ISSN:0975-9646

Comparative Study of Segmentation Methods applied on Cervical Cytology Images

G.Karthigai Lakshmi[1], K.Krishnaveni[2]

[1]Department of Computer Science, V.V.Vanniaperumal College for Women
[2]Department of Computer Science, Sri Ramasamy Naidu Memorial College

***Abstract* - Features of the cervical cytology cells determine the severity of cervical cancer. They are used to differentiate benign and malignant cancer cells. Cervical cytology images obtained from Papanicolaou test are segmented into background, nucleus and cytoplasm by various segmentation algorithms like Level Set, Active Contour, Watershed, K-Means Clustering, Morphological reconstruction and Gaussian Mixture model. A comparative study of these methods is carried out in this paper. The features of nucleus and cytoplasm extracted using the above mentioned methods are used to decide the stage of cancer.**

***Index Terms*- Segmentation methods, K-means Clustering, Gaussian Mixture Model, Watershed Segmentation, Active Contour**

I. INTRODUCTION

Cervical cancer is one of the most common cancers affecting women worldwide. It is caused by HPV or Human Papilloma Virus. Cervical cancer can be successfully treated when it is detected at an early stage through a Pap test. Cytology images obtained by Pap smear test are preprocessed and segmented into its components viz. background, cytoplasm and nucleus. Features like nucleus-cytoplasm ratio, position of nucleus with respect to center of cytoplasm, compactness, luminance, perimeter, diameter of nucleus and cytoplasm are used to find out the stage of cancer in a patient.

Automated Cervical smear screening is preferable to visual screening. Changes in the morphology of cervical cells can be accurately monitored with the help of system. Human diagnosis is tedious, error-prone, fatigue and expertise-dependent. Diagnosis is mind-numbing and unreliable due to the presence of obscuring elements like blood stains, non-uniform smear thickness and variable staining intensity. Automated diagnosis is error-free and accurate, due to pre-processing and noise removal prior to segmentation.

This paper deals with the techniques functional for the automated segmentation of cervical cytology images. Different categories of image segmentation are briefly mentioned in Section II. Techniques of image segmentation are described in Section III. A comparative analysis is done in Section IV. Concluding remarks are given in Section V.

II. DIFFERENT CATEGORIES OF IMAGE SEGMENTATION

Any image is made up of pixels. Pixels of same region are homogeneous in intensity. Segmentation is the process of splitting an image into its components based on homogeneity or heterogeneity in parts. Cervical cytology images have three parts: background, cytoplasm and nucleus and these parts may be segmented using pixel based, edge based or region based techniques. A region in an image can be defined by its border (edge) pixels or its interior pixels (pixels of region). The discontinuity in the image or variations in the intensity is prominent in the edges of regions and is used to differentiate them.

a) Pixel based Segmentation:

Pixels are the smallest addressable units of an image. Pixel based image segmentation techniques set a threshold value for partitioning the image. Most of these techniques use histogram of intensity values for choosing the intensity of each component and splitting the image based on this value. Clustering in color space is an indicator of similarity of image regions, and may be used for segmentation purposes. K means clustering labels the pixels of an image, as a member of a region whose mean values are determined at random or by choosing the intensity values of dominant colors in the image. So it can be classified as a hybrid technique using pixel and region based techniques.

b) Edge based Segmentation:

Edges in an image are marked by abrupt changes in the gradient values. Pixels that form the edge are linked and regions are partitioned using edges detected in this way. Active contours and Level set methods use edge based segmentation.

c) Region based Segmentation:

A region R of an image I is defined as a connected homogenous subset of the image based on some criterion such as intensity of color, gray level or texture. A region based segmentation of an image I, is a partition of I, into several homogeneous subsets or sub-regions. Region splitting and merging, Region growing and Watershed segmentation are examples of methods that employ region based segmentation technique.

III. TECHNIQUES OF IMAGE SEGMENTATION APPLIED ON CERVICAL IMAGES

Any image I can be represented as a function f(x, y) which specifies the intensity value of pixel at position (x, y). When this function is visually interpreted, one can see the image. Similarly, the cervical cytology image can be represented as a matrix of intensity values of order m x n (width x height of image) in the two dimensional space. The image is a RGB image, with each colour channel represented as a matrix. This matrix undergoes pre-processing for noise removal and contrast enhancement. The pre-processed image is given as input to segmentation procedures.

The segmentation procedures may work on RGB colour space or may convert from this colour model to L a* b*, YCbCr colour models or to a gray scale image. Some procedures use binary image.

