Australian Journal of Basic and Applied Sciences, 9(36) December 2015, Pages: 449-453

ISSN:1991-8178

Australian Journal of Basic and Applied Sciences

Journal home page: www.ajbasweb.com

Automatically Detecting Various Stages Of Benign Tumor By Investigating Pathological Behavior Of Breast Tissues

[1]S.Thiruvalar Selvan and [2]R.Kayalvizhi

[1]*Research Scholar, Department of E&I, Annamalai University, Chidambaram, Tamilnadu, India.*
[2]*Professor, Department of E&I, Annamalai University, Chidambaram, Tamilnadu, India.*

ARTICLE INFO

Article history:
Received 10 November 2015
Accepted 30 December 2015
Available online 18 January 2016

Keywords:
Brain Tumor, Image Segmentation, MRI, Markov Random Field, Gray Level Co-Occurrence Matrix, Feature Extraction.

ABSTRACT

Detecting and analyzing benign tumor in breast tissues using electron microscopic takes more time and high magnification. To detect and decide the tumor the pathologists utilize the latest computer technology derived from digital pathology which can provide a result fast and effective. By investigating the digital tissues through computer system can provide more effective result with accurate classification based on the tissue – tumor affected states. There are various diagnosis system were proposed in the earlier researches in dissimilar areas with dissimilar physiognomies. The tumor can be classified in terms of tissue quality, the size of the tumor affected region and pathological behavior of the tissues in the region. To make detecting and classifying the tumor tissues, it is essential to extract the features of the region after segmentation. In this paper, Markov Random Field [MRF] method is utilized for segmenting the tumor regions, GLCM method is utilized for feature extraction and finally multi-class-SVM method is used for classifying the tumor. According to the classification a pathologist can advise the patient immediately for urgent treatment and also helps medical experts to take speedy actions and avoid death. The experiment is carried out in MATLAB software and the results show that the efficacy of the proposed approach is better than the existing approaches.

To Cite This Article: S.Thiruvalar Selvan and R.Kayalvizhi., Automatically Detecting Various Stages Of Benign Tumor By Investigating Pathological Behavior Of Breast Tissues, ***Aust. J. Basic & Appl. Sci.,*** *9(36): 449-453, 2015*

INTRODUCTION

An image is an array or a matrix, consists of square pixels, arranged in rows and columns. Intensity is assigned in an 8 bit grey scale image from picture element has an intensity values that ranges from 0 to 255. Gray scale image normally a white and black image. The various file formats are GIF, JPEG, TIFF, PS, PSD (photo shop format). Color images include RGB, CMYK uses subtractive color model (Iizulca, M., 1987). When images of astronomical objects are considered, it is usually taken by electronics detectors such as CCD (Charge coupled device). Telescope images are mostly gray scale. Color filter is used to take astronomical image. Filters used to generalize the images can be broadband or narrow band. Enhanced color image uses a chromatic order for an image. Digital remotely sensed image is composed of pixels located at the intersection of each row i and column j. Each pixel is a number called as Digital Number (DN) or Brightest Value (BV) (Chan, T. and L.A. Vese, 2001). Digital cameras are very expensive now a day, which includes a larger and larger number of pixels within one sensor. It is normally defined as size of image that can be captured by the sensor. One of the prime process is image enhancement to obtain the result is more accurate than the original image for a certain applications. This process makes the result more accurate by sharpening the image features like contrast, boundary regions and edges. The enhancement doesn't acquire properties of information content of data. It includes (i) Point operation where it involves contrast stretching, news clipping, histogram modeling and window slicing techniques, (II). Spatial Operation where it involves zooming, noise smoothing, median filtering and BP, HP and LP filtering, (III). Transform Operation where it involves in linear filtering, root filtering and homomorphic filtering and (iv). Pseudo coloring where it do false coloring states in pseudo coloring. Image enhancement consists of spatial and frequency based enhancement techniques. Spatial Domain of an image is represented as:

$$g(x,y) = f(x,y) * h(x,y) \quad (1)$$

And the frequency Domain representation is:

