Shape Feature Based Automatic Abnormality Detection of Cervico-Vaginal Pap Smears

Mrinal Kanti Bhowmik[1], Niharika Nath[2], Abhijit Datta[3], and Anjan Kumar Ghosh[1]
[1] Department of Computer Science and Engineering, Tripura University, Suryamaninagar-799022, Tripura, India
[2] Department of Life Sciences, New York Institute of Technology, 1855 Broadway, New York, NY 10023
[3] Department of Pathology, Agartala Government Medical College, Agartala-799006, Tripura, India
Email: mrinalkantibhowmik@tripurauniv.in, nnath@nyit.edu, agmc@rediffmail.com, anjn@ieee.org

***Abstract*—Early detection of cervical cancer involves visual screening for changes in cellular morphology through microscopic analysis of Pap smears. Cytological interpretation by conventional microscopy of abnormal Pap smears performed manually is time-consuming, observer dependent and error prone. The aim of this study is to discriminate abnormal squamous cells from normal ones by quantitative image analysis of cervico-vaginal single cells with specific focus on the structure of the nuclei. In this study: 1) Six discriminative features such as nuclear area, nuclear perimeter, equivalent diameter, major axis length, minor axis length and convex area were selected and statistically justified, 2) A new dataset of 100 Pap smear cell images were collected from North-East Indian Regional population for the experimentation, and 3) Ground truth images of Pap smear cell dataset created by medical experts were compared with the automatically-segmented images with respect to the selected shape features. The cell boundary was segmented using greedy active contour model. Based on these six discriminating features, relevant cell images were classified as normal and abnormal using Support Vector Machine. Our method reports accuracy of 97.33%. Additionally, the proposed framework was applied to a known Pap smear benchmark dataset, to which we report an accuracy of 90.21%.**

***Index Terms*—cervical cancer, cell imaging, pap smear test, nucleus, segmentation, greedy active contour model, feature extraction, support vector machine**

I. Introduction

Cytology interpretation is an efficient and well-established technique for analysis and diagnosis of many diseases. It is considered as an essential method of early detection of dysplasia or precancerous cells. Cervical cancer is one of the most common cancer problems among women in the world [1]. In developing countries, it is the most common cause of cancer-related deaths in women [2]. There are various risk factors; however cervical cancer is mainly caused by Human Papillomavirus (HPV) types 16 and 18. Early detection and immediate follow-up treatment are considered as strong control measures against this cancer. Routine screening such as through a Pap smear test using Pap smear images is the most effective procedure for examining the precancerous cervix [3]-[5]. Dysplastic cells show precancerous changes; however, the visual interpretation of microscopic images are time-consuming, annoying and sometimes an error-prone process [5], [6]. Furthermore, due to different staining procedures, there exist variances in dye concentration and illumination of the cells and additional disturbances such as air drying, mucus or bacterial presence which make the recognition and interpretation process a more difficult task [7], [6]. In visual analysis of cells in Pap smears, the correct characterization of microscopic slides and drawing any conclusion depend mostly on the characteristic appearance of the nucleus in the cells. From pathological point of view, the nucleus of the infected cell may be excessively enlarged, or display irregularity in shape, with possibility of increased DNA content and irregular chromatin density [4]. The identification and estimation of these significant changes in the nucleus contribute, in part, to the discrimination of normal and abnormal cells in the microscopic images [8]. Therefore, it is essential to have an accurate determination of the nucleus area, nucleus shape and other nuclear characteristics of cells in cell images obtained by microscopy for accurate analysis. The major challenges for the detection of normal and abnormal cells are the exact detection of the location of cell nuclei automatically and accurate determination of the nuclear boundaries. Reliable cell nucleus segmentation methods are necessary, and many reports have used segmentation methods [9]-[12]. However, incorporation of nuclear structural features including accurate boundary determination has the potential to further improve the discrimination. Therefore, in this study several shape features for the nucleus were considered. The present study focuses on the quantitative analysis of the structure of cell nuclei from Pap smears in order to differentiate between normal and abnormal (premalignant/ malignant) squamous cells. The contributions of this article comprise i) creation of a new dataset of 100 numbers of Pap smear cell images of North-East Indian regional population, ii) creation of ground truth images of Pap smear cell dataset by medical experts, iii) comparison with the automatically segmented images with respect to their shape features, and iv) identification of abnormality present in the cervico-vaginal microscopic images. Initially, Pap smear images were enhanced and the boundary of the cell nuclei

