Microscopic Image Analysis for Detection of Breast Cancer

Bhagyashree Chavan, Dr.Anand Jatti, Rajasree P M

Bhagyashree Chavan, PG Scholar,Department of Electronics &Instrumentation Engg,R.V College of Engineering, Bengaluru, Karnataka-560059, Email: bhagya2416@gmail.com
Dr Anand Jatti, Associate professor, Department of Electronics & Instrumentation Engg,R.V College of Engineering, Bengaluru, Karnataka-560059. Email: anandjatti@rvce.edu.in
Rajasree P M, Assistant professor, Department of Electronics & Instrumentation Engg,R.V College of Engineering, Bengaluru, Karnataka-560059. Email: rajasreepm@rvce.edu.in

---***---

Abstract - *This study proposes the detection of cells in a cytological image of breast cancer using circular Hough transform instead of doing nuclei segmentation. Cytological images are choosen beacuse of freely distribution of cells in an image compare to histopathological images. The detected circles on each cells of a microscopic image are cropped and features are extracted for each cell using regionprops and GLCM algorithm in MATLAB software (Image processing tool). SVM classifier filters the incorrect cells from all detected cells using extracted features of each cell. Only correct cells are found on original microscopic image and counted.*

Key Words: Breast cancer, Circurlar Hough Transform (CHT), Gray level co-occurence matrix (GLCM), Support vector machine (SVM) classifier.

1. INTRODUCTION

The incidence of breast cancer and mortality due to breast cancer in India predicted in numbers for the year 2015 is 155863 and 75957 respectively based on the survey of year 2012. Hence the detection of breast cancer is very important task for perticular treatment to the patient in order to reduce the death rate. This paper focuses on analysis of cytology images. Counting of cells and feature extraction of each cell of cytological images for classification is very important for applications in practical midicine. Many papers have been shown the difficulties to do segmentation efficiently. Early segmentation techniques are used to extract foreground objects from noisy background. This preprocessing method includes ostu's thresholding method and adaptive thresholding [2] [7]. Graph based segemenation, watershed algorithm, Marker controlled watershed algorithm, 'snakes' which is a semi automatic segmentation method [2] [4] [8] [9] have been developed for segemenatation. The Complex wavelet based decomposition method, curvelet transform, sheartlet transform have been used to characterise the microscopic image [7] [6] [3]. The classification builded by NN classifier with multiplay perceprton and back propagation learning algorithm, GMM-SVM, Naive Bayes, Decision trees, K-NN, and Voting procedure to do classification, Probabilistic neural network (PNN)[1][4][5][9].

In this research the circular hough transform (CHT) is applied to detect each cell in a cytology image instead of using segmentation techniques. Ostu's thresholding method has been applied in order to seperate forground objects from background automatically and canny egde detection algorithm is applied for edge detection. Organization of paper is as follows: section2 includes methodology of the proposed work. Section3 explains about results obtained. Section4 conclusion

2. METHODOLOGY

The original image is convered into grayscale image. Ostu's thresholding technique is applied to seperate the forground required objects from background automatically. Thresholded image is operated by erosion and dilation to detect edges neatly using canny edge detection algorithm. Here the canny edge detector is used beacuse it removes noisy edges and detects only strong edge points. Circular hough transform detects circles on original image by using those edges as feature points. The detected circles are cropped and features extrected for each cell in order to filter incorrect cells from all detected circles using SVM classifier. Finally only correct cells are detected on image and counted. The praposed method is as given in the figure 1.

2.1 Ostu's Threshold

Segmentation by thresholding has to be done to extract cell nuclei from background pixels. The Ostu's thresholding method is choosen beacuse it is one of the automated method where two classes are made and it predicts threshold value itself. This algorithm gives optimum threshold value by assuming that an image containing two classes of pixels (foreground & background pixels). It seperates pixels into two classes, one class consists of combined pixels whose intra-class variance is minimal or equivalent and second class consists of pixels whose inter-class variance is maximal. Hence it determines an optimum threshold value is by separating those two classes.