Corresponding Author: S.Thiruvalar Selvan, Research Scholar, Department of E&I, Annamalai University, Chidambaram, Tamilnadu, India.
E-mail: thiru20585@gmail.com@gmail.com

$$G(w1,w2) = F(w_1,w_2).H(w_1,w_2) \quad (2)$$

Processing methods are based on intensity of single pixels. It controls and coordinates the flow of pixel intensity from one to one pixel entity. Contrast stretching is generally increases the range of the gray levels dynamically.

$$Grey\_output(H) = grey\_scale(x,y) + H^{-1}(x,y) \quad (3)$$

The dynamic range of a processed image far exceeds the capability of a display device, in which case only the brightest part of the image is visible on the display screen. $S = C \log(1 + |a|)$, where log specifies the transform function and C is a constant. The digital image which holds gray scaling limits ranging from [L, L-1]. Number of images in the gray levels will be considered for the k^{th}gray level (You, Y. and M. Kaveh, 2000). The use of spatial mask for image processing is called spatial filtering. Blurring and noise reduction is removed by smoothing filters (Unser, M., 1986). Segmentation algorithm has been used for various applications like optical character recognition, automatic target acquisition of motion pictures detection and measurement of bone, tissue ethic in medical images (Arivazhagan, S. and L. Ganesan, 2003). It is required and mandatory for an image to pass pre-processing to correct image defects. Thresholding may be viewed as a function that involves test against a function F of the form

$$F = F[x,y, p(x,y),f(x,y)] \quad (4)$$

Where, f (x, y) is the gray level, and p (x, y) is a local property. Simple thresholding schemes compare each pixel gray level with a single global threshold. If T depends on both f (x, y) and p (x, y) then this is called as local thresholding (Bae, E., J. Yuan and X. Tai, 2011). It is used to filter the output or input to other operators. Multilevel thresholding classifies a point (x, y) as belonging to one object class if $T1 <(x, y) < T2$. Edge is the boundary between two regions. The point in which image brightness changes sharply are typically organized into a set of curved line segments called edges. Edge detection is a methodology to locate, identify and verify the discontinuity of the edges in an image. Image segmentation is an elementary step in most of the areas of computer vision which includes object recognition, target tracking and medical analysis etc. Image segmentation also provides some additional information about the image such as texture, color, edge and regions. The first step in the high level computer vision task, there are more challenges to ideal image segmentation. Segmentation is subdividing all the separate objects into constituent regions or objects. Means the segmentation should stop whenever the region of the objects has been detected. Segmenting objects from medical images has more practical difficulties due to the image quality, noise and resolution. In the past decade combination of various algorithms like graph cut, watershed segmentation has been successfully applied to get fairly accurate object segmentation and it still requires more user interaction. Imaging has become an essential component in many fields of bio-medical research and clinical practice. Breast cancer is one of the major problems among women's in the medical field. The new techniques of early diagnostic detection of breast cancer are critical for women's quality of life. Cancer forms in the tissues of breast.

Proposed Approach:

Tumor Segmentation:

In this paper the input image is breast mammogram collected from a dataset which is available publicly and referred in (Jegelka, S. and J. Bilmes, 2011). To improve the efficiency there are three image processing methods are combined together and applied. They are Markov Random Field based image segmentation, GLCM (Fritz Albregtsen, 2008) based feature extraction and Multi-Class SVM based tissue classification. According to the features the tissues are classified as benign or malignant. The entire functionality of the proposed approach used in this paper is shown in Figure.1.