Manuscript received June 28, 2017; revised November 22, 2017.

doi: 10.18178/joig.5.2.52-58

subsequently detected by greedy active contour model. The abnormality of the microscopic images was detected by observing the nucleus part of the cell. For this, eleven different shape features were considered for extraction from both normal and abnormal Pap smear cell images. Based on the t-test for adjudicating statistical significance on these eleven features only six of them could qualify the test and accordingly exploited for classification task. Classification of these cell images based on discriminant features was done using support vector machine (SVM) classifier. Experimentation and analysis are further conducted on a benchmark database to validate the effectiveness of the technique. Section II includes the relevant literature survey. Section III elaborates the proposed method; and Section IV illustrates the process of new regional database creation. Brief discussions on the statistical analysis of different features and performance of SVM classifier are demonstrated in Section V, followed by conclusion in Section VI.

II. Related Work

Classification of cervico-vaginal cells from Pap smear images on the basis of histology and cytology have used several approaches according to reported research studies [9]. In order to detect the abnormality of microscopic cell images, the primary task is the detection of concerned cell nuclei, which reflects significant changes when the cell is cancerous. The localization of cell nuclei and the detection of the boundary of the cell nuclei are one set of the challenging tasks. Bamford *et al.* [9] suggested a useful method for nucleus boundary detection. They implemented Viterbi search-based dual active contour model to detect the nucleus boundary where the search space is bounded by the initial positions of inner/outer contours within/outside of the nucleus, and obtained a remarkable 99% success rate on a dataset including 20130 pap stained cell images [9]. Median filter has been found effective for reducing noise in Pap smear cell images [13], [8]. The Hough transformation [10] and morphological reconstruction process [11], [14] were also used successfully to detect the cell nuclei. The Fuzzy c-means clustering algorithm [15], [12], [14] was used effectively to split the segmented images in terms of abnormality and normality. Classification using support vector machine [17] and K-means clustering [8], [16], [13] yielded impressive results. Deformable model or GVF [15], [10], Active contour model [16], watershed transform algorithm [15], [16] have also been proved as effective ways of nucleus boundary detection. Yung-Fu Chen *et al.* [17] utilized semiautomatic cellular image analysis to segment nucleus and cytoplasm by Normalized Cross Correlation (NCC) detector and extracted different texture features to discriminate dysplastic cells from normal ones. Agarwal *et al.* used De-correlation stretching to enhance the discrimination between nucleus and background [18]. Cell nucleus segmented by Mean-shift based approach and adaptive image thresholding were also performed on the resultant image to convert into a binary image [18]. The Norup benchmark database [19] was used for validation.

III. Proposed Framework

Appropriate choice of shape based features useful for ascertaining the presence of abnormality in Pap smear images is absolutely important. An adequate and more appropriate feature set if used may provide improved classification accuracy. The framework of this study to classify a Pap-smear cytology image as depicted in Fig. 1.

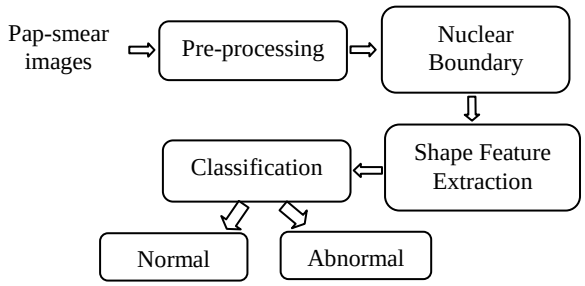

Figure 1. Framework of the proposed technique.

A. Pre-Processing

For abnormality detection of cervico–vaginal microscopic images, the boundary between nucleus and cytoplasm in the cell are required to be clearly defined. The microscopic images are usually poor quality color images which are stained to color the cells and are affected by noise. To ensure homogeneity in the image content for subsequent processing, such cervico-vaginal Pap smear color images are converted into corresponding grey representation followed by contrast enhancement.

B. Region of Interest (ROI) Segmentation

As the size and shape of the nucleus indicate the presence of abnormality, identification of changes in the nucleus contributes to the discrimination of normal and abnormal cells in microscopic images. Dysplastic cells (abnormal cells) generally have longer or stretched nucleus and often oddly shaped whereas normal cells have a comparatively smaller and round/oval nucleus [8].