A. Intensity Thresholding

The first and foremost method used in early stages of cervical cell segmentation is intensity thresholding. The cytoplasm and nucleus have intensities that differ significantly and consistently from the background. Automated threshold selection using the histogram is based on statistical analysis of the global or local image intensities. Since the size of nucleus is relatively high with respect to cytoplasm in single cells in the advanced stages of cancer, accurate choice of local intensity values is unfeasible in most cases. This method may be used as a primitive step in segmentation.

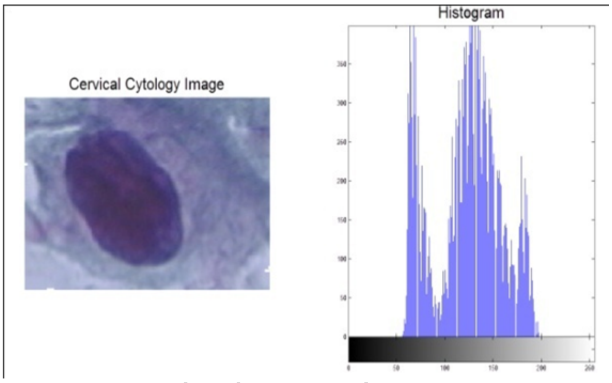


Fig 1. A Cervical Cytology Image and its Histogram

Fuzzy thresholding[1] can yield better results than normal thresholding.

B. K-means Clustering

K-Means Clustering is an unsupervised clustering algorithm. It classifies the pixels of input image into multiple classes based on their inherent distance from each other. The algorithm assumes that the intensity of pixels form a vector space and tries to group them into clusters. In cervical cytology image, the pixels are clustered around centroids μ_i (of the background, cytoplasm and nucleus) where i = 1,2,3 The centroids are obtained by minimizing the objective function F

$$F = \sum_{i=1}^{3} \sum_{x_j \in C_i} (x_j - \mu_i)^2 \qquad (1)$$

where there are 3 clusters C_i , i = 1, 2, 3 and μ_i is the centroid or mean point of all the points $x_j \in C_i$.

The steps in the algorithm are as follows:

1. Initialize the centroids of 3 clusters with 3 dominant colour intensities.
2. Cluster the points based on difference of their intensities from the centroid intensities.
$$c_{(i)} = arg\ min_j\ ||x_i - \mu_j||^2 \quad (2)$$
3. Compare the current and previous centroid values.
4. If the difference is very small, clustering is finished.
5. Otherwise, compute the new centroid for each of the clusters.
6. Repeat the steps 2 and 3 until the cluster labels of the image do not change significantly.

K-means clustering can be done both on RGB image and Gray level image.

If pre-processing is not appropriately done, then the noise may form additional clusters or it may merge with significant objects of the cervical image, resulting in poor segmentation.

C. Segmentation using Active Contours

An Active Contour is a deformable continuous closed spline curve[1]. It is also called as a snake. The deformation of active contour is determined by internal and external forces of gradient vectors.

$$E_{snake} = \int_0^1 (E_{int}(V(s)) + E_{image}(V(s)) + E_{con}(V(s))ds \qquad (3)$$

The internal force (E_{int}(V(s)) acts as constraint on smoothness of snake. $E_{image}(V(s))$ and $E_{con}(V(s))$ are the external forces. $E_{con}(V(s))$ is the constraint energy that drives the active contour towards the features of the image. $E_{image}(V(s))$ represents the energy function of image as a combination of energies of lines, edges and terminations in the image. Active contours perform global thresholding and edge detection by persistent assessment of continuity, curvature and strength of local edge. An active contour has significant setbacks like initial contour positioning and initial values of parameters. Figure 2 shows the output of Active Contour method on a Cervical Cytology image.

D. Watershed Segmentation

The watershed transform[3] finds "catchment basins" and "watershed ridge lines" in an image by treating it as a surface where light pixels lie in higher level and dark pixels are in low level. Segmentation using Watershed transform starts by computing a segmentation function out of an image whose dark regions are the objects to be segmented. Foreground markers which are connected blobs of pixels within each of the objects are formed. Pixels which do not belong to any object form the background markers. The segmentation function is then modified so that it only has minima at the foreground and background marker locations. The watershed transform is applied on the modified segmentation function. Watershed directly employed on the image may lead to over-segmentation. A combination of h-minima and watershed transform[4] may lead to proper segmentation of cervical cytology images. Fig 3 shows the application of Watershed transform on a cervical cytology image.