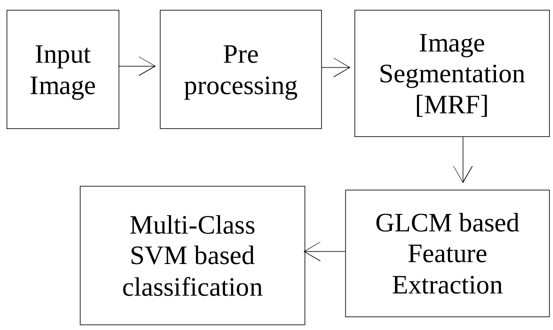


Fig. 1: Proposed Model for Breast

Australian Journal of Basic and Applied Sciences, 9(36) December 2015, Pages: 449-453

In medical imaging, most of the regions are often homogeneous where the neighboring pixels usually have similar properties such as intensity, color and texture. So, it is essential to utilize an automatic method having a strong theoretical background. One of the random probabilistic models can capture certain contextual constraints is Markov Random Field (Lempitsky, V.S. and Y. Boykov, 2007). MRF is a theoretical background, well studied and capable to detect homogeneous regions. Markov random field can be defined as labeled fields, neighboring pixels, energy points and etc. In this paper the standard pair wise Markov Random Field model is applied for interactive segmentation in breast mammogram images, which is proposed in (American Cancer Society, 2009a).

Markov Random Field Model:

Let Ibe the image, and the set of all pixels in the image isν. All the pair of neighboring pixels is represented as ε, which is the edges in the MRF. Each pixel i in the image **I**belongs to ν and it can be represented as $i \in \nu$ and it is obtained from a binary random variable x_i. The label forx_i is 0 or 1. The posterior distribution **P (|X|) = (1/Z) exp (-E (x))** of pair wise model factors into unary potentials $\emptyset_i(x_i)$ and pairwise potentials$\boldsymbol{\psi_{i,j}(x_i, x_j)}$**.** The Gibbs energy E is:

$$E(x) = \sum_{i \in \upsilon} \phi_i(x_i) + \sum_{i,j \in \varepsilon} \psi_{ij}(x_i, x_j) \quad (5)$$

The above function ϕ_iEncode the likelihood of pixel I belong to figure or ground, while ψ_{ij} is a contrast sensitive prior

$$\psi_{ij}(x_i, x_j) = \theta(I_i, I_j) | x_i - x_j | \\ = \theta(I_i, I_j)(x_i + x_j - 2x_i x_j) \quad (6)$$

The function $\theta(I_i, I_j)$ modulates the penalty for assigning different labels to pixels i and j based on

GLCM Feature Extraction:

A statistical process of examining structure that estimate the spatial relationship of pixels is the gray-level co-occurrence matrix (GLCM), also understood as the gray-headed-flat spatial concatenation table. The GLCM performance describes the structure of a semblance by scheming how often suit of pixel with specifying import and in a mention spatial relationship appear in a semblance, produce a GLCM, and then an extraction statistical degree from this grid. Gray Level Co-occurrence Matrix is used to extract the texture features from the segmented image by using second order statistical texture features. It is a matrix which has rows and columns equals to the number of gray levels. The element of matrix **P(i, j |** $\boldsymbol{\Delta x, \Delta y}$)is a relative frequency with two pixels separated by pixel distance called **(**$\boldsymbol{\Delta x, \Delta y}$**).** Let assume **P X Q** neighborhood of an input image consists of **G** gray levels from 0 to **G-1**. Let **f(p, q)** be the intensity at sample **p**, line **q** of the neighborhood. Then,

$$M(i, j \mid \Delta x, \Delta y) = KZ(i, j \mid \Delta x, \Delta y) \quad (7)$$

Where

$$K = \frac{1}{(P - \Delta x)(Q - \Delta y)} \quad (8)$$

$$Z(i, j \mid \Delta x, \Delta y) = \sum_{p=1}^{Q-\Delta y} \sum_{q=1}^{P-\Delta x} A \quad (9)$$

And

$$A = \begin{Bmatrix} 1 & if\ f(p,q) = i\ and\ f(p+\Delta x, q+\Delta y) = j \\ 0 & elsewhere \end{Bmatrix} \quad (10)$$

