Computer-Aided Detection and Classification of Masses in Digitized Mammograms

Mohammed Jahirul Islam

November 20, 2009
Presentation Outline

• Introduction
• Breast Cancer Statistics
• Motivation, Objectives
• Principal Stages of Breast Cancer Detection
• Literature Survey
  – Image Enhancement
  – Image Segmentation
  – Feature Extraction and Selection
  – Classification
• Proposed Method of Computer-Aided Diagnosis (CAD) System
• Simulation Results and Performance Evaluation
• Conclusions
Introduction

- Breast cancer continues to be a public health problem in the world.

- Breast cancer is the second leading cause of death by disease in Canada for women, after lung cancer.
Breast Cancer Statistics

• **Breast Cancer Statistics in 2009***
  – an estimated 22,700 Canadian women will be diagnosed with breast cancer and 5,400 will die from it.
  – An estimated 170 men will be diagnosed with breast cancer and 50 will die of it.
  – 1 in 9 women (11%) is expected to develop breast cancer during her lifetime (by age 90) and one in 28 will die from it.

• **Early detection of breast cancer, allowing treatment at an earlier stage, can significantly reduce breast cancer mortality.**

*Source: Canadian Cancer Society / National Cancer Institute of Canada; Canadian Cancer Statistics 2009, Toronto, Canada*
Motivation

• Mammography has been one of the most reliable methods for early detection of breast carcinomas.

• X-ray mammography is currently considered the “gold standard” for breast cancer diagnosis.

• However, it is difficult for radiologists to provide both accurate and uniform evaluation for the enormous mammograms.

• The estimated sensitivity of radiologists in breast cancer screening is only about 75% [80]

• Computer-aided diagnosis (CAD) system can be used as a second opinion to aid the radiologist by indicating the locations of suspicious abnormalities in mammograms
Objectives

- Develop a CAD system for breast cancer diagnosis and detection based on automated segmentation of masses in mammograms.

- The ultimate diagnosis of all types of breast disease depends on a biopsy. In most cases the decision for a biopsy is based on mammography findings.

- Biopsy results indicate that 65-90% of suspected cancer detected by mammography turned out to be benign [81]

- The objective of the automated methods for classifications is to provide a tentative diagnosis (the final decision is produced by human expert) of individual masses, based on their physical attributes.
Challenges in Breast Cancer Detection

- Microcalcifications and Masses are two important early signs of breast cancer.

- Masses are often indistinguishable from the surrounding parenchyma because their features can be obscured or similar to the normal inhomogeneous breast tissues.

- This makes the automatic mass detection and classification challenging
Principal Stages of Breast Cancer Detection

1. Digital Mammograms
2. Image Enhancement (MF, HE)
3. Image Segmentation
4. Feature Extraction and Selection
5. Classification
   (Mass, Calcification- Benign, Malignant)
Sample Mammograms

Image 1: Normal Breast

Image 2: Cancerous Breast
Survey Over Image Enhancement Techniques

- **Global approach (HE) [33-36]**
  - Reassign the intensity values of pixels to make the new distribution of the intensities uniform to the utmost extent
  - Effective in enhancing the entire image with low contrast
  - Can not enhance the textural information.
  - Working only for the images having one object

- **Local Approach [33, 37-38]**
  - Feature-based or using non-linear mapping locally (Median filtering)
  - Effective in local texture enhancement
  - Can not enhance the entire image well
Survey Over Image Segmentation Techniques

• **Image Segmentation** - Recognize homogeneous regions within an image as distinct and belonging to different objects.

