The 2010 IEEE/ICME International Conference on Complex Medical Engineering
July 13-15, 2010, Gold Coast, Australia

Computer Aided Abnormality Detection for Microscopy Images of Cervical Tissue

Yue Cui, Jesse S. Jin, Mira Park, Suhuai Luo, Min Xu, Yu Peng, W. S. Felix Wong and Leonardo D. Santos

***Abstract*— Cervical cancer is the second most common malignancy among women worldwide, if it is detected in early stage, cure rate is relatively high. Computer aided abnormality detection for cervical smear is developed to assist medical experts to handle the microscopy images, examine cell abnormalities and diagnose dyskaryosis. The microscopy images of cells in cervix uteri are stained by the tumor marker Ki-67, so that the abnormal nuclei present brown while normal ones are bluish. Segmentation is the most important and difficult task to calculate the ratio of abnormal nuclei to all nuclei. In order to achieve accurate segmentation of nuclei, we propose a multi-level segmentation approach for abnormality identification in microscopy images. First level segmentation aims to partition abnormal (stained) nuclei regions and all nuclei regions. Because of under-segmentation after first level segmentation, second level segmentation is applied to further partition the clustered nuclei. In order to classify touching regions of clustered nuclei and separate regions of single nucleus, relevant meaningful features are extracted from regions of interest. Consequently all the nuclei regions are separated and in conjunction with the abnormal nuclei regions in the first level segmentation, the abnormality *i.e.* ratio of abnormal nuclei to all nuclei is obtained. Experimental results indicate that our method achieved an accuracy of 93.55% and 95.8% in term of abnormal nuclei and all nuclei respectively for identification of abnormalities. Our proposed method produces a satisfactory segmentation.**

I. Introduction

CERVICAL cancer is the second most common malignancy among women worldwide, if it is detected in early stage, cure rate is relatively high. Pap smear, invented by Georgios Papanicoloau, is a screening procedure to diagnose pre-invasive and early invasive cancer [1]. The procedure involves taking a sample of cervical cells from a specific area on the cervix. The slide is viewed for the presence of abnormalities. Manual detection requires a high level of skill, and is labor intensive, exceptionally boring and time consuming. Therefore, there is an increasing demand for computer aided abnormality detection which can analyze abnormalities in microscopy images accurately.

Some researches on computer aided cervical cancer detection have been investigated. Bamford [2] used active contours for cervical cell nucleus segmentation. B.L. Luck *et al.* [3] developed a model in conjunction with anisotropic median-diffusion, Gaussian Markov random fields and a Bayesian framework for automatically segmenting nuclei in confocal reflectance images of cervical tissue. Instead of dealing with reflectance confocal images, Ref. [4] used colposcopy images of cervical cancer and proposed a methodology to evaluate temporal changes of tissue color. Ref. [5] applied Support Vector Machines in Fourier-Transform Infrared data to enhance and improve upon the standard Pap test. However, our work is different from theirs in which we use microscopy images for abnormalities. Ki-67 is a cellular proliferation marker used in numerous clinical or lab researchers. Based on research by Jeffers *et al.* [6], Ki-67 index is significantly higher in leiomyosarcomas than in benign leiomyomas. The same as our previous work in [7], the Pap smear slide from cervix was dyed by Ki-67 in order to assist identification of abnormality for uterine tumor detection. Consequently, abnormal nuclei were stained brown while normal nuclei are still bluish as shown in Fig. 1.

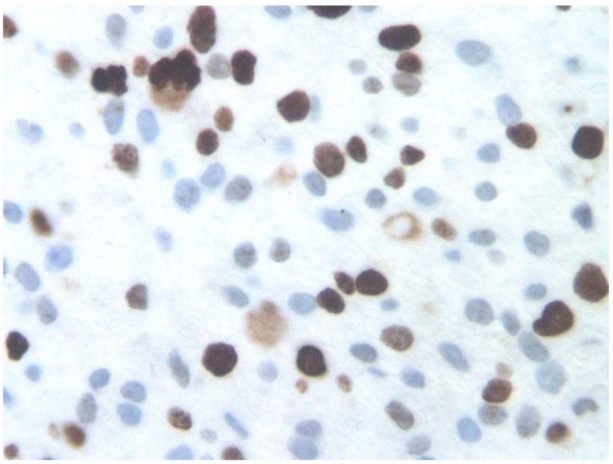


Fig. 1. Microscopy image of cervical tissue slide stained by Ki-67 with a resolution of 1200 * 1600 pixels. We observe abnormal nuclei are stained brown while normal nuclei are bluish.