Image

0	1	1	2	3
1	1	0	2	2
1	1	1	3	3
0	0	2	2	3
0	0	0	1	2

P(i, j: 1, 0)

i, j	0	1	2	3
0	1/20	0	2/20	0
1	1/20	0	2/20	3/20
2	1/20	2/20	5/20	0
3	0	0	3/20	3/20

A G X G matrix must be accumulated for each window and each parameter set (d, θ). It is usually restricted to the values to be tested to a limited number of FOV-[Field Of View] values. The matrix contains a large occupancy level, it can be achieved by restricting the number of gray value quantization levels or by a relatively large window. The scalar texture measures $T(d, \theta)$ is extracted by the matrix

$$M_T(d) = \frac{1}{N_\theta} \sum_\theta T(d, \theta) \quad (11)$$

Where, $\boldsymbol{T(d, \theta)}$ gives overall addition of angular measurements and $\boldsymbol{N_\theta}$ gives number of measurements which gives angular mean$\boldsymbol{M_T(d)}$. And range is given as

$$R_T(d) = \max_\theta[T(d, \theta)] - \min_\theta[T(d, \theta)] \quad (12)$$

Where, the variance of angular second movement gives a difference of scalar texture measures $\boldsymbol{T(d, \theta)}$ mean value.

$$V_T^2(d) = \frac{1}{N_\theta} \sum_\theta [T(d, \theta) - M_T(d)]^2 \quad (13)$$

Australian Journal of Basic and Applied Sciences, 9(36) December 2015, Pages: 449-453

The area of the Breast tumor is calculated and computed by counting the pixels in an iterative manner by row wise or by column wise.

$$Area = \sum_{i=0,j=0}^{m,n} p(i,j) \quad (14)$$

$$mean = \frac{1}{m \times n}\sum_{i=1}^{m}\sum_{i=1}^{n} I(i,j) \quad (15)$$

$$variance = \frac{1}{m \times n}\sum_{i=1}^{m}\sum_{j=1}^{n}(I(i,j) - mean)^2 \quad (16)$$

$$Standard\ Deviation\ [\sigma] = \sqrt{\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}[P(i,j) - \mu]^2} \quad (17)$$

$$Third\ moment\ [\mu_i] = \sum_{i=0}^{L-1}(Z_i - m)^3 p(Z_i) \quad (18)$$

$$Entropy = -\sum \mathrm{p} \log_2 p \quad (19)$$

The closeness of the distribution of the image GLCM diagonal can be mathematically written as

$$Homogeneity = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\frac{P_{ij}}{(1+|i-j|)} \quad (20)$$

$$\text{Colours mean } (\mu) = \frac{1}{N}\left(\sum_{i=1}^{N} Pi\right) \quad (21)$$

Using a GLCM methodology the co-occurrence matrix of the tumor image is extracted. A total of 12 features (Shape, Area, Number of Pixels, color values, Mean, STD, Entropy, Homogeneity, Mean-Color and Variance.) is extracted from mammogram images and these features are representing the gray levels of the texture pattern of the tumor portion.

Multi-Class SVM Classifier:

Common Support Vector Machine-[SVM] classifier will classify the input data into two classes under True or False. Since the dataset may have a number of classes SVM cannot provide more number of classifications than true and false. Since, it is essential to classify numbers data under same dataset. One of the statistical classifier is Multi-SVM, which first nonlinearly map data with a high-dimensional space by kernels and then tries to find the hyper plane that separates data with a maximum margin in that new space. This feature searching and comparing process is repeated in the feature vector. The searching process is done in row wise, column wise and in orthogonal wise etc. Original Multi-SVM proposed for 2-class problems [True and False Class], SVM can be easily extended to multiclass problems by one-against-one or one-against-all strategies, where the latter is used in this work. Here the features extracted by GLCM methodology will be classified using multi-class SVM. The selected features will be the input to the Multi-class SVM classification and it decided the segmented tissues are normal tissues or abnormal tissues. According to the various features and the values the multi-class SVM classifier classifies the tissues are normal or abnormal.

