Segmentation of cervix using minimum spanning superpixel tree detector

Sheela Shiney TS 1*, Albert Jerome S1*, Jemila Rose2*
1Assistant Professor, Department of Computer Science and Engineering, St. Xavier’s Catholic College of Engineering, Chunakadai, Tamil Nadu, India.
1Associate Professor, Department of Biomedical Engineering, Noorul Islam Centre for Higher Education, Kumaracoil, Tamil Nadu, India.
2Associate Professor, Department of Computer Science and Engineering, St. Xavier’s Catholic College of Engineering, Chunakadai, Tamil Nadu, India.
*Corresponding Author sheela@sxcce.edu.in, albertjerome@niuniv.com, jemila.rose@gmail.com

Abstract: Cervical cancer occurs in cervix region of women which is life threatening and is a key research in the field of medical diagnosis. Cervical cancer can be prevented if precancerous changes is diagnosed earlier and cured properly. But finding cancer at the earlier stage is a tedious process. Cervical cancer screening is crucial since utmost of the screening procedures are invasive in nature. The objective is to automatically extract the area where in the cervical cancer starts to occur. In order to diagnose cervical cancer and distinguish the malignant and benign tissues, a cytology image obtained by cervicographic or colposcopy device is used. A robust segmentation algorithm is proposed that uses superpixel for the image segmentation method. This method is regularized by the combination of structural properties such as color, texture, and spatial data present in the superpixel graph.

1. Introduction

Cervical cancer is the fourth cancer worldwide which develops in woman’s cervix. Nearly all cervical cancer cases get infected with high-risk human papillomaviruses (HPV), an extremely common virus transmitted through sexual contact. The pre-cancerous changes may lead to cervical cancer and only few women gets affected by cancer. It may take nearly ten years for the pre-cancerous changes to change to cancer but in worst cases it may occur within a year. But the symptoms of cervical cancer can be seen only in advanced stage and leads to death. Getting a regular screening or pelvic examination can highly reduce the risk of pre-cancers turning into cancer. The screening procedure involves the great deals of skill which is associated in the screening procedure. It can be done only by a trained and knowledgeable cytotechnologist or a pathologist. Hence a novel method is required to overwhelm the death caused by cervical cancer in women by diagnosing at the initial stages.

2. Literature survey

Alan Conrad Bovik et al., suggested a representation to approximate the cell contour as a set of sparse contour points [1]. For the partial segmentation of overlapping cells from images of cervical smear contour points are extracted. To enhance the true edges a gradient decomposition- method is applied which is based on edge enhancement. Furthermore, to identify the feeble contour points a contour search method is implemented. Finally, to remove the exact cell boundary Gradient Vector Flow Snake model is performed.

The cervical cancer is diagnosed earlier by the automatic segmentation of nuclei from the microscopic images proposed by Brette L. Luck et al., [2]. The model is implemented by Gaussian Markov random fields which is combined with anisotropic median-diffusion filter. Finally classification is performed using Bayesian framework. The algorithm is implemented using cervical dataset images and it has been observed that it predicts 90% of manual-segmented nuclei having around six false positives per image.

An automated segmentation of major lesions is proposed by Changick Kim et al., so that the cervical cancer is diagnosed in the early stages [3]. The main aim of segmentation is to find the location for a biopsy automatically. The aceto-white region is segmented by using a non-convex optimization algorithm.
An automatic cell detection process is performed to overcome the overlapped nuclei is proposed by Chanho Jung et al., [4]. The distance transform is applied to overlapped nuclei. In this proposed algorithm the distance transform displays the area to be extracted. Next the parametric expectation-maximization is done in order to acquire knowledge about the area. To determine number of overlapping nuclei cluster validation is accomplished.

Christophoros Nikou et al., suggested a fully automated approach for detecting cell nuclei in Pap smear images [5]. Using morphological analysis the contour of the cell nuclei are detected and they are refined, which includes a priori knowledge about the circumference of each nucleus. Thirty eight cytological images was evaluated by using the proposed method recognized around 5617 cells are squamous epithelial cells.

P. Dollar et al., recommended that accurate segmentation of cervical cells in Pap smear images is significant to detect the precursor changes in cervix [6]. This study introduced a learning-based method to fragment distinct cell in Pap smear data in order to overcome the overlapping cell nuclei. To demonstrate the segmentation accuracy of our proposed work an evaluation metrics is performed using 2 dissimilar datasets.

