International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878, Volume-8 Issue-4, November 2019

Estimating Cell Count in Overlapped Cluster of ThinPrep Pap Smear Cell Microscopic Image

Deepa.T.P., Nagaraja Rao A

Abstract: ***The cell adequacy is one of the important factor is Pap smear test as per Bethesda reporting system. The adequate number of cells in a given sample helps to effectively check for abnormality of cells. The Pap smear sample with inadequate number of cells will be rejected and patient will be called for collecting sample once again. The cell adequacy is the number of squamous epithelial cells present in the given Pap smear sample. This paper presents an automated method which estimates the number of cells in a given Pap smear image. This method estimates a cell count even when cells are overlapped. The proposed approach first separates disjoint cells and overlapped cell cluster from cervical Pap smear image. Then the overlapped cluster is further segmented to obtain individual cells. The overlapped cluster is selected based on the area factor. That is area of overlapped cluster will be more than disjoint cells. This method uses Bitplane slicing technique for segmentation of overlapped cells. Then the total cell counts is estimated as sum of disjoint cells count and number of cells separated from overlapped cluster. The method is tested on the ThinPrep images publicly available in open access database.***

Keywords: **Bethesda, Bitplane Cervical Cell, Cell count.**

I. INTRODUCTION

The main objective of Pap test is to find abnormalities in transformation zone of uterine cervix. This test is very important for pathogenesis of cervical cancer. It is primary screening method to detect cervical infections. These infections are most commonly detected in women younger than 40. The results of Pap test help to avoid potential physical and psychological damage to the patient. The reporting of Pap test should be clear and well understandable. The Pap test reporting is done according to the Bethesda system for Cervicovaginal cytology. This Bethesda system was developed during a workshop sponsored by the National Cancer Institute in the United States. It was developed to ensure standard in Pap test reporting. In contrast to traditional Pap test classification system, the Bethesda system classifies all Pap test results with respect to adequacy. A Pap test should be repeated if a sample or specimen is unsatisfactory. Satisfactory Pap smear sample should contain an adequate number of squamous epithelial cells from transformation zone. Pap smears that are poorly preserved, with debris or inflammatory cells are considered as satisfactory but limited by certain condition. Hence evaluation of specimen adequacy is considered as one of the most important quality assurance component of Bethesda system. One of the main reason developed countries adopted liquid-based cytology (LBC) is due to its low frequency of unsatisfactory and inadequate Pap smear samples [2]. But, finding this adequacy is still challenging and tedious task even in case of LBC Pap sample when cells are overlapped. It is even more difficult due to more number of cells in the cluster, bad staining, poor contrast [3] , also, there is possibility of cells are hidden below the other cells in case of overlapping, which leads to wrong interpretation of adequacy of Pap smear sample. Hence there is a need for method which can automatically identify such overlapped cells and accurately find the adequacy. To find adequacy, it is necessary to count total number of cells in a given sample. Hence, there is a need for automated method which can find cell count in the given sample. In manual method, this cell count is not accurate for sample containing overlapped cells. As it is difficult for human to visually find the cell underneath the overlapped cells. Hence there is a need for computer assisted automated method to accurately find the total number of cells in a given Pap smear sample.

Table I: Cell count needed to perform to ascertain the correct number of cells for adequacy [2]

Type of Sample	10 x Eyepiece FN20 X 40 Objective	10 x Eyepiece FN22 X 40 objective
Surepath™	Cells per field for 15,000 cells 26	Cells per field for 15,000 cells 31

As per Bethesda System (TBS) 2001, Liquid based cytology (LBC) sample should contain a minimum of 5000 squamous cells. The cell count for LBC sample [2] is as shown in the Table I.

This paper proposes a method which uses image processing techniques to automatically segment and separates overlapped cells and count the number of cells in the given Pap smear sample.

II. LITERATURE SURVEY

Author et al in [3,4] proposed a method in which unsupervised cell segmentation using extended depth of field images which are created from different focal planes.