Manuscript received February 5, 2010. This work was supported in part by the CSC-Newcastle Scholarship, ARC LP0669645 and IntelliRAD.

Yue Cui, Jesse S. Jin, Mira Park, Suhuai Luo, Min Xu and Yu Peng are with the School of Design, Communication and IT, University of Newcastle, Callaghan 2308 Australia (phone: 61-2-49854509; e-mail: Yue.Cui@uon.edu.au).

W. S. Felix Wong is with the Department of Obstetrics and Gynaecology, School of Medicine, University of New South Wales, Sydney, NSW 2052, Australia.

Leonardo D. Santos is with Department of Anatomical Pathology, Sydney South West Pathology Service, Liverpool Hospital, Liverpool NSW 2170, Australia (e-mail: Leonardo.Santos@sswahs.nsw.gov.au).

Segmentation plays an important and tough role in image processing due to the complexity and diversity of images. A number of automated methods on nuclei segmentation have been proposed, such as thresholding [8], watershed algorithm [9, 10] and active contour [2, 11]. Thresholding is widely used for segmentation tasks [12]. An algorithm based on thresholding in nuclei segmentation is proposed in [8]. However thresholding is easily affected by noise and uneven illumination. Watershed is another powerful tool in cell and nuclei segmentation. Yang *et al.* [10] proposed a novel marker-controlled watershed which can effectively segment clustered nuclei with less over segmentation. In Ref. [9], a mathematical morphology-based watershed algorithm is applied to segment clustered nuclei. In addition, active contour method is used in [2, 11]. Fok *et al.* [11] extracted boundaries of each axons based on active contour model in conjunction with a rough identification of all the axon centers by using elliptical Hough transform.

Although the existing methods are reasonably successful for the tasks, relatively few studies focused on specific cervical tissue images. In this paper, we propose a multi-level segmentation approach for abnormality identification in microscopy images. The first level segmentation aims to partition abnormal nuclei regions and all nuclei regions. Whereas there are touching regions of clustered nuclei, second level segmentation is applied to further partition touching regions. Finally all the nuclei regions are separated. The ratio of abnormal nuclei to all nuclei is obtained by combining abnormal nuclei regions with all separated nuclei regions.

The rest of this paper is organized as follows. Section II introduces the overview of the system. In Section III, our multi-level segmentation method is described. Experimental results on dataset are presented in Section IV. Finally Section V concludes the paper.

II. System Overview

Fig. 2 illustrates an overview of our proposed multi-level segmentation approach. At first a powerful smoothing method anisotropic diffusion is applied in the original image for noise reduction. Then the first level segmentation is employed in order to partition abnormal nuclei and all the nuclei regions based on k-means clustering in L*a*b* color space. We find that under segmentation is predominant and over segmentation is rare because clustered nuclei are segmented as a touching region. In order to classify clustered nuclei and single nucleus, relevant meaningful features including solidity, intensity standard deviation and area are extracted from regions of interest. After classification of touching regions and separated regions, second level segmentation is applied in touching regions in order that all the clustered nuclei are separated. After separating all nuclei, combined with the abnormal nuclei information obtained in the first level segmentation, abnormal nuclei are marked in the final segmented image. Finally, the ratio of abnormality is obtained.