Jacob Goldberger et al., discussed a segmentation scheme which segments automatically of main lesions is detected in primary phases of cervix region [7]. To identify the position of cervix for a biopsy segmentation is done automatically. By using a non-convex optimization approach the aceto-white region of cervix, is first segmented. Within the aceto-white region, by texture analysis further abnormal features are classified. The extracted features evidently discriminate the normal and abnormal regions using fuzzy c-means algorithm.

To segment nuclear and cytoplasmic contours, a semiautomatic PC-based cellular image analysis system was developed by Marina E. Plissiti et al., [8]. The first experiment result shows the cross-validation results of the average accuracies of 97.16% and 98.83%, respectively for four-cluster and two-cluster classifiers. The second experiment result shows the accuracy of 96.12% for four-cluster and 98.61% for two-cluster classifiers.

Micheal Hocket et al., suggested that visual information management image coding for effective communication [9]. To develop tools for analysing the image content is a perplexing mission for vision researchers and engineers In this paper using color and texture features, the properties of three tissue types are modelled.

Philip King et al., suggested an alternative screening tool named Cervicography that uses color cervix images which has been widely used [10]. An automated registration of Cervigram TM images are introduced in this technique. In this work a segmentation-based registration is defined and a combination of clustering- and active contour-based methodology is utilized.

Wei Wang et al., proposed a distance-guided learning technique implanted in a multiple classifier system (MCS) for segmentation of tissue from the optical images of the uterine cervix [11]. In order to compute the decision boundary shapes similarity between a pair of statistical classifiers distance based metric has been proposed.

Sadao Isotani et al., proposed a system for the determination of image density and morphometry for the analysis of malignant nuclei [12]. In this work Nuclei Demarcation is carried out for regions which are dark, the segmentation is performed by using the area threshold and finally the regions are validated to optimize the functionality of SPCIM.

A novel technique is proposed by V. Pallavi et al., [13] to grade the cancer severity. The Lugol’s iodine image is segmented to classify the abnormal tissues to diagnose the abnormal tissues accurately. This information is further used for grading the cancer.

Wei Wang et al., proposed that regular screening of cervix can detect pre cancer changes earlier [14]. In this work a lightweight yet elegant computer supported automatic method is proposed.
Zhiyun Xue et al., presented learning class-specific boundaries a method for extraction and segmentation of a class-specific object [15] automatically. Using a Markov random field (MRF) the watershed segmentation map of the input image is modelled which indicates whether the region is cancerous tissue or not.

It is important to develop tools to improve the reliability of cancer diagnosis for automated cancer diagnosis. The cancer diagnosis tool consists of four main computational steps: pre-processing, segmentation, feature extraction and classification. The cytology image obtained by cervicographic or colposcopy device is subjected to noise during staining process. The pre-processing step contains noise removal phase to remove the noise and to enrich the image quality to classify the objects and segmentation phase for extracting nucleus/cell information and abnormal regions. The feature extraction and diagnosis phase depends on the successful rate of segmentation results.

3. PROPOSED SYSTEM

The proposed architecture of the framework is exemplified in Figure. 1. The main objective of this work is to segment the tumour region from cervicography image.

![Figure 1 Block Diagram](image)

The noise information present in the input image is pre-processed by suitable pre-processing algorithm. Pre-processing is performed by using contrast enhancement method and cervical boundary growth. Then the image is being processed in feature extraction model. The feature extraction model consists of Illumination Layer, Reflectance layer, Texture layer. Finally the Minimum Spanning Superpixel Tree Detector is used to segment the cervical image.

An effective Simple linear iterative clustering (SLIC) algorithm is implemented to segment the cervical image into a group of superpixels.

**a. Feature Extraction**

The Feature Extraction comprises of illumination layer, reflectance layer and Texture Layer.