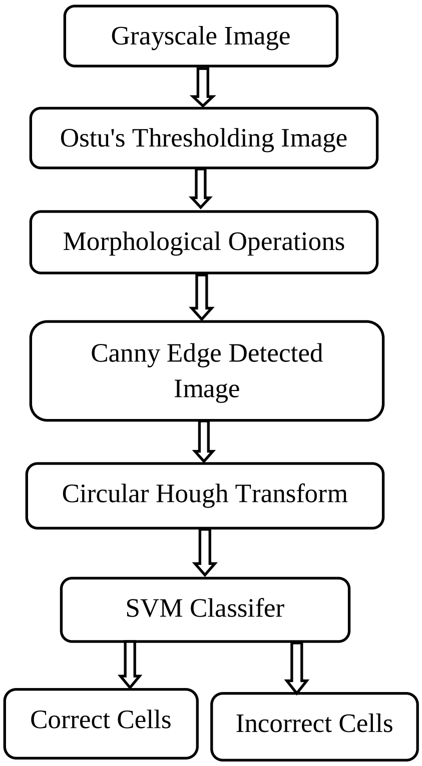


Fig 1: Scheme of the method

2.2 Morphological Image Processing

Morphological technique can be defined as making a structuring element to move over the binary image. Structuring element can be of any size and it can contain compliments of 1s and 0s. A logical operation is performed between underlying binary image and structuring element at each pixel position. It performs logical operation only when the structuring element is perfectly fit on the binary image else it does not perform any operations.

2.2.1 Erosion

Erosion is one of the two fundamental morphological operators. It is used to apply for original image to reduce the size of forground pixels. This operator takes two data sets as inputs. First one is input image which has to be eroded and second is a structuring element which is having set of coordinate points. The below equation $A \,\Theta\, B$ denotes erosion of set of forground pixels of A by set of coordinate points of structuring element B. Erosion of A by B gives a set of all points of b such that B translated by b, is contain A given by equation (1).

$$A \,\Theta\, B = X - b = \{z: (B + z) \subseteq X\} \qquad \text{..... (1)}$$

2.2.2 Dilation

Dilation is one of the two fundamental morphological operators. It is used to apply for original image in order to enlarge the size of areas of forground pixels. This operator takes two data sets as inputs. First one is input image which has to be dilated and second is a structuring element which is having set of coordinate points. In below equation $A \oplus B$ denotes dilation of set of forground pixels of A by set of coordinate points of structuring element B. Dilation of A by B gives a set of all points of b such that B translated by b, is contain A given by equation (2).

$$A \oplus B = X + b = \{x + b: x \in X \,\&b\, B\} \qquad \text{..... (2)}$$

2.3 Canny Edge Detection Algorithm

Steps for algorithm of canny edge detector

1. Smoothing: The input image is blurred to remove noise from an image.
2. Detection of gradients: Detects gradients of an image wherever there are large magnitudes of gradients and those edges should be marked.
3. Remove Non-maximum edges: Only local maxima points are taken as edges.
4. Double thresholding: Double thresholding determines the potential edges.
5. Tracking of edges: Finally edges are deteted by removing the dges which are not connected to a very certain (strong) edge.

2.4 Circular Hough Transform

The Edges detected in an image by canny edge detector are considered as feature points by CHT in order to detect circles on an image. Circular Hough Transform for detection of circles follows the same principle which is followed by Linear Hough transform for detection of line. The Circular Hough Transform will transform feature points in the image space into accumulated votes in Hough space. Votes for each feature point are accumulated in an accumulator array for all parameter combinations. The array elements indicate the presence of the pattern which consists of the highest number of votes. A circle pattern is described by given below equation.

$$(X - a)^2 + (y - b)^2 = r^2 \qquad \text{..... (3)}$$

Where r is the radius of the circle and (a, b) are the coordinates of the centre in the (x, y) direction respectively.