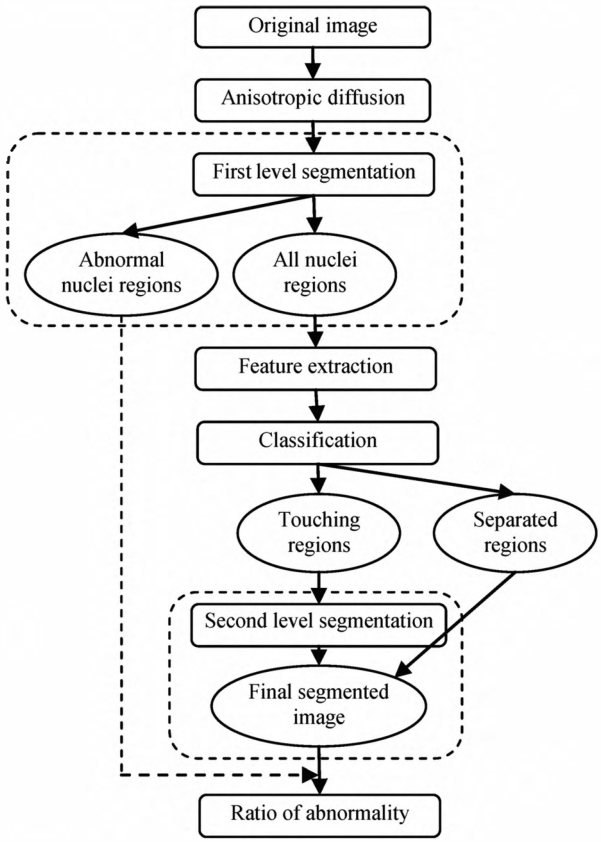


Fig. 2. Overview of computer aided abnormality detection process

III. Method

Microscopy images are acquired from a specific area in the cervix uteri and stained with proliferation marker Ki-67. As illustrated in Fig. 1, what we observe are nuclei in cells and abnormal ones are marked brown, while normal ones are bluish. Note that each cell has only one nucleus, counting the cell number is corresponding to counting the nuclei number.

A. Preprocessing

At first a powerful enhancement method anisotropic diffusion is applied in original image. Anisotropic diffusion is a nonlinear, iterative process that smoothes images while maintaining edges [3, 13]. As the gradient has the highest value perpendicular to the edge and is dilated along the edge, we increase the smoothing function parallel to the edge and stop the smoothing perpendicular to the edge. Therefore, anisotropic diffusion effectively reduces noises while maintaining nuclei edges.