Objects having closed boundary in an image are usually subjected to different edge detection algorithms for isolation of their boundaries. In the present scope, we are inclined to use greedy active contour model instead, for identifying object contour present in the image. This is because greedy active contour model is adaptive to extraction of wide range of objects with complex and irregular shape as well as contour with broken boundaries [8]. For the sake of self-sufficiency, we illustrate the active contour model in brief. First, the user needs to specify contour points outside the region of interest and then the initialized points are involved in an iterative method for searching the local neighborhood around them to identify a set of new contour points having lower energy [20]. By moving towards lowest energy position, it extracts the desired object. The traditional snake is represented by a vector (v(s)) as in (1) [20].

$$v(s) = (x(s), y(s)) \quad (1)$$

In (1), x(s) and y(s) are the snake points in x-coordinate and y-coordinate, respectively. The energy function is a combination of internal energy due to bending and stretching of contour and image energy. The energy function (E) is represented by (2) [20].

$$E = \int (\alpha(s)E_{cont} + \beta(s)E_{curve} + \gamma(s)E_{image})ds \quad (2)$$

In (2), the E_{cont} and E_{curve} are first and second-order continuity constraints with their weighted coefficients. The E_{image} measures image quantity, namely the edge strength or intensity. Weighted coefficients in (2) are considered as Elastic constant (α)=1, Curvature constant (β)= 1, Image energy constant (γ)=1.2 as per William *et al.* [20].

C. Shape Based Feature Extraction and Selection

After segmentation of the cell nucleus as the ROI using greedy Active Contour Model (ACM), eleven shape features, are illustrated bellow, were considered because of their significant role in identifying the abnormality.

1) Nucleus Area (NA)

The nucleus area was determined from the total number of pixels in the segmented region [5], [20]-[23].

2) Nucleus Perimeter (NP)

The nucleus perimeter was calculated by counting the number of boundary points [5], [16], [20]-[23].

3) Nucleus Roundness (NR)

The nucleus roundness was calculated as the ratio between the actual area of the nuclear and the length of the perimeter of the nuclear. Circle has the roundness value exactly 1. The roundness value signifies the deviation of the shape of a cell from the circular shape. The nuclear roundness sometimes called circularity has been calculated by using the formulas [16], [13], [22].

$$Roundness = \frac{4\pi area}{perimeter^2} \quad (3)$$

4) Equivalent Diameter (ED)

Equivalent diameter specifies the diameter of a circle with the same area as the region [16], [21].

$$Equivalent\_diameter = \sqrt{\frac{4area}{\pi}} \quad (4)$$

5) Major Axis Length (MAJ)

To find major axis length, a pair of points on the boundary is found whose Euclidean distance from each other is larger than any other pair of points on the boundary line. The line used to connect these two points is called major axis length [20], [24].

6) Minor Axis Length (MIN)

Minor Axis length is the axis perpendicular to the major axis length [20], [24].

7) Elongation (EN)

The elongation was calculated through minor axis length divided by major axis length. Circle has the elongation value exactly 1. From the elongation value, assumptions are generally made regarding how much a cell deviated from the circle [20], [24].

$$Elongation = \frac{Minor\_Axis\_Length}{Major\_Axis\_Length} \quad (5)$$

8) Eccentricity (ECC)

It defines to evaluate the deviation of the objects shape from a symmetric or regular circular shape. It allows for tracking how much an abnormal cell nucleus differs how much from a normal cell nucleus. The values of eccentricity vary between 0 to 1 [21].

$$Ecentricity = \sqrt{1 - \frac{Minor\_Axis\_Length^2}{Major\_Axis\_Length^2}} \quad (6)$$

9) Convex Area (CA)

Convex area specifies the total number of pixels in the convex hull image [20].

10) Solidity (SO)

Solidity describes the quantity of the common pixels in the convex hull and in the region [21].

$$Solidity = \frac{Area}{Convex\_Area} \quad (7)$$

11) Extent (EX)

Extent defines the ratio of pixel in the region to pixels in the total bounding box [21].

$$Extent = \frac{Area}{Area\_of\_the\_bounding\_box} \quad (8)$$

Among these eleven shape features, some features may contribute significantly in comparison to the others. Pap smear images were evaluated for all eleven shape features to identify which features contribute with significance with $p < 0.001$ [25], [26].