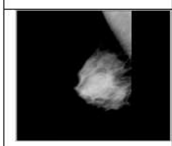
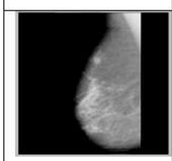
The overall steps followed in this paper are given below

$Tissue\_Analsis\_Steps$
1. $Read\ the\ image.$
2. $Segment\ the\ input\ image\ using\ MRF.$
3. $Feature\ Extraction\ Using\ GLCM.$
4. $Classification\ Using\ Multi - Class - SVM.$
5. $Suggest\ Treatment$

Experimental Results:

The proposed approach discussed in this paper is implemented in MATLAB software and results are verified. For the experiment, the sample input images are taken from MIAS database available in (Available at: http://marathon.csee.usf.edu/ Mammography/Database.html). The first input image is read from the user and the image will be segmented using MRF method. The input image, segmented image and the classified results are given in the following Table-1.

Table 1: Experimental Results of Proposed Approach

Input Image	Segmented Image	Classification
		Benign
		Malignant

The first column in Table-1 depicts the input image, the second column depicts the segmented image and the third column depicts the classified result obtained for the relevant segmented image given in the proposed approach experiment. This experiment is applied for more number of images and the performance of the proposed approach is verified in terms of classification. The classified result is compared with the ground truth image results and verified.

Conclusion:

The main objective of the paper is to detect and classify the breast tumor by analyzing the pathological behavior of the tissues. The combinatorial behavior of MRF, GLCM and Multi-Class SVM methods can detect, segment, extract the features and classify the tissues as normal or abnormal. The pathological behavior of the tissues is analyzed using the gray level co-occurrence matrix based feature variations. The experimental results are compared with the true label results and verified the performance of the proposed approach. From the verification, it is concluded that our proposed approach is efficient in terms of segmentation and classifying breast tumor.

Future Enhancement:

The proposed approach can be simplified in terms of number of iterations carried out and the performance can be investigated by comparing the true positive and false positive results with the other best approaches.

REFERENCES

Jegelka, S. and J. Bilmes, 2011. “Sub modularity beyond sub modular energies: Coupling edges in graph cuts”, In CVPR, pp: 1897-1904.

Lempitsky, V.S. and Y. Boykov, 2007. “Global optimization for shape fitting”, In CVPR.

American Cancer Society, 2009a. “Breast cancer facts and figures 2009–2010”, Atlanta: American Cancer Society, Inc.

Fritz Albregtsen, 2008. “Statistical Texture Measures Computed from Gray Level Co-occurrence Matrices”, Computer Vision, Graphics, and Image Processing; Graphical Models and Image Processing.

Available at: http://marathon.csee.usf.edu/ Mammography/Database.html)

Arivazhagan, S. and L. Ganesan, 2003. Texture segmentation using wavelet transform.Pattern Recognition Letters, 24: 3197-3203.

Bae, E., J. Yuan and X. Tai, 2011. Simultaneous convex optimization of regions and region parameters in image segmentation models.UCLA CAM Report, 11-83.

Chan, T. and L.A. Vese, 2001. Active contours without edges.IEEE Trans. Image Process.,10: 266-277.

You, Y. and M. Kaveh, 2000. Fourth-order partial differential equation for noise removal.IEEE Trans. on Image Processing, 9: 1723-1730.

Unser, M., 1986. ”Sum and Difference Histograms for Texture Classification”, IEEE Trans. on Pattern Analysis and Machine Intell., PAMI-8 pp: 118-125.

Iizulca, M., 1987. ”Quantitative evaluation of similar images with quasi-gray levels”, Computer Vision, Graphics, and Image Processing, 38: 342-360.