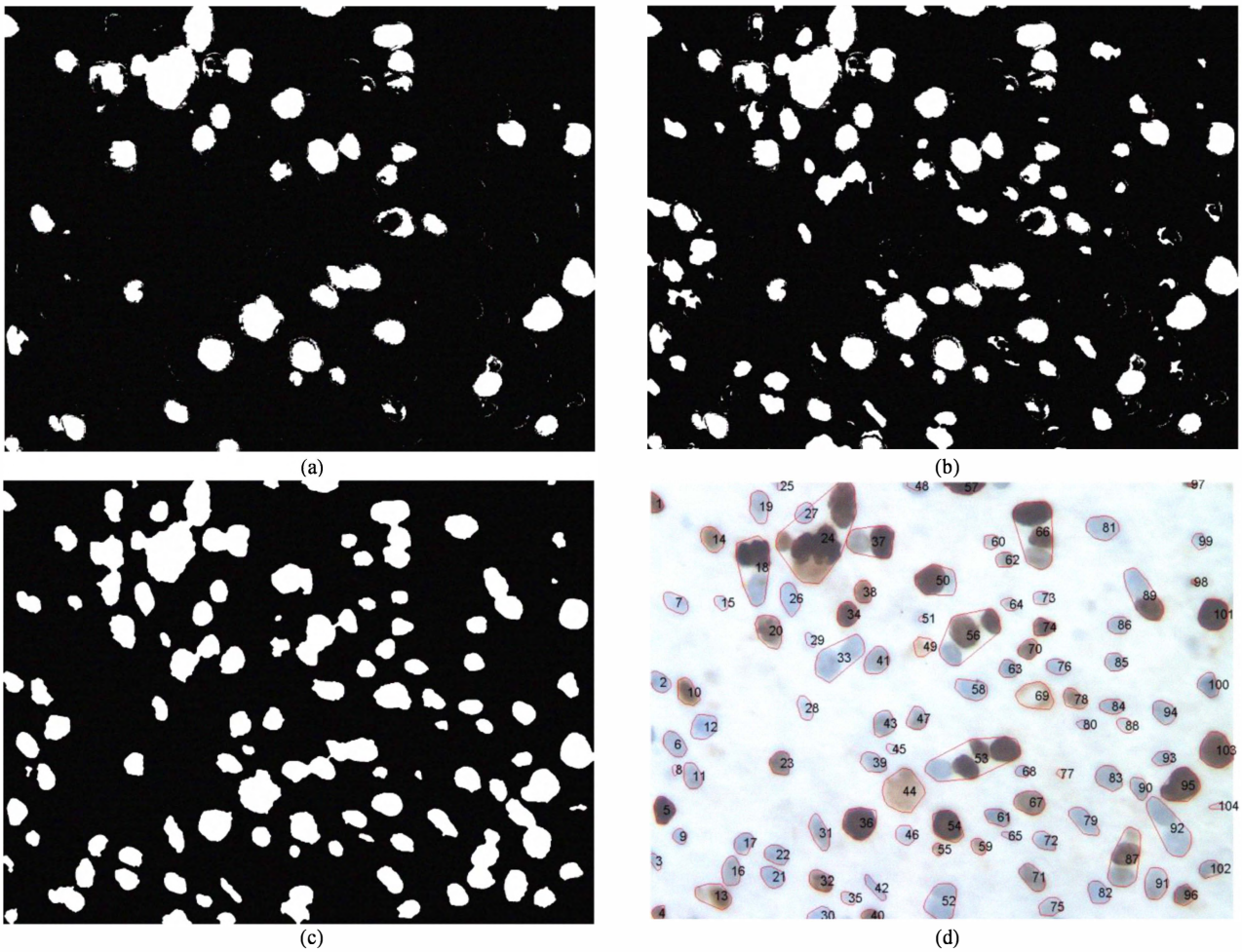


Fig. 3. First level segmentation. (a) Abnormal nuclei regions and (b) all nuclei regions using k-means clustering. (c) Edge information added to (b), (d) Convex hull of each region in (c). We observe in (d), most of the regions are separated regions, however, there are some clustered nuclei regions because of under-segmentation.

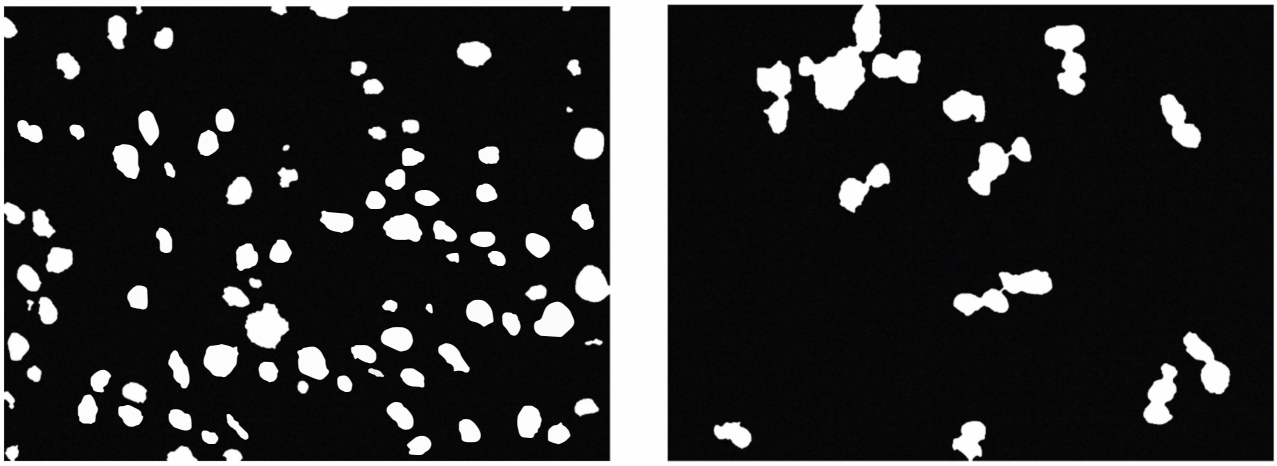
B. First Level Segmentation

Microscopy data from a color camera are digitized RGB images. It is more common to describe a color by RGB components on computers, but as human beings, we cannot distinguish each color from RGB representation. Unlike RGB, L*a*b* is designed to approximate human vision. L*a*b* color space which is also known as CIELAB or CIE L*a*b* can quantify these visual differences. In the three coordinates, L* represents the lightness of the color, a* indicates the position between red/magenta and green and b* indicates the position between yellow and blue. Thus, all of the color information is in the 'a*' and 'b*' layers. We perform k-means clustering in 'a*b*' space to classify abnormal nuclei regions and all nuclei regions based on the color information. We set k equals to 6 experimentally and obtain abnormal nuclei regions, all nuclei regions and background respectively. Fig. 3(a) and (b) shows abnormal nuclei regions and all nuclei regions. We average centers of each training image as six clustering centers for classification for testing images. However some of the nuclei are not segmented completely as shown in Fig. 3(b), edge information and morphological operation are added in order to segment complete nuclei. Result is shown in Fig. 3(c). Convex hull of each region in (c) is shown in (d).