D. Classification

These shortlisted six features were fed to the support vector machine to classify cell images as normal or abnormal cells.

IV. New Database Design and Development

Performance of the proposed framework was evaluated on the created North-East Indian regional population dataset.

A. Database Description

Our database was created in the Pathology Departments of Agartala Government Medical College in North East India and composed of 100 conventional Pap-stained cervical cell images. The images were taken randomly from 20 patients (5 images from each patient) of North-East Indian region population and found befitting for image analysis, acquired through a digital camera (OLYMPUS SP 350) adapted to a light microscope (OLYMPUS CX41) having 80 megapixels resolutions and stored in JPEG format of 132 X 158 pixels. As illustrated in Table I, among 20 patients, 10 patients had normal and 10 patients had abnormal cells based on the observation of the pathology expert at the hospital.

B. Cell Staining

Cells obtained from cervico-vaginal swabs were stained. Papanicolaou-type stains were used for identification of features related to nucleus [27]. Fig. 2 depicts the steps used for cell preparation.

Papanicolaou stain require immediate and rapid fixation in alcohol to preserve the details of cells and to avoid the smears drying out. Ethanol (95%) was used [27] and haematoxylin stain, which is a natural dye which stains the nucleus part of a cell and Orange green 6 which stain the cytoplasm of keratinized cells, were used. Polychromic stain, a mixture of Light Green SF and Eosin G, were used as the cytoplasmic stain. Clearing in xylol produced cellular transparency prior to mounting. Mounting was performed with DPX which is a mixture of distyrene, a plasticizer, and xylene. Representative normal microscopic cell images and abnormal microscopic cell images are shown in Fig. 3.

TABLE I. SAMPLES IN THE CREATED DATABASE

Category of Pap Smears	Number of Patients	Number of Pap Smears Per Patient	Number of Pap Smear Images
Normal	10	5	50
Abnormal	10	5	50

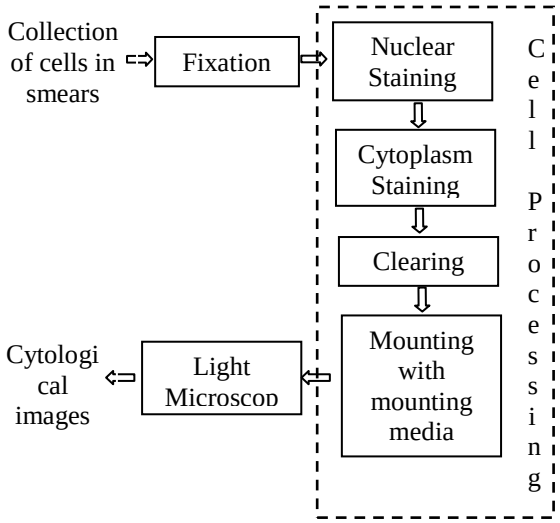

Figure 2. Stages of data preparation.

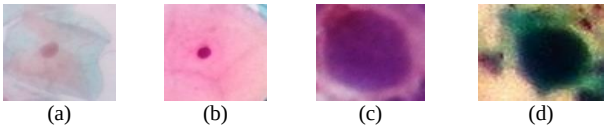

Figure 3. Created dataset of Pap smear images. (a) and (b) indicate Normal Pap smear cell images, (c) and (d) indicate abnormal Pap smear cell images.

C. Ground Truth Generation

Ground truth generation is a crucial work to verify the competence of the proposed framework. Generally ground truths are generated using manual segmentation by the physicians and medical experts to visually identify the presence of abnormality. In the terms of medical image processing, these ground truths can also be used to measure the performance of different segmentation algorithms. Norup and another group Martin *et al.* consider CHAMP software to segment Pap smear images. Unfortunately, CHAMP software cannot provide a satisfying segmentation performance, especially for abnormal cells [8].

Three distinct set of ground truths for Pap smear cell database were generated by three independent medical experts individually using the widely used Sefexa image segmentation tool [28] while keeping all parameters same for the whole database. Sefexa makes the manual segmentation of nucleus region of Pap smears appear as depicted in Fig. 4(d).