• The segmentation process can be based on finding the maximum homogeneity in grey levels within the regions identified. The segmentation doesn’t perform well if the grey levels of different objects are quite similar [M.A. Sid-Ahmed]

• **Global Thresholding** [39, 40, 43]
  – Based on global information, such as histogram of the mammograms
  – Widely used, easy to implement
  – Not good for identifying ROIs
  – FNs and FPs may be too high
Survey Over Image Segmentation Techniques

• **Local Thresholding [7, 41, 42, 44]**
  – Thresholding value is determined locally
  – It can refine the results of global thresholding, and is better for mass detection than global thresholding
  – It can not accurately separate the pixels into suitable sets. It is often used as an initialization of other algorithms

• **Edge Detection [35, 43, 48-53]**
  – Traditional method for image segmentation and it detects the discontinuity in mammograms

• **Template Matching [54-59]**
  – Segments possible masses from the background using prototypes
  – Easy to implement; if the prototypes are appropriate, it can provide good results
  – It depends on the prior information of the masses, it may result high number of false positives
Survey Over Image Segmentation Techniques

• **Region Growing [44-47]**
  – Finds a set of seed pixels first, then grow iteratively and aggregate with the pixels that have similar properties

• **Bilateral Image Subtraction [60-66]**
  – It is based on the normal symmetry between the left and right breast
  – Easy to implement, and the difference between the left and right mammogram images can be identified as suspicious regions
  – It is difficult to register the left and right breast correctly

• **Fuzzy Techniques [67-71]**
  – The fuzzy techniques including fuzzy thresholding and fuzzy region growing; it can handle the unclear boundary between normal tissue and tumors
  – It is not easy to determine suitable membership functions and rules
Survey Over Features Extraction and Selection

• Features extraction and selection is a key step in mass detection and classification

• Features are calculated from the region of interest (ROI) characteristics such as size, shape, density and smoothness etc. [72]

• Feature space is very large and complex due to wide diversity of the normal tissues and the variety of the abnormalities.

• Feature space can be divided into 3 sub-spaces [73]
  – Intensity features
  – Shape features
  – Texture features
Haralick Texture Features- GLCM

- The basis for Haralick features [79] is the Gray-Level Co-occurrence Matrix (GLCM) or Spatial Gray Level Dependence (SGLD) Matrix.

- This matrix is square with dimension $N_g$, where $N_g$ is the number of gray levels in the image.

- Element $[i,j]$ of the matrix is generated by counting the number of times a pixel with value $i$ is adjacent to a pixel with value $j$ and then dividing the entire matrix by the total number of such comparisons made.

- Each entry is therefore considered to be the probability that a pixel with value $i$ will be found adjacent to a pixel of value $j$.

\[
G = \begin{bmatrix}
  p(1,1) & p(1,2) & \cdots & p(1,N_g) \\
  p(2,1) & p(2,2) & \cdots & p(2,N_g) \\
  \vdots & \vdots & \ddots & \vdots \\
  p(N_g,1) & p(N_g,2) & \cdots & p(N_g,N_g)
\end{bmatrix}
\]
Example of GLCM

- Adjacency can be defined to occur in four directions in a 2D, square pixel image (horizontal, vertical, left and right diagonals, four such matrices can be calculated.

- Rotation invariance is a primary criterion for any features used with these images, a kind of invariance was achieved for each of these statistics by averaging them over the four directional co-occurrence matrices.
## Haralick Texture Features (1-6)

<table>
<thead>
<tr>
<th>Feature</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angular Second Moment</td>
<td>$\sum_{i} \sum_{j} p(i, j)^2$</td>
</tr>
<tr>
<td>Contrast</td>
<td>$\sum_{n=0}^{N_{g}-1} n^2 \left{ \sum_{i=1}^{N_{g}} \sum_{j=1}^{N_{g}} p(i, j) \right}, \</td>
</tr>
<tr>
<td>Correlation</td>
<td>$\frac{\sum_{i} \sum_{j} (ij)p(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y}$</td>
</tr>
<tr>
<td>Sum of Squares: Variance</td>
<td>$\sum_{i} \sum_{j} (i - \mu)^2 p(i, j)$</td>
</tr>
<tr>
<td>Inverse Difference Moment</td>
<td>$\sum_{i} \sum_{j} \frac{1}{1+(i-j)^2} p(i, j)$</td>
</tr>
<tr>
<td>Sum Average</td>
<td>$\sum_{i=2}^{2N_{g}} i p_{x+y}(i)$</td>
</tr>
</tbody>
</table>

where $\mu_x$, $\mu_y$, $\sigma_x$, and $\sigma_y$ are the means and std. deviations of $p_x$ and $p_y$, the partial probability density functions.