The circle is represented by parametric equations as given below.

$$x = a + r\cos\theta \quad \text{..... (4)}$$

$$y = b + r\sin\theta \quad \text{..... (5)}$$

The circle is represented by the three parameters: - radius (r), centre coordinates (a, b) and the Hough space (Three dimensional spaces with the Z-axis representing the radius). All the cells in a microscopic image are detected by using CHT algorithm and croping is done to extract features of each cell and classify as correct and incorrect cells. SVM classifier is builded to detect only correct cells on an image by filtering incorrect cells so that we can count the number of cancerous cells present in a microscopic image.

2.5 Feature Extraction

2.5.1 Regionprops

The 'regionprops' is a function used to extract the features of an image in MATLAB. The extracted features are given below:

- Area: Actual number of pixels in the region.
- Centroid: Center of mass of the region. It consists of horizontal and vertical coordinate (x-y coordinates respectively)
- Orientation: Angle between the long axis and the horizontal axis.
- Perimeter: Boundary of the labeled component.
- Mean and variance of R, G, B of each cell in a cytology image.

2.5.2 Gray-Level-Co-occurence Matrix (GLCM)

This method considers spatial relationship of pixels in GLCM for examination of texture. It calculates how often pairs of pixel with specific values and in a specified spatial relationship occur in an image to determine the texture of an image. It is done by creating a gray level co-occurence matrix and extracting statistical measures from this matrix. The extracted features using GLCM are as given below equations.

- Contrast: It is a measure of local variations of contrast or intensity in Gray-level-co-occurence matrix

$$\sum_{i,j=0}^{N-1} P_{i,j}(i-j)^2 \quad \text{..... (6)}$$

- Homogeneity: It is closeness of distribution of elements in GLCM to the diagonal of GLCM measurement.

$$\sum_{i,j=0}^{N-1} \frac{P_{i,j}}{1+(i-j)^2} \quad \text{.....(7)}$$

- Entropy: High entropy have inhomogeous view and low entropy have homogeneous view

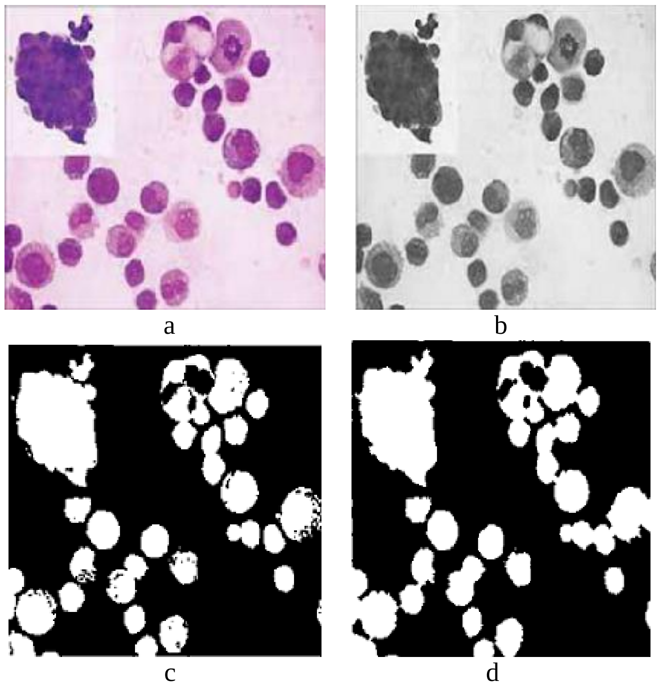
$$\sum_{i,j=0}^{N-1} P_{i,j}\left(-\ln P_{i,j}\right) \quad \text{..... (8)}$$

- Energy: Energy is defined as sum of squared elements in GLCM. It is a measure of homogeneity of an image.