V. EXPERIMENTAL RESULTS AND DISCUSSION

The objective of this section is to investigate the performance of the automated framework compared to ground truth. Fig. 4 illustrates subsequent phases of nuclear delineation using greedy active contour model and typical one among the three separate ground truths (due to space limitation) prepared manually by three individual medical experts.

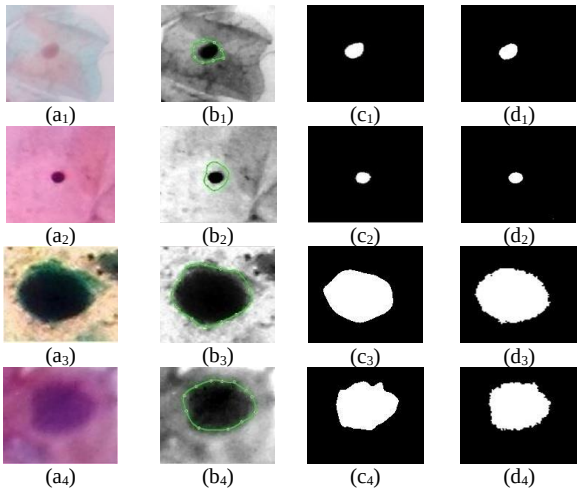

Figure 4. ($a_1 - a_4$) Normal and abnormal Pap smear cell images of created dataset, (b_1-b_4) represents initialization of ACM on input, (c_1-c_4) segmented image using ACM, (d_1-d_4) Ground truth image.

Comprehensive range of eleven shape features extracted from automated segmented nucleus was compared with the range of features extracted manually from the nucleus. The comparative results are shown in Table II, wherein a sharp discrimination was observed between normal and abnormal cell images only for the shortlisted six shape features.

Extraction of the shape features was followed by the evaluation of the statistical significance of theses shape features in abnormality detection, followed by SVM classification [29]. The average of these feature values along with their standard deviation are listed in Table III and statistical significance of these features are tested by using independent sample t-test [30], [31] with significance level of 0.1%. The shape features whose t-value is greater than the critical value or the p-value<0.001 are considered to be statistically significant and the features, whose t-value is less than the critical

value or p-value>0.001 are considered to be statistically insignificant in abnormality detection. As illustrated in Table III, out of 11 shape features, only 6 features: Nucleus Area, Nucleus Perimeter, Equivalent diameter, Major Axis Length, Minor Axis length and Convex area were found to be statistically significant (p<0.001) in discriminating abnormal cells from the normal ones.

TABLE II. ELEVEN SHAPE FEATURES OBTAINED AUTOMATICALLY AND FEATURES FROM GROUND TRUTH IMAGE THEREOF

Shape Features	Range of Features Obtained Automatically From Segmented Image		Range of Features Obtained Automatically From Ground Truths	
	Normal cell	Abnormal cell	Normal cell	Abnormal cell
NA (μm^2)	99-1282	3052-6085	107-1203	1534-10239
NP (μm)	54-166	212-496	52-164	188-508
NR	0.415-0.585	0.343-0.595	0.429-0.611	0.295-0.602
ED (μm)	11.227-40.402	50.802-120.784	11.672-39.137	44.194-114.178
MAJ (μm)	15.183-48.365	55.123-132.43	15.725-46.069	47.633-137.198
MIN (μm)	8.386-34.802	44.2-111.117	8.760-35.179	32.094-99.896
EN	0.525-0.934	0.360-0.977	0.514-0.952	0.312-0.924
ECC	0.357-0.852	0.215-0.933	0.304-0.857	0.382-0.950
SO	0.887-0.973	0.916-0.988	0.326-0.981	0.580-0.973
EX	0.605-0.819	0.449-0.783	0.163-0.869	0.273-0.785
CA (μm^2)	105-1326	2112-11690	112-2083	1618-10718