**Illumination Layer and Reflectance layer**

The input image is expressed mathematically as illustrated in equation 1. where $L_1$ denotes illumination layer and $L_2$ denotes reflectance layer.

$$I = L_1 + L_2$$  \hspace{1cm} (1)
Texture Layer

In order to analyse the texture features Gabor transform is mathematically expressed in the equation.

\[ T_\varphi(b, \theta, a) = c_\varphi^{-1/2}a^{-1} \int \varphi^*(a^{-1}r_{-\theta}(x - b))I(x) \, d^2x \]  

(2)

Where, \( I \) is the input image, \( A = \text{diag}[\eta^{-1/2}, 1] \) represents a 2x2 matrix in which non-diagonal values are zero, \( k_0 \) designates a vector frequency. \( \psi, \psi^*, a, \) and \( b \) are the two dimensional Gabor variables.

To attain clear cervical position in different angles, the maximum wavelet feedback for each pixel is calculated using the equation 3.

\[ M_\varphi(a, b) = \max_\theta |T_\varphi(b, \theta, a)| \]  

(3)

Where \( \theta \). Ranges from 0 to 170

c. Superpixel based Segmentation

Group of pixels which share similar characteristics is termed as superpixel. Superpixel carry more information than the pixels. The input image is partitioned into number of segments that constitute a set of pixels or superpixels. The number of superpixels in input image can be generated by using minimum spatial distance among various pixels. For computing super pixel the total number of pixels in the image is required. Then all the superpixels have been merged, till the number of superpixels is equivalent to the number of groupings contained in image.

SLIC based Superpixel Segmentation

Based on the color and proximity of image the Simple Linear Iterative Clustering (SLIC) algorithm generates superpixels by grouping similar pixels in an image. Superpixel can be generated by using five dimensional \( [r \, g \, b \, x \, y] \) space where \( r \, g \, b \) specifies pixel color values and \( x \, y \) denotes the coordinate position of pixel. To cluster pixel in this five dimensional space a distance measure is introduced.

Minimum Spanning Superpixel Tree Detector

In this paper, in order to employ the SLIC algorithm take illumination and texture layer information, to divide the input image into a cluster of superpixels. The cluster centres are located at each pixels. In order to prevent positioning cluster centres at the edges move the cluster centres towards the deepest gradient position. Subsequently the clustering process is applied on each pixels to the nearby cluster center. The overall distance measure \( D \) is expressed in equations (4) to (7).

\[ d_{rgb} = \sqrt{(R_{k1} - R_{k2})^2 + (G_{k1} - G_{k2})^2 + (B_{k1} - B_{k2})^2} \]  

(4)

\[ d_t = \sqrt{(g_{k1} - g_{k2})^2} \]  

(5)

\[ d_s = \sqrt{(x_{k1} - x_{k2})^2 + (y_{k1} - y_{k2})^2} \]  

(6)

\[ D = \sqrt{m_{rgb}d_{rgb}^2 + m_td_t^2 + m_sd_s^2} \]  

(7)

where \( R, G, B \) denote the red, green, and blue frequencies of the illumination layer, the pixel position, and the cluster center coordinate points respectively. Furthermore, \( d_{rgb}, d_t, \) and \( d_s \) represent
the illumination, texture, and spatial proximities, respectively, and $m_{rgb}$, $m_t$, and $m_s$ represents the weights of the illumination distance, texture distance, and spatial distance, respectively.

4. Results and Discussion

a. Pre-processing

The input image is given in Figure. 2 and it is resized as shown in Figure. 3 and this image is converted into grayscale image as illustrated in Figure. 4. In order to enhance the contrast and to detect the boundary of cervix pre-processing step is applied. The local contrast enhancement is achieved by using Contrast Limited Adaptive Histogram Equalization (CLAHE) algorithm and is shown in Figure. 5. In addition, to drawbacks generated around camera aperture border. The wavelet transformation is applied to eliminate the artefacts and is shown in Figure. 6.

b. Feature Extraction

In order to extract the features, a probability-based method is applied so as to obtain illustration of the input image. From the input image, Illumination Layer is obtained and is given in Figure. 7. Reflectance Layer is also obtained and is shown in Figure. 8. The reflectance layer seems to be, illuminated wherein some portions are distributed unevenly. The Texture Layer given in Figure. 9 is computed from Cervical Boundary image.

c. Segmentation

The Super Pixel Image is shown in Figure. 10 and the final segmented output is shown in Figure. 11.
5. Conclusion

Segmentation of cervix is a crucial task in the field of medical imaging. Although the image has artefacts and complex pathologies, existing methods have failed to appropriately segment poor contrast cervical image. Henceforth a superpixel based segmentation algorithm is introduced which operates on group of pixels rather than a single pixel. Furthermore a Minimum Spanning Tree (MSST) is implemented to evaluate each superpixels.