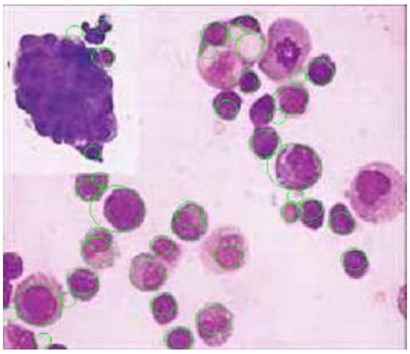
$$\sum_{i,j=0}^{N-1} \left\{\left(P_{i,j}\right)^2\right\} \quad \text{..... (9)}$$

3. Experimental Results

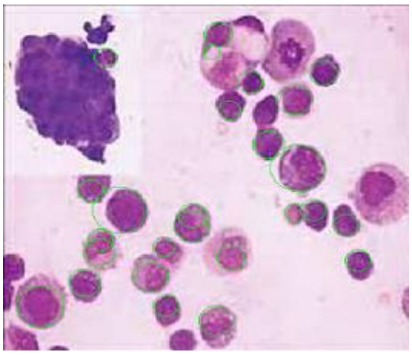
Figure2. Shows the steps result for detection of cancer cells in a cytological image of breast cancer. Figure c. shows the Ostu's thrsholded image of grayscale image and figure d. is obtained after applying morphological operations. Canny edge detection algorithm detects edges as shown in figure e. CHT performs detection of circles on original image by considering these edge points as feature points. Figure f shows detection of all circles. Cropping of all these cells has been done and features are extracted for each cell. Figure g shows only corrects cells detection after filtering of incorrect cells by SVM classifier using extracted features of each cell.



e



f



g

Fig 2: Cells detection (a) image (b) Converted grayscale image (c) Ostu's thresholded image (d) After morphological operations(erosion and dilation) (e) canny edge detected image (f) circles detected after CHT (g) detection of correct cells after filtering incorrect cells.

4. CONCLUSIONS

This paper proposes the work based circular hough transform instead of using other segmentation tachniques which are having some difficulties. Cells are detected by means of Ostu's thresholding method, Morphological operations, and circular hough transform (CHT). Features are extracted using regionprops and GLCM algorithm.The SVM classifier filters incorrect cells by analysing the extracted features of each cell. Finally correct cells are detected and counted. In this paper we have focused only on circles so future research can be focused on usage of ellipes which can provide more accurate model for nuclei.

REFERENCES

[1] S. Jitaree, A. Phinyomark, K. Thongnoo, P. Boonyapiphat, P. Phukpattaranont “Classifying Breast Cancer Regions in Microscopic Image Using Texture Analysis and Neural Network”, The 2013 Biomedical Engineering International Conference (BMEiCON-2013).

[2] Amresh Nikam, Arpita Gopal “Nuclei Segmentation from Breast Cancer Histology Images”; ICETACS 2013.

[3] Hadi Rezaeilouyeh, Mohammad H. Mahoor, S. Mohammad Mavadati, Jun Jason Zhang “A Microscopic Image Classification Method using Shearlet Transform”; International Conference on Healthcare Informatics 2013 IEEE.

[4] Sathish kumar.M, Dinesh.E “Involuntary Diagnosis of Intraductal Breast Images using Gaussian Mixture Model”;

[5] Paweł Filipczuk, Marek Kowal and Andrzej Obuchowicz “Breast Fibroadenoma Automatic Detection using K-Means Based Hybrid Segmentation Method”; ISBI 2012.

[6] Bailing Zhang “Breast Cancer Diagnosis from Biopsy Images by Serial Fusion of Random Subspace Ensembles”; 2011 4th International Conference on Biomedical Engineering and Informatics (BMEI)

[7] Issac Niwas S, Palanisamy P, Sujathan K “Wavelet based feature extraction method for Breast cancer cytology images”; 2010 IEEE Symposium on Industrial Electronics and Applications (ISIEA 2010), October 3-5, 2010, Penang, Malaysia

[8] Geeta P.K. Dinesh kumar A “Cell Segmentation and NC Ratio Analysis for Biopsy Images UsingMarker Controlled Watershed Algorithm”;

[9] Issac Niwas S, Palanisamy P, Sujathan K “Complex wavelet based Texture features of Cancer cytology Images”; 2010 5th International Conference on Industrial and Information Systems, ICIIS 2010, Jul 29 - Aug 01, 2010, India.