TABLE III. MEAN AND STANDARD DEVIATION OF SHAPE FEATURES OF CREATED DATASET WITH SIGNIFICANCE VALUE

Shape Features	Mean ±Standard Deviation		t value	p value
	Normal cell	Abnormal cell		
NA (μm^2)	456.94±250.10	4029.78±1410.06	864.523	**p<0.001**
NP (μm)	103.28±26.18	315.36±84.17	209.688	**p<0.001**
NR	0.50±0.037	0.52±0.047	1.926	p>0.001
ED (μm)	23.30±6.26	72.48±19.85	102.253	**p<0.001**
MAJ (μm)	27.46±7.81	85.98±25.40	106.733	**p<0.001**
MIN (μm)	20.26±5.52	62.87±18.57	94.270	**p<0.001**
EN	0.74±0.11	0.74±0.14	0.466	p>0.001
ECC	0.64±0.13	0.62± 0.16	0.466	p>0.001
SO	0.95±0.01	0.96±0.01	1.224	p>0.001
EX	0.70±0.04	0.70±0.06	0.766	p>0.001
CA (μm^2)	478.04±258.03	4589.22±2294.19	892.571	**p<0.001**

For each data sample in the created dataset and benchmark dataset, the discriminative shape features were used as features for the support vector machine classifier. In this study the classification is treated as a two class pattern classification (Normal and Abnormal) problem. For training, support vector machine linear kernel was used. To obtain proper parameters for kernel function, 5-fold cross validation was performed for both the databases. Here, 70% of the data were selected randomly from both datasets for training and 30% were for testing. As a result, the accuracy of the method was 97.33%. Additionally, the proposed framework was applied to a known Pap smear benchmark dataset, and we determined an accuracy of 90.21%. There are only few reports in literature on this benchmark database that classify cell images as normal and abnormal. For example, Sa *et al.* reported 89.64% accuracy on the same Pap smear benchmark database for two-class (normal and abnormal) problem using gravitational method [32]. Mbaga *et al.* mentions a promising result with an average accuracy of 92.961% [33]. Paul *et al.* report impressive results of 92.37% and 98.31 % accuracy for minimum distance and K-nearest neighbor classifiers, respectively [34]. Sun *et al.* report accuracy up to 94.44% using random forest classifier [35].

VI. CONCLUSION

In this study a North-East Indian regional population dataset were developed with its corresponding ground truth images that were provided by medical experts. Out of the eleven shape features, six discriminative and statistically justified shape features were extracted to classify Pap smear cell images into normal and abnormal smears using support vector machine. The results of these studies reveal additionally that support vector machine gives an impressive performance with potential to provide higher accuracy.

ACKNOWLEDGMENT

The research work was supported by the Grant BT/533/NE/TBP/2013 from the Department of Biotechnology, Government of India to M. K. Bhowmik conducted at the Biometrics Laboratory of Computer Science and Engineering. NYIT-ISRC-2016 and -ISRC-2017 to N. Nath supported histochemical analysis and software.

Mrinal Kanti Bhowmik and Niharika Nath are equal corresponding authors.

REFERENCES

[1] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics," *CA Cancer J. Clin.*, vol. 67, no. 1, Jan. 2017.

[2] L. A. Torre, R. L. Siegel, E. M. Ward, and A. Jemal, "Global cancer incidence and mortality rates and trends--An update," *Cancer Epidemiol Biomarkers Prev.*, pp. 16-27, Dec. 2016.

[3] H. Lee, "Segmentation of overlapping cervical cells in microscopic images with superpixel partitioning and cell-wise contour refinement, " in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition Workshops*, 2016, pp. 63-69.

[4] R. Saha, M. Bajger, and G. Lee, "Spatial shape constrained fuzzy C-Means (FCM) clustering for nucleus segmentation in pap smear images," presented at Digital Image Computing: Techniques and Applications (DICTA), Gold Coast, Queensland, Australia, Nov. 2016.

[5] B. Kangkana, M. Chowdhury, L. B. Mahanta, M. K. Kundu, and A. K. Das, "Automated classification of Pap smear images to detect cervical dysplasia," *Computer Methods and Programs in Biomedicine*, vol. 138, pp. 31-47, Oct. 2017.

[6] Y. Song, *et al.*, "Accurate cervical cell segmentation from overlapping clumps in Pap smear images," *IEEE Trans. on Medical Imaging*, vol. 36, pp. 288-300, Jan. 2017.