C. Touching Regions Classification

We observe most of the regions are separated regions in Fig. 3(d), which means each region has only one nucleus. However, there are still some touching regions cannot be isolated in the first level segmentation. They need to be classified and further segmented. Note that the shape of nucleus is nearly convex. If the region shape is not convex, it is probably a touching region because of under-segmentation. We define solidity of each region as a meaningful feature, which is given by

$$s(n) = \frac{a(n)}{ca(n)} \tag{1}$$



(1) (2)

Fig. 4. Classification of touching nuclei regions. (1) Separated nuclei regions; (2) Touching nuclei regions.

where ***a(n)*** and ***ca(n)*** are region area and convex hull area of the n^{th} region. Intuitively from (1), clustered nuclei regions have small solidity.

Intensity standard deviation is another feature used to measure the average difference of the intensity from the mean. The n^{th} region intensity standard deviation takes the equation

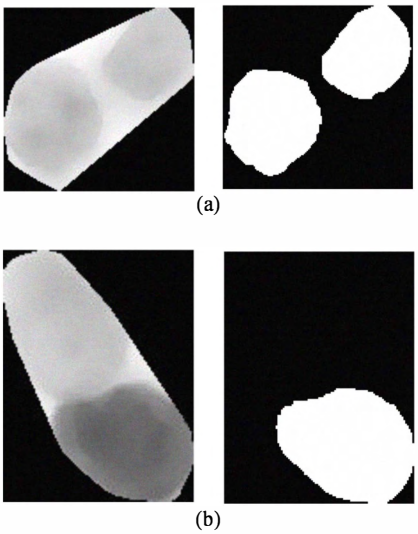
$$\sigma(n) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(I_i - \mu)^2} \qquad (2)$$

where N is the pixel number of the region, I_i is the intensity of each pixel, μ is the mean intensity of the region. High $\sigma(n)$ indicates that the region intensities are spread out over a large range of values, *i.e.* it is probable that stained and normal nuclei coexist in the region.

If the region with small solidity or high intensity standard deviation, it is probable a touching region. We also note that region area is also an important factor for classification. Regions with small areas should be single nucleus. Based on the meaningful features extracted from each region, we experimentally set thresholds of each feature for classification. In this way, the touching regions are identified for further segmentation. Results of classification are shown in Fig. 4.

D. Second Level Segmentation

Touching nuclei regions need to be further segmented. In each convex hull region, Otsu's algorithm to threshold the region is applied in order to split clustered nuclei. This method produces a satisfactory segmentation for the clustered nuclei whose intensity standard deviation is low. However, for clustered region in which abnormal and normal nuclei coexist, only abnormal nuclei can be segmented because the region with low intensity is regarded as background. Fig. 5 illustrates the different segmentation results using Otsu's algorithm in each convex hull region.



(a)

(b)

Fig. 5. Applying Otsu's algorithm in convex hull regions (a) with low intensity standard deviation, (b) with high intensity standard deviation. Left column is the convex regions; right column is the segmentation results. We observe in (a), two nuclei are segmented clearly; in (b), only nucleus with high intensity is segmented.

In the case of Fig. 5(b), we apply condition erosion method in [10]. Condition erosion means eroding only when the size of the object is larger than the predefined threshold. Two erosion steps are applied which are coarse erosion and fine erosion. The rule is that first coarse erosion using coarse structures, then fine erosion using fine structures. Considering the shapes of nuclei are like ellipses, two structures are defined as shown in Fig. 6.

We set two thresholds T1 and T2 by experiment. Firstly erode iteratively with coarse structure until the size of objects is smaller than T1. Then apply fine erosion iteratively until the size of the objects is smaller than T2. Finally, all the touching nuclei regions are separate as shown in Fig. 7. Final segmented image is shown in Fig. 8.

0	1	0
1	1	1
0	1	0

(a)

0	0	0	1	0	0	0
0	0	1	1	1	0	0
0	1	1	1	1	1	0
0	1	1	1	1	1	0
0	1	1	1	1	1	0
0	0	1	1	1	0	0
0	0	0	1	0	0	0

(b)

Fig. 6. Erosion structures. (a) Fine structure. (b) Coarse structure.