where $x$ and $y$ are the coordinates (row and column) of an entry in the co-occurrence matrix, and $p_{x+y}(i)$ is the probability of co-occurrence matrix coordinates summing to $x + y$. 
# Haralick Texture Features (7-13)

<table>
<thead>
<tr>
<th>Feature</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sum Variance</td>
<td>[ \sum_{i=2}^{2N_g} (i - f_S)^2 p_{x+y}(i) ]</td>
</tr>
<tr>
<td>Sum Entropy</td>
<td>[- \sum_{i=2}^{2N_g} p_{x+y}(i) \log{p_{x+y}(i)} = f_S ]</td>
</tr>
<tr>
<td>Entropy</td>
<td>[- \sum_i \sum_j p(i, j) \log(p(i, j)) ]</td>
</tr>
<tr>
<td>Difference Variance</td>
<td>[ \sum_{i=0}^{N_g-1} i^2 p_{x-y}(i) ]</td>
</tr>
<tr>
<td>Difference Entropy</td>
<td>[- \sum_{i=0}^{N_g-1} p_{x-y}(i) \log{p_{x-y}(i)} ]</td>
</tr>
<tr>
<td>Info. Measure of Correlation 1</td>
<td>[ \frac{H_{XY} - H_{XY1}}{\max{H_X, H_Y}} ]</td>
</tr>
<tr>
<td>Info. Measure of Correlation 2</td>
<td>[ (1 - \exp[-2(H_{XY2} - H_{XY})])^{\frac{1}{2}} ]</td>
</tr>
<tr>
<td>Max. Correlation Coeff.</td>
<td>[ \text{Square root of the second largest eigenvalue of } Q ]</td>
</tr>
</tbody>
</table>

Where \( H_{XY} = - \sum_i \sum_j p(i, j) \log(p(i, j)) \), \( H_X \), \( H_Y \) are the entropies of \( p_x \) and \( p_y \), \( H_{XY1} = - \sum_i \sum_j p(i, j) \log\{p_x(i)p_y(j)\} \), \( H_{XY2} = - \sum_i \sum_j p_x(i)p_y(j) \log\{p_x(i)p_y(j)\} \).
Survey Over Classification Techniques

• **Linear Discriminant Analysis (LDA) [74, 75]**
  – Traditional method for classification
  – Construct decision boundaries by optimizing certain criteria to classify cases into one of mutually exclusive classes
  – High performance for linear separable problems, poor for non-linear separable data

• **Artificial Neural Network [76-78]**
  – Construct non-linear mapping function as a decision boundaries.
  – Two kinds: 3 layers back propagation of NN and Radial Basis Function (RBF) network
  – Robust, no rule or explicit expression is needed, widely applicable
  – No common rule to determine to size of the ANNs, long training time, over training.
**Image Database**

- **Mini-MIAS**- The Mammographic Image Analysis Society [Suckling et al., 1994]
  - An organization of UK research group.
  - Films taken from UK National Breast Screening Programme
  - Includes radiologist's "truth"-markings on the locations of any abnormalities that may be present
  - Available Online at the Pilot European Image Processing Archive (PEIPA) at the University of Essex
    - [http://peipa.essex.ac.uk/info/mias.html](http://peipa.essex.ac.uk/info/mias.html)

- Total Patients: 161
- Total Image: 322 (Left and Right Breast)
  - Images are digitized at a resolution of 1024x1024 and 8 bit gray level scale
  - Each image includes the location of abnormality, its radius, breast position, type of breast and tumor type
Proposed Method- ROI Extraction

• **X-ray label Removal- Global ROI**
  – Global thresholding (Otsu)
  – Connected component labeling
  – Calculate no. of pixels in each region
  – Pick the biggest region- Extracted breast region

• **Pectoral Muscle Removal**
  – Mass and pectoral region may have similar texture characteristics, causing a high number of FPs when detecting suspicious masses.
  – It is a higher density than the surrounding tissue
  – Automated region growing