[7] Y. Song, *et al.*, "Segmenting overlapping cervical cell in Pap smear images," presented at 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), Prague, Czech Republic, April 13-16, 2016.
[8] M. H. Tsai, Y. K. Chan, Z. Z. Lin, S. F. Yang-Mao, and P. C. Huang, "Nucleus and cytoplast contour detector of cervical smear image," *Pattern Recognition Letters*, vol. 29, no. 9, pp. 1441-1453, July 2008.
[9] P. Bamford and B. Lovell, "Unsupervised cell nucleus segmentation with active contours," *Signal Processing*, vol. 71, no. 2, pp. 203-213, Dec. 1998.
[10] A. Garrido and N. P. D. L. Blanca, "Applying deformable templates for cell image segmentation," *Pattern Recognition*, vol. 33, no. 5, pp. 821-832, May 2000.
[11] M. E. Plissiti, "Automated segmentation of cell nuclei in PAP smear images," in *Proc. IEEE International Special Topic Conf. on Information Technology in Biomedicine, Greece*, 2006.
[12] M. E. Plissiti, "Accurate localization of cell nuclei in PAP smear images using gradient vector flow deformable models," in *Proc. Third International Conf. on Bio-inspired Systems and Signal Processing*, Valencia, Spain, January 20-23, 2010, pp. 284-289.
[13] G. Naghdy, "Morphological characteristics of cervical cells for cervical cancer diagnosis," in *Proc. 2nd International Congress on Computer Applications and Computational Science*, 2012, pp. 235-243.
[14] M. E. Plissiti, C. Nikou, and A. Charchanti , "Automated detection of cell nuclei in Pap smear images using morphological reconstruction and clustering," *IEEE Trans. on Information Technology in Biomedicine*, vol. 15, no. 2, pp. 233-241, Mar. 2011.
[15] M. E. Plissiti, "Watershed-based segmentation of cell nuclei boundaries in Pap smear images," in *Proc. 10th IEEE International Conf. on Information Technology and Applications in Biomedicine*, Nov. 2010, pp. 1-4.
[16] M. E. Plissiti, C. Nikou, and A. Charchanti, "Combining shape, texture and intensity features for cell nuclei extraction in Pap smear images," *Pattern Recognition Letters*, vol. 32, no. 6, pp. 838-853, April 2011.
[17] Y. F. Chen, *et al.*, "Semi-automatic segmentation and classification of pap smear cells," *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 1, pp. 94-108, Jan. 2014.
[18] P. Agarwal, A. Sao, and A. Bhavsar, "Mean-shift based segmentation of cell nuclei in cervical PAP-smear images," presented at Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG), Fifth National Conference, Dec. 2015, pp. 1-4.
[19] J. Norup, "Classification of Pap-smear data by tranduction neuro-fuzzy methods," Master's thesis, Technical University of Denmark, DTU, DK-2800 Kgs, Lyngby, Denmark, 2005.
[20] D. J. Williams and M. Shah, "A fast algorithm for active contours and curvature estimation," *CVGIP: Image Understanding*, vol. 55, no. 1, pp. 14-26, Jan. 1992.
[21] Y. M. George, H. H. Zayed, M. I. Roushdy, and B. M. Elbagoury, "Remote computer-aided breast cancer detection and diagnosis system based on cytological images," *IEEE Systems Journal*, vol. 8, no. 3, pp. 949-964, Sept. 2014.
[22] M. E. Plissiti and C. Nikou, "Cervical cell classification based exclusively on nucleus features," presented at International Conference Image Analysis and Recognition, Springer Berlin Heidelberg, June 2012, pp. 483-490.
[23] Y. Marinakis, G. Dounias, and J. Jantzen, "Pap smear diagnosis using a hybrid intelligent scheme focusing on genetic algorithm based feature selection and nearest neighbor classification," *Computers in Biology and Medicine*, vol. 39, no. 1, pp. 69-78, Jan. 2009.
[24] I. Muhimmah and R. Kurniawan, "Analysis of features to distinguish epithelial cells and inflammatory cells in Pap smear images," presented at 6th International Conference on Biomedical Engineering and Informatics (BMEI), IEEE, Dec. 2013.
[25] R. Russo, *Statistics for the Behavioural Sciences: An Introduction*, New York: Psychology Press, 2004, pp. 153-155.
[26] S. C. Gupta and D. V. Kapoor, "Fundamentals of mathematical statistics: A modern approach, Sultan Chand," New Delhi, pp. 14.16-14.17, 2000.
[27] L. G. Koss and M. R. Melamed, *Koss Diagnostic Cytology and Its Histopathologic Bases*, Lippincott Williams & Wilkins, 2006, vol. 1, p. 1067.
[28] Sefexa Image Segmentation Tool. [Online]. Available: http://www.fexovi.com/sefexa.html
[29] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining and Knowledge Discover*, vol. 2, pp. 121–167, June 1998.
[30] L. G. Grimm and P. R. Yarnold, *In Reading and Understanding Multivariate Statistics*, American Psychological Association, 1995, pp. 169-215.
[31] S. W Huck, W. H. Cormier, and W. G. Bounds, *Reading Statistics and Research*, New York: Harper & Row, 1974, pp. 69-71.
[32] J. J. D. M. Sa and A. R. Backes, "A color texture analysis method based on a gravitational approach for classification of the pap-smear database," presented at IEEE International Conference on Image Processing, 2014, pp. 2280-2284.
[33] A. H. Mbaga and P. ZhiJun, "Pap smear images classification for early detection of cervical cancer," *International Journal of Computer Applications*, vol. 118, no. 7, pp. 10-16, May 2015.
[34] P. R. Paul, M. K. Bhowmik, and D. Bhattacharjee, "Automated cervical cancer detection using Pap smear images," *Advances in Intelligent Systems and Computing, Springer*, vol. 335, pp. 267-278, Dec. 2015.
[35] G. Sun, S. Li, Y. Cao, and F. Lang, "Cervical cancer diagnosis based on Random Forest," *International Journal of Performability Engineering*, vol. 13, no. 4, pp. 446-457, July 2017.