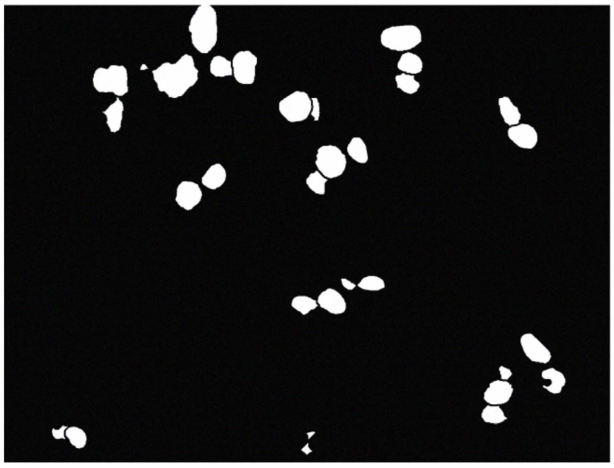


Fig. 7. Second level segmentation.

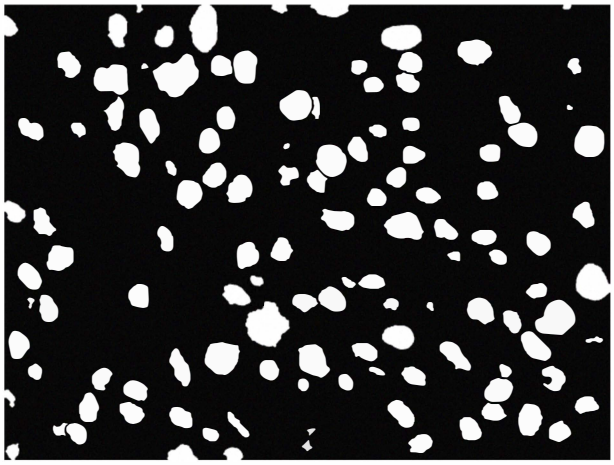


Fig. 8. Final segmented image.

E. Abnormality Identification

Combined with the abnormal nuclei information got in the first level (Fig. 3(a)), abnormal nuclei are marked in the final segmented nuclei image. We define abnormality by (3)

$$\text{abnormalit y} = \frac{\text{abnormal nuclei}}{\text{all nuclei}} \tag{3}$$

Note that the abnormality is the ratio of abnormal nuclei to all nuclei. Fig. 9 presents the abnormality detection result. In this image, N represents normal nuclei, A represents abnormal nuclei. The abnormality is 41.5%.

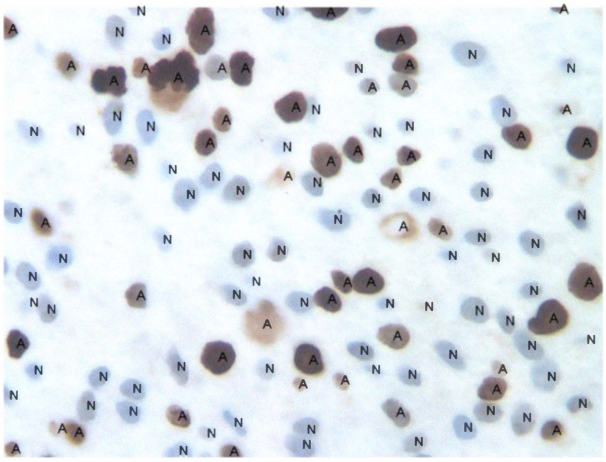


Fig. 9. Final automatic abnormality detection results. N represents normal nuclei; A represents abnormal nuclei. The abnormality is 41.5%.

I. Experimental Results

Our dataset has 10 images with resolution of 1200 * 1600 pixels. There are around 120~180 nuclei in each image, 1674 nuclei in total. We randomly choose five images as training images, others as testing images. We evaluate our method using abnormal nuclei accuracy and nuclei accuracy by (4).

$$\begin{aligned} \text{abnormal accuracy} &= \frac{\text{detected as abnormal nuclei}}{\text{abnormal nuclei}} \\ \text{nuclei accuracy} &= \frac{\text{detected as nuclei}}{\text{all nuclei}} \end{aligned} \tag{4}$$

Equation (4) indicates the detection accuracy in terms of abnormal nuclei and all nuclei respectively. Our method achieves abnormal nuclei accuracy 93.55% and nuclei accuracy 95.8%. Experimental results demonstrate our method is capable of detecting most of the nuclei and abnormal nuclei. It provides important diagnosis information effectively.