• **Mass Extraction**
  – Mass is slightly brighter than its surrounding areas, produces a sharp peak of unusual gray level intensity pixels
  – Peak analysis of the histogram
  – Extract significant peak regions- **ROIs**
Proposed Method- Image Enhancement

• **Removal of Noise Effects**
  – Median filtering- very powerful in removing noise from 2D signals without blurring edges
  – Noise pixels generally have little correlation with mass pixels, median filter can smooth out these pixels so as to reduce their effects.

\[
I_2(x, y) = \text{median } \{I_1(x, y)\} = \text{median } \left\{ \sum_{i=-1}^{1} \sum_{j=-1}^{1} I_1(x + i, y + j) \right\}
\]

• **Contrast Enhancement**

\[
I_{en}(i, j) = \left( \frac{I(i, j)}{I_{max}} \right)^k \ast I_{max}
\]

\[
k = 2, 3, 4 \ldots
\]
Proposed Method- Image Segmentation

- **Automatic Seeded Region Growing using Haralick texture features**
  - Divide enhanced ROI into R x R non-overlapping blocks
  - If block is too small, the difference of the mass textures from normal textures can not be well characterized.
  - If it is too large, the result may be too coarse
  - Calculate the Haralick texture features from Spatial Gray Level Dependence Matrix (SGLD) of each block
  - Select the significant features that can easily discriminate mass and non mass region.
  - Select the blocks that contains mass based on the features
Proposed Method- Image Segmentation (Cont’d)

- Maximum gray level of that block is the seed point
- Region growing starts from that point and then grow iteratively and aggregate with the pixels that have similar properties
- Approximate the segmented mass to a circle
- Estimate the radius of the circle and compare it with the ground-truth data
- This comparison will provide the results how close the segmented mass to the ground-truth mass determined by the expert radiologists.
- Extract the mass region from the original image that is used as an input for classification
Proposed Method - Classification (Benign or Malignant)

- **Classifier**: Artificial Neural Network
- **Input**: 7 texture features
- **Output**: Benign (0) or Malignant (1)

<table>
<thead>
<tr>
<th>Feature</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>$\mu_{ij} = \frac{1}{(2n+1)^2} \sum_{x=i-n}^{i+n} \sum_{y=j-n}^{j+n} p(x, y)$</td>
</tr>
<tr>
<td>Skewness</td>
<td>$S_{ij} = \frac{1}{(2n+1)^2} \sum_{x=i-n}^{i+n} \sum_{y=j-n}^{j+n} \left[ \frac{p(i, j) - \mu_{ij}}{\sigma_{ij}} \right]^3$</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>$\sigma_{ij} = \frac{1}{(2n+1)} \sqrt{\sum_{x=i-n}^{i+n} \sum_{y=j-n}^{j+n} (p(x, y) - \mu_{ij})^2}$</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>$k_{ij} = \left{ \frac{1}{(2n+1)^2} \sum_{x=i-n}^{i+n} \sum_{y=j-n}^{j+n} \left[ \frac{p(i, j) - \mu_{ij}}{\sigma_{ij}} \right]^4 \right} - 3$</td>
</tr>
<tr>
<td>Smoothness</td>
<td>$R_{ij} = 1 - \frac{1}{1 + \sigma_{ij}^2}$</td>
</tr>
<tr>
<td>Uniformity</td>
<td>$U_{ij} = \sum_{k=0}^{L-1} \Pr_k^2$</td>
</tr>
<tr>
<td>Entropy</td>
<td>$h_{ij} = -\sum_{k=0}^{L-1} \Pr_k \left( \log_2 \Pr_k \right)$</td>
</tr>
</tbody>
</table>
Simulation Results- X-ray Label Removal

Original Mammogram mdb132.pgm

Histogram of the Original Mammogram

Outlined Breast Region

Extracted Breast Region
Simulation Results - Pectoral Muscle Removal

Breast Region

Pectoral Muscle

Outlined Pectoral Muscle

Pectoral Muscle suppressed Breast
Simulation Results- ROI Extraction

Original Mammogram mdb132.pgm

Histogram of the Breast Region

Extracted Significant Peak Region

Region of Interest
Sample ROIs
Simulation Results - Image Enhancement

Original Image

Histogram Equalization

Image Adjust

Adaptive HE
Simulation Results- Image Enhancement- Proposed

Original Image

Enhanced Image-2

Enhanced Image-3

Enhanced Image-4
Simulation Results - Image Segmentation

Non-overlapped Enhanced Image
Image Segmentation - Haralick Features (1-4)
Image Segmentation- Haralick Features (5-8)
Image Segmentation - Haralick Features (9-13)
Image Segmentation- Region Growing