Mrinal Kanti Bhowmik received his Bachelor of Engineering (B.E.) degree in Computer Science & Engineering from Tripura Engineering College, Govt. of Tripura, in 2004 and Master of Technology (M.Tech) degree in Computer Science and Engineering from Tripura University (A Central University), India, in 2007. In 2014, he received his Ph.D. (Engineering) degree from Jadavpur University, Kolkata, India. Since 2010, he has been working as an Assistant Professor at Tripura University. He has successfully completed two DeitY funded project and currently is the principal investigator of the DBT-Twinning project, Govt. of India. His current research interests are in the field of medical imaging, biometrics, information fusion, information security etc. He is also a member of the IEEE (USA).

Niharika Nath specializes in the cell biology of cancer and cancer prevention. She obtained B. Pharm degree from Delhi University, M.Tech Biotechnology from Jadavpur University. She earned a PhD from the Indian Institute of Technology-Delhi in Biochemical Engg and Biotechnology; and pursued post-doctoral transnational research on carcinogenesis and drug design studies of preventable cancers at Columbia University, New York and at City College, New York. She is currently Associate Professor and Chairperson, the Department of Life Sciences, New York Institute of Technology, New York. She promotes undergraduate research and scholarship as Councilor at the Council on Undergraduate Research, Washington DC.

Abhijit Datta received his Bachelor of Medicine, Bachelor of Surgery (MBBS) degree in Medical Science from Kottayam Medical College, Kerela, India, in 1986. In 1999, he received his Doctor of Medicine (MD) in Pathology from All India Institute of Medical Sciences, New Delhi, India. He received FAIMER fellowship (CMCL) in Medical Education from Christian Med. College, Ludhiana, India in 2017. Since 2011, he has been working as an Associate Professor at Department of Pathology, Agartala Govt. Medical College, and Tripura, India. His interests are in the fields of Cytopathology, Haematopathology and Medical education system.

Anjan Kumar Ghosh obtained his MS from the SUNY at Stony Brook and his Ph.D. (Electrical Eng.) in 1984 from Carnegie-Mellon Univ., Pittsburgh. Dr. Ghosh served as a faculty member in the Univ. of Iowa, Iowa City, USA, IIT Kanpur, India, Nanyang Tech. Univ., Singapore, the University of Oklahoma, Tulsa, USA and DA-IICT, Gandhi Nagar, India. He was the Head of the Dept. of Electrical Eng., Adv. Center for Electronic Sciences, Laser Tech. Program and the Center for Laser Tech., all at IIT Kanpur. Currently, He is the Vice-Chancellor of Tripura University (A Central University). He is a Senior Member of IEEE, a Fellow of the IETE (India) and members of SPIE, OSA and Optical Soc. of India. He has over 28 years of research and teaching experience in various areas including photonic sensors, optical communications and, optical information processing. He is a Senior Member of IEEE.