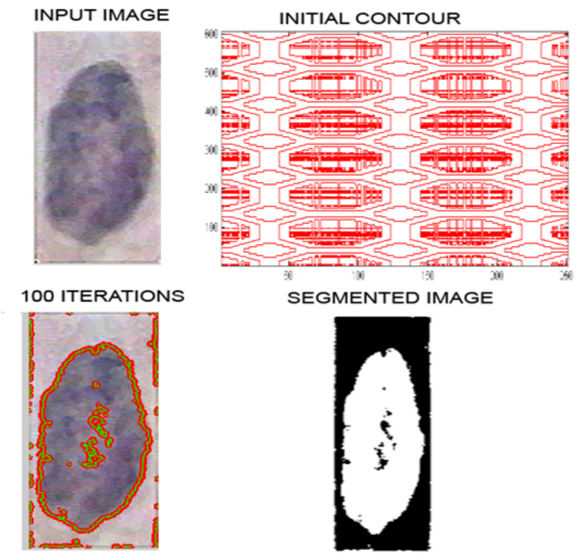
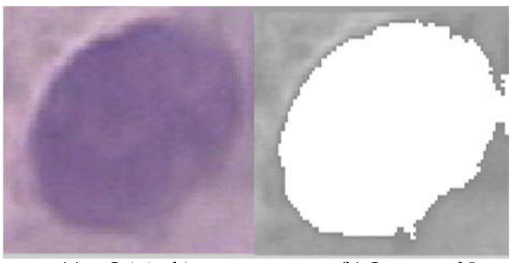


Fig 2. A Cervical Cytology Image subjected to Active Contour Segmentation



(a) Original image (b) Segmented Image

Fig 3. Segmentation using Watershed Transform

IV. COMPARATIVE ANALYSIS OF SEGMENTATION METHODS

Cervical cytology images can be partitioned into background, cytoplasm and nucleus using various segmentation techniques. Features of cytoplasm and nucleus are used to decide the severity of cancer. A comparison of the methods is given in table I.

TABLE I
COMPARISON OF SEGMENTATION METHODS

Method	Advantages	Disadvantages
Intensity Thresholding	Simple, fast, better for high quality image without noise.	If the image has patches of dye stains or blood stains, post-processing using morphological dilations and erosions is mandatory.
K-means Clustering	Efficient for single cell images.	If there is no homogeneity in the intensity of regions, erroneous segmentation may occur. The initial value of centroids should be carefully selected.
Active contours	Useful for background removal	Positioning of initial contour is crucial in deciding segmentation results
Watershed Segmentation	Helps in isolating cells from a group.	Over-segmentation may cause the components to be divided into further parts.

IV. CONCLUSION

In this paper, different segmentation techniques applicable on cervical cytology images were discussed. All the methods cannot be applied independently to get better results. Promising results can be obtained by hybrid techniques. Pixel based and edge based algorithms may be used as the preliminary processes. Region based systems can then be used to get better results. Thresholding techniques help in detecting nuclei boundaries in bi-level images. Morphological operations like dilation and erosion can also be used for nucleus discrimination. Canny edge detector followed by morphological processes help in separation of nucleus. Image subtraction can then be used to extract cytoplasm. Local and Global graph-cuts use boundary and regional information for partitioning. Combination of Fuzzy thresholding and Active Contours help in isolation of background. Gaussian Mixture Model [2] applied to image segmented using K-means Clustering yields enhanced outcomes. H-minima transformed image can be segmented using Watershed algorithm to isolate nucleus in a better way.

REFERENCES

[1] G.Karthigai Lakshmi, K. Krishnaveni, “Automated Extraction of Cytoplasm and Nuclei from Cervical Cytology Images by Fuzzy Thresholding and Active Contours”, *International Journal of Computer Applications (0975 – 8887) volume 73- No. 15, July 2013.*

[2] G.Karthigai Lakshmi, K.Krishnaveni. , “Multiple Feature Extraction from Cervical Cytology Images by Gaussian Mixture Model”, World Congress on *Computing and Communication Technologies (WCCCT), 2014 , pp.309-311, Feb. 27 2014-March 1 2014, published by IEEE.*

[3] http://in.mathworks.com/help/images/examples/marker-controlled-watershed-segmentation.html

[4] Aslı Kale, Selim Aksoy, “Segmentation of Cervical Cell Images”, *Proceedings of 20th IAPR International Conference on Pattern Recognition, pages 2399–2402, Istanbul, Turkey, August 23–26, 2010.*

[5] M. E. Plissiti and C. Nikou, “A review of automated techniques for cervical cell image analysis and classification”, *Biomedical Imaging and Computational Models in Biomechanics. Editors: D. Iacoviello and U. Andreaus. Lecture Notes in Computational Vision and Biomechanics, Vol.4, pp. 1-18, Springer 2013. ISBN 978-9400742697.*

[6] http://labs.fme.aegean.gr/decision