II. Conclusions

Computer aided abnormality detection in cervical tissue provides a basis in early diagnosis of cervical cancer. In the paper, a multi-level segmentation approach for abnormality detection of microscopy images in the cervix uteri is proposed. The method is composed of two level segmentations. The first level segmentation is used to segment abnormal nuclei regions and all nuclei regions. However, under-segmentation exists in all nuclei regions because of clustered nuclei. In order to classify touching nuclei and separated nucleus, meaningful features such as solidity, intensity standard deviation and area are extracted from regions of interest. Thresholds are set experimentally for classification. After classifying touching regions and separated regions, second level segmentation is performed in

the touching regions in order to partition clustered nuclei into separated nucleus. Consequently all nuclei regions are separated nucleus. Combined with the abnormal information obtained in the first level segmentation, abnormality *i.e.* ratio of abnormal nuclei to all nuclei is achieved. The proposed method is tested on dataset consists of 10 images. The experimental results demonstrate our method produces a satisfactory segmentation. In the future we will focus on testing more cervical tissue data and comparing our method with state-of-the-art method.

ACKNOWLEDGMENT

The authors would like to thank the Department of Pathology, Liverpool Hospital, Australia for providing microscopy images.

REFERENCES

[1] P. C. Bamford, "Segmentation of cell images with application to cervical cancer screening." PhD thesis, the University of Queensland, 1999.

[2] P. Bamford and B. Lovell, "Unsupervised cell nucleus segmentation with active contours," *Signal Processing,* vol. 71, pp. 203-213, 1998.

[3] B. L. Luck, K. D. Carlson, A. C. Bovik, and R. R. Richards-Kortum, "An image model and segmentation algorithm for reflectance confocal images of in vivo cervical tissue," *IEEE Transactions on Image Processing,* vol. 14, pp. 1265-1276, 2005.

[4] H. G. Acosta-Mesa, Z. Barbara, H. V. Ríos-Figueroa, N. Cruz-Ramírez, A. Marín-Hernández, R. Hernández-Jiménez, B. E. Cocotle-Ronzón, and E. Hernández-Galicia, "Cervical cancer detection using colposcopic images: a temporal approach," in *Proceedings of the Sixth Mexican International Conference on Computer Science (ENC'05)*, Puebla, Mexico, pp. 158-164, 2005.

[5] E. Njoroge, S. R. Alty, M. R. Gani, and M. Alkatib, "Classification of cervical cancer cells using FTIR data," in *Proceedings of the 28th IEEE EMBS Annual International Conference*, New York City, USA, pp. 5338-5341, 2006.

[6] M. D. Jeffers, S. J. Oakes, J. A. Richmond, and E. M. Macaulay, "Proliferation, ploidy and prognosis in uterine smooth muscle tumours," *Histopathology,* vol. 29, pp. 217-223, 1996.

[7] M. Park, J. S. Jin, M. Xu, W. S. F. Wong, S. Luo, and Y. Cui, "Microscopic image segmentation based on color pixels classification," in *The First ACM International Conference on Internet Multimedia Computing and Service*, Yunnan, China, pp. 61-77, 2009.

[8] G. Xiong, X. Zhou, and L. Ji, "Automated segmentation of drosophila RNAi fluorescence cellular Images using deformable models," *IEEE Transactions on Circuits and Systems,* vol. 53, pp. 2415-2424, 2006.

[9] N. Malpica, C. O. d. Solórzano, J. J. Vaquero, A. Santos, I. Vallcorba, J. M. García-Sagredo, and F. d. Pozo, "Applying watershed algorithms to the segmentation of clustered nuclei," *Cytometry,* vol. 28, pp. 289-297, 1997.

[10] X. Yang, H. Li, and X. Zhou, "Nuclei segmentation using marker-controlled watershed, tracking using mean-shift, and Kalman filter in time-lapse microscopy," *IEEE Transactions on Circuits and Systems,* vol. 53, November 2006.

[11] Y. L. Fok, J. C. K. Chan, and R. T. Chin, "Automated analysis of nerve-cell images using active contour models," *IEEE Transactions on Medical Imaging,* vol. 15, pp. 353-368, 1996.

[12] R. C. Gonzalez and R. E. Woods, *Digital image processing*: Prentice Hall, 2008.

[13] J. S. Jin, "Applications of Statistical Methods in Medical Imaging," in *Advanced Medical Statistics*, Y. Lu and J.-Q. Fang, Eds.: World Scientific Publishing Company, pp. 379-406, 2003.