Original Image

Enhanced Image

Segmented Image

Estimated=29, Original=33
Image Segmentation - Contour Extraction

Original Image

Enhanced Image

Segmented Image

Extracted Contour
Image Segmentation- Mass Extraction

Original Image

Segmented Image

Estimated Mass Region

Extracted Mass
Contour Extraction Using Proposed Method
Mass Extraction Using Proposed Method
Observations

- **Image Enhancement**
  - Since noise is also enhanced, noise is removed using median filtering.
  - For $k=4$, the image enhancement is satisfactory for this application.

  \[
  I_{en}(i, j) = \left( \frac{I(i, j)}{I_{\text{max}}} \right)^4 * I_{\text{max}}
  \]

- **Image Segmentation**
  - Block size 32x32 is found satisfactory.
  - Out of 13 Haralick texture features Sum Average is found very much significant to discriminate mass and non mass blocks.

  \[
  Sum \text{ } Average = \sum_{k=2}^{N_g} k * p_{x+y}(k)
  \]

  \[
  p_{x+y}(k) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j)
  \]

* $N_g$ is distinct gray levels $i + j = k$
Observations

- **Image Segmentation**
  - Seed point is automatically selected that corresponds the maximum Sum Average feature value
  - Segmented image is smoothed using some morphological operators like dilation, erosion, imopen and imclose.
  - Extracted mass is approximated to a circle and compare it’s radius to the original radius
  - Extracted mass is used as an input to the classification stage
Segmentation Validation- Maple Leaf

Original Image

Enhanced Image
Segmentation Validation- Image Division
Segmentation Validation

Original Image

Enhanced Image

Segmented Image

Outlined Contour
Performance Evaluation

- **Efficiency**: Computational time is less as it starts from the seed point and grows iteratively to its neighborhood.

- **Adaptibility**: Evaluated on 82 images containing malignant and benign masses with different size, shape and contrast. The algorithm works properly in all cases.

- **Robustness**: Evaluated with and without preliminary denoising steps. The two results are found to be comparable.
Performance Evaluation

- Estimated Region (ER): Segmented Region
- Reference Region (RR): Circular area estimated by radiologist
- Area Difference (AD) = |area(RR) - area(ER)|
- True Positive (TP) regions: Intersection of ER and RR
- False Positive (FP) regions: The area not identified in RR
- False Negative (FN) regions: The area in RR not identified in ER
- Completeness (CM)
- Correctness (CR)

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

\[
CR = \frac{TP}{TP + FP}
\]
Performance Evaluation and Result Analysis

Estimated Mass Region

Reference Mass Region

True Positive (TP) Region

False Positive (FP) Region

False Negative (FN) Region
Performance Evaluation and Result Analysis

Estimated Mass Region

Reference Mass Region

True Positive (TP) Region

False Positive (FP) Region

False Negative (FN) Region
## Performance Evaluation- Result Analysis

<table>
<thead>
<tr>
<th>Image</th>
<th>ER</th>
<th>RR</th>
<th>AD</th>
<th>TP</th>
<th>FP</th>
<th>FN</th>
<th>CM (%)</th>
<th>CR (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>010_B</td>
<td>2714</td>
<td>3545</td>
<td>831</td>
<td>2665</td>
<td>49</td>
<td>880</td>
<td>75</td>
<td>98</td>
</tr>
<tr>
<td>012_B</td>
<td>4080</td>
<td>5172</td>
<td>1092</td>
<td>3485</td>
<td>595</td>
<td>1687</td>
<td>68</td>
<td>86</td>
</tr>
<tr>
<td>023_M</td>
<td>2541</td>
<td>2753</td>
<td>212</td>
<td>2403</td>
<td>138</td>
<td>350</td>
<td>87</td>
<td>95</td>
</tr>
<tr>
<td>028_M</td>
<td>6705</td>
<td>10053</td>
<td>3348</td>
<td>6705</td>
<td>0</td>
<td>3348</td>
<td>67</td>
<td>100</td>
</tr>
<tr>
<td>092_M</td>
<td>6017</td>
<td>5962</td>
<td>55</td>
<td>5486</td>
<td>531</td>
<td>476</td>
<td>92</td>
<td>91</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>78</td>
<td>94</td>
</tr>
</tbody>
</table>