References

[1] Alan Conrat Bait, Barinova O, Kuzmishkina A, Vezhnvets A and Vezhnvets V 2007 Accurate Segmentation of Partially Overlapping Cervical Cells Based on Dynamic Sparse Contour Searching and GVF Snake Model in Proc. Graphicon 162–165.

[2] Brette S, Luck L, Belongie S, Malik J and Puzicha J 2002 An Image Model and Segmentation Algorithm for Reflectance Confocal Images of In Vivo Cervical Tissue IEEE Trans. Pattern Anal. Mach. Intell. 24 4 509–522.

[3] Changick Kim, Boykov Y and Funka-Lea G 2006 Graph cuts and efficient N-D image segmentation Int. J. Comput. Vis 70 2 109–131.

[4] Chancho Jung G, Canny J 1986 Unsupervised Segmentation of Overlapped Nuclei Using Bayesian Classification IEEE Trans. Pattern Anal. Mach. Intell. 8 6 679–698.

[5] Cristopher Nikon PM, Gerbaldo D, Perino AR, Piccoli FJ, Montz and Captianio GL 1995 Automation Detection of Cell Nuclei Using Pap Smear Obstet. Gynecol 85 1011–1016.

[6] Dollar P, Tu Z and Belongie S 2006 Accurate Cervical Cell Segmentation from Overlapping Clumps in Pap Smear Images in Proc. Comput. Vis. Pattern Recognit. (CVPR) 1964–1971.

[7] Jacob Gold Berger B, Eluf-Neto J and Nascimento CM 2001 Segmentation and Classification of Cervix by Pattern and Texture Semin. Oncol 28 2 188–97.

[8] Marina.E.Plissiti C, Fei-Fei L and Perona P 2005 Semi-Automatic Segmentation and Classification of Pap Smear Cells in Proc. Comput. Vis. Pattern Recognit. (CVPR) 524–531.

[9] Micheal Hocket H, Felzenszwalb P 2001 Image Segmentation of Uterine Cervix Images for Indexing in PACS in Proc. Comput. Vis. Pattern Recognit. (CVPR) 56–62.

[10] Philip King J, Geraud T, Margin J-F, Bloch I and Maitre H 1995 An Automated, Segmentation-Based, Rigid Registration System for Cervigram TM Images Utilizing Simple Clustering and Active Contour Techniques in Proc. Int. Conf. Image Process 548–551.

[11] Wei Wang M, Gordon S, Zimmerman G, Long R, Antani S, Jeronimo J and Greenspan H 2006 Automated and Interactive Lesion Detection and Segmentation in Uterine Cervix Images in Proc. SPIE Med. Imag 6144 1549–1556.

[12] Sadao Isotani R, Greenspan H, Gordon S, Zimmerman G, Lotenberg S, Jeronimo J, Antani S and Long R 2009 A System For the Density of Image and Morphometric Analysis in Proc. IEEE Trans. Med. Imag 454–468.

[13] Pallavi V, Haris K, Efstratiadis S N, Maglaveras N and Katsaggelos A K 1998 Segmentation and Classification of the Cervix by Pattern and Texture Analysis IEEE Trans. Image Process 7 12 1684–1699.

[14] Wei Wang, Herrero R, Hildesheim A, Bratti C, Sherman M E, Hutchinson M, Morales J, Balmaceda I, Greenberg M D, Alfaro M, Burk R D, Wacholder S, Plummer M and Schiffman 2000 Automated and Interactive Lesion Detection and Segmentation in Uterine Cervix Images J. Nat. Cancer Inst 92 6 464–474.

[15] Zhiyan Xine, Herrero R, Schiffman MH, Bratti C, Hildesheim A, Balmaceda I, Sherman ME, Greenberg M, Cárdenas F, Gomez V, Helgesen K, Morales J, Hutchinson M, Mango L, Alfaro M, Potischman NW, Wacholder S, Swanson C and Britton LA 1997 Computer assisted Cervical cytological Nucleus Pan Amer. J. Public Health 1 15 362–375.