer Academic Publishers, 1998, pp. 457-460.
4. Kopans DB. *The positive predictive value of mammography*, AJR, 1991, vol. 158, pp. 521-526.
5. Kupinski MA, Giger ML, *Automated Seeded Lesion Segmentation on Digital Mammograms*, IEEE Transactions on Medical Imaging, 1998, vol. 17, no. 4, pp. 510-517.
6. Li L, Zheng Y, Zhang L, Clark R, *False-positive reduction in CAD mass detection using a competitive classification strategy*, Medical Physics, 2001, Vol. 28, no. 2, pp. 250-258.
7. Li H, Wang Y, Liu KJR, Lo S-C, Freedman MT, *Computerized Radiographic Mass Detection - Part I: Lesion Site Selection by Morphological Enhancement and Contextual Segmentation, IEEE Transactions on Medical Imaging*, 2001, vol. 20, no. 4, pp. 289-301.
8. Lo SC, Li H, Wang J, Kinnard L, Freedman MT, A Multiple Circular Path Convolution Neural Network System for Detection of Mammographic Masses, IEEE Transactions on Medical Imaging, 2002, vol. 21, No. 2, (Accepted for publication).

9. Mendez AJ, Tahoces PG, Lado MJ, Souto M., Vidal JJ, *Computer-aided diagnosis: Automatic detection of malignant masses in digitzed mammograms*, Medical Physics, 1998, vol. 25, no. 6, pp. 957-964.
10. Metz C, LABROC Program, ftp://radiology.uchicago.edu/roc.
11. Nystrom L, Rutqvist LE, Wall S, Lindgren A, Lindqvist M, Ryden S, et. al., *Breast cancer screening with mammography: Overview of Swedish randomized trials*, Lancet, 1993, vol. 341, pp. 973-978.
12. Petrick N, Chan H-P, Sahiner B, Wei D, *An Adaptive Density-Weighted Contrast Enhancement Filter for Mammographic Breast Mass Detection, IEEE Transactions on Medical Imaging*, 1996, vol. 15, no. 1, pp. 59-67.
13. Pohlman S, Powell KA, Obuchowski NA, Chilcote WA, Grundfest-Broniatowski S, "Quantitative classification of breast tumors in digitized mammograms", Medical Physics, 1996, vol. 23, no. 8, pp. 1336-1345.
14. Sahiner B, Chan HP, Wei D, Petrick N, Helvie MA, Adler DD, Goodsit MM, *Image feature selection by a genetic algorithm: Application to classification of mass and normal breast tissue*, Medical Physics, 1996, vol. 23, no. 10, pp. 1671-1684.
15. Shapiro S, *Screening: Assessment of current studies*, Cancer, 1994, vol. 74, pp.231-238.
16. Tabar L, Fagerberg G, Duffy S, Day NE, Gad A, Grontoft O. *Update of the Swedish two-country program of mammographic screening for breast cancer*, Radiology Clinics of North America: Breast Imaging - Current Status and Future Directions, 1992, vol. 30, pp. 187-210.