Sample Size: 82

<table>
<thead>
<tr>
<th>Correct Segmentation (%)</th>
<th>Radiologist Sensitivity (%)</th>
<th>Incorrect Segmentation (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>84.15% (69/82)</td>
<td>75%</td>
<td>15.85% (13/82)</td>
</tr>
</tbody>
</table>
Mass Classification- Artificial Neural Network (ANN)

- **Layers:** 3
- **Input Units:** Texture features (7)
- **Output Unit:** 1 (Benign=0, Malignant=1)
- **Hidden Units:** \((7+1)*2/3=5\)
- **Total Weights:** \((7*5)+5=40\)
- **Total Samples:** 69
- **Training Samples:** Benign (8), Malignant (8) - 25% (approx)
- **Testing Sample:** 53 - 75% (approx)
Training Data Preparation

Feature Extraction From the Image

- Load Image
- Benign
- Malignant
- Features Calc
- Save Features
- Exit

Extracted Features

- Mean: 0.757731854915619
- Std. Deviation: 0.32649302482605
- Smoothness: 0.0963292270898819
- Entropy: 0.999977350234985
- Skewness: -3.05705809593201
- Kurtosis: 0.246905758976936
- Uniformity: 0.140974879264832
- Output: 1
Mass Classification- ANN Testing

![Image of a software interface for mass classification with ANN testing, showing input and output controls, and data such as Height: 82, Width: 100, Execution Time: 0.031 seconds.](image-url)
Simulation Result- Mass Classification

- Sample Size: 53 (Testing mass only)
- Benign: 31
- Malignant: 22

<table>
<thead>
<tr>
<th>Correct Classification (%)</th>
<th>Misclassification (%)</th>
<th>Radiologist Misclassification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benign</td>
<td>Malignant</td>
<td>Benign</td>
</tr>
<tr>
<td>83.87% (26/31)</td>
<td>90.91% (20/22)</td>
<td>16.63% (5/31)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>65-90%</td>
</tr>
</tbody>
</table>
Conclusions

- X-ray label is removed using global Otsu thresholding technique followed by connected component labeling
- Pectoral muscle is removed using automatic region growing
- ROI is extracted using peak analysis from the histogram of the breast tissue
- Image enhancement of the ROI is done using nonlinear operator
- Automated seed region growing is proposed for image segmentation.
- Automated seed selection is done using Haralick texture features. Sum Average is found the most discriminative features among 13 features.
- Segmented image is smoothed using mathematical morphology
Conclusions and Future Works

• Proposed segmentation technique is validated using artificial maple leaf image. It was blurred artificially and then segmented correctly using the proposed techniques.

• Performance of the proposed method is evaluated using two quantitative features Completeness (CM) and Correctness (CR).

• Correct segmentation is achieved 84.15% that is very much very promising compare to the radiologist’s sensitivity 75%

• Artificial Neural Network is proposed for mass classification. Correct classification for benign is achieved 83.87% and for malignant 90.91%.

• Results are encouraging and have shown promise of our proposed system.

• Future Works: Optimal features selection for classification of masses
Thanks for your Patience. Questions?
References

References

References


References


References


References


References


References


References


References


75. P.A. Lachenbruch, Discriminant analysis, Hafiner, New York, 1975


How does ANN work?

- ANNs are adjusted or trained so that a particular input leads to a specific desired or target output.
Multi-Layer Perceptron (MLP)

- Most common NN model
- Uses supervised training methods to train the NN