Manuscript published on November 30, 2019.
* Correspondence Author
Deepa.T.P.*, Research Scholar, SCOPE, VIT University, Vellore, Tamil Nadu, India. Email: deepa.sce@gmail.com
***Dr. Nagaraja Rao A,** Associate Professor, SCOPE, VIT University, Vellore – 632014, Tamil Nadu, India. Email: nagarajaraoa@vit.ac.in

Apriori weights were used in modified Otsu method for segmentation of nuclei in a cell cluster. Depending on the nuclei segmented, cells are counted and detected to further segment cytoplasm using level set algorithm. This method used ISBI 2015 challenge dataset. The proposed method could achieve dice coefficient of 0.86.

In [4], author et al proposed a method to detect cervical cell in thin liquid based cytology samples. They used logical regression classifiers with 28 features to identify cells. There was two levels of classification, first level includes acquisition of image and separation of cells from the background, measurement of morphological features of cells. The second level of classification includes extraction of features same as level 1 and applying LR classifier. The proposed method was tested on 20,000 cells. Cells were classified into epithelial, lymphoid, neutrophils and junk cells. The recognition rate was 93.2%. The false negative and false positive rate was much lower compared to traditional Pap smear analysis.

M. E. Plissiti et al proposed a method [5] in which automatically detects nuclei in Pap smear images. The locations of centroids in each nuclei if found and used as priori knowledge to find the circumference. Using distance dependent rule and classification algorithms, unwanted artifacts were removed. In every step, the locations of centroids were refined to get better classification results. The proposed method was tested on 5617 squamous cells.

Y. Jusman et al reviewed 103 scholarly and indexed journal paper on cervical cancer detection, feature extraction, classification using image processing. Authors explored different techniques to detect cervical cancer, comparison of manual and automated analysis. The various classification methods like artificial neural network, Support vector machine, k-nearest neighbor, linear discriminant analysis etc. They also explored that there are four ways to perform segmentation of cervical cells, based on shape, color, texture and contour. The size of the cell is expressed in terms of radius, area and perimeter. The radius is defined as the average length of radial lines towards every boundary point. Area is the number of pixels within the boundary. Perimeter is sum of distances between every neighboring boundary points. The point where major and minor axis intersects is a center [6].

Authors published a paper in which they proposed a method to classify Pap smear cells into normal and dysplastic cells using hybrid classifier. They used both neural network and support vector machine for classification. Six features like GLCM, LBP, histogram, texture, grey level and wavelet were extracted [7].

III. METHODOLOGY

The proposed method counts the total number of cells in the given ThinPrep Pap smear sample image at 40x magnification. ThinPrep is a technique in which smear sample is collected from the patient is washed in a bottle of special liquid. Then these cleaned cells are pressed on to the glass slide which prepares a monolayer of cells which are evenly distributed on the slide. Pap smear slides prepared using this technique is free from biological debris with clear background of cells. This makes cells clear and easy for further analysis as there is no additional effort is needed to remove background. Even though cells with clear background is obtained, but still there exists overlapped cells in ThinPrep slides. The proposed method segments cells in two phases. The phase 1 separates cell cluster from background, phase2 separates nuclei from cytoplasm and extracts its features. The phase 1 uses bit plane slicing to segment overlapped cells, extract the features which helps to identify nuclei and then count the number of nuclei which further gives the cell count. The steps to arrive at this result is as shown in the fig.1 and the algorithm A-1.

