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 chosen because of the freely distribution of cells in an image compared 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, Circular Hough Transform (CHT), Gray level co-occurrence matrix (GLCM), Support vector machine (SVM) classifier.

2. METHODOLOGY

The original image is converted into grayscale image. Ostu's thresholding technique is applied to separate the foreground 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 because 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 extracted 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 proposed 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 chosen because it is one of the automated methods 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 separates 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.

\[ A \ominus B = \{x: (B + z) \subseteq X\} \quad \text{--- (1)} \]

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 foreground 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 foreground 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).

\[ A \oplus B = X + b = \{x + b: x \in X \& b B\} \quad \text{--- (2)} \]

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 detected by removing the edges 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.

\[(x - a)^2 + (y - b)^2 = r^2 \quad \text{--- (3)} \]
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.

\[
\begin{align*}
    x &= a + r \cos \theta \\
    y &= b + r \sin \theta
\end{align*}
\] .... (4) .... (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 cropping 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-occurrence 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-occurrence 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-occurrence matrix
  
  \[
  \sum_{i,j=0}^{N-1} P_{ij} (i-j)^2 
  \] .... (6)

- **Homogeneity**: It is closeness of distribution of elements in GLCM to the diagonal of GLCM measurement.
  
  \[
  \sum_{i,j=0}^{N-1} \frac{P_{ij}}{1+(i-j)^2} 
  \] .... (7)

- **Entropy**: High entropy have inhomogeneous view and low entropy have homogeneous view
  
  \[
  \sum_{i,j=0}^{N-1} P_{ij} (-\ln P_{ij}) 
  \] .... (8)

- **Energy**: Energy is defined as sum of squared elements in GLCM. It is a measure of homogeneity of an image.
  
  \[
  \sum_{i,j=0}^{N-1} \left( P_{ij} \right)^2 
  \] .... (9)

3. Experimental Results

Figure 2. Shows the steps result for detection of cancer cells in a cytological image of breast cancer. Figure c. shows the Ostu’s thresholded 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.
4. CONCLUSIONS

This paper proposes the work based circular hough transform instead of using other segmentation techniques 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 ellipses which can provide more accurate model for nuclei.

REFERENCES