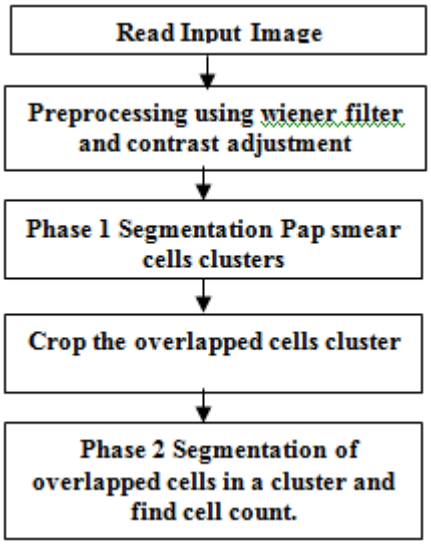

Fig. 1. Estimating Cell count in overlapped Pap smear cell cluster.

The ThinPrep Pap smear image with 670X 500 and 400x magnification is given as input to the method. It is preprocessed to remove device induced noise using Adaptive wiener filter with window size [7 7]. Also, contrast of an image is enhanced to clearly distinguish pale grey cytoplasmic area, dark grey nuclei region from white background. As ThinPrep method of slide preparation results in samples free from biological debris and with clear background, there is no extra effort needed to remove such debris. Then Bit plane slicing segmentation method is applied to preprocessed image which segments cellular region in the image. The segmented cellular region contains both disjoint cells and overlapped cells forming cluster. As this paper concentrates on finding number of cells in the overlapped cells region, disjoint cells are ignored for the study. After retaining only overlapped cells region or cluster, that region is cropped using bounding box based on area of segmented region with area > 5.66 pixels (1500 μm^2). As the maximum area of squamous cervical cell, i.e., area of superficial and intermediate cells is 1500 μm^2 [1].

Each cropped area has overlapped cell clusters in a given sample. Later in phase 2 segmentation these overlapped cells are separated again using bit plane slicing.

Retrieval Number: C4572098319/2019©BEIESP
DOI:10.35940/ijrte.C4572.118419
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

A-1 Algorithm to count cells in the given Pap smear sample

Step 1: Input ThinPrep Pap smear image
Step 2: Preprocessing of input image using Adaptive Wiener filter with window size [7 7].
Step 3: Increase the contrast of image by saturating bottom 1% and top 1 % of all pixel values.
Step 4: obtain 8 bit planes of the image. This is done by storing kth bit from each pixel in an image. Here, k=7.
Step 5: The binary image with kth bits is obtained. Apply holes filling and border removal.
Step 6: Find and mark the edges in 7^{th} Bits binary image using Laplacian of Gaussian filter. Label each region bounded by edges. These regions are corresponding to cytoplasmic area.
Step 7: The regions obtained in step 6 are corresponding to both disjoint cells and overlapped cells.
Step 8: Consider a disjoint cell based on the area, identify and label its nuclei region using thresholding. The number of labels for nuclei area will provide the count of disjoint cells.
Step 9: Consider a overlapped cell cluster, identify, label and count nuclei region as in step 7.
Step 10: Estimate the cell count by summing disjoint cell count and overlapped cell count.

Bit plane slicing was used twice as it was giving more accuracy in a less time compared to other segmentation methods authors have tried. After separating overlapped cells in a cropped area, cells are counted. Then total cells in a given image is given by equation-

Total_Cell_Count= Number_of_disjoint_cells +
∑Number_of_cells_in_each_cropped_region…. (1)

The number of disjoint cells can be found by counting the number of regions ignored in phase 1 segmentation and also whose area is between 1.13 to 5.66 pixels (300 μm^2 -1500 $\mu m^{2)}$ as size of different squamous cells vary between these values [1].

IV. RESULT AND DISCUSSION

This work concentrated on counting number of cells in overlapped ThinPrep Pap smear microscopic cervical cells. The following images shows step-by-step working of proposed method to count number of cells in given ThinPrep Pap smear cell microscopic image. The following images shows sample output that is cell count in a cropped overlapping cluster of fig. 2-

The fig. 3 to fig. 6 shows cropped overlapped cell cluster and its corresponding cell count. The proposed method were tested over 500 overlapped cells clusters with number of cells are atleast two in each cluster. Due to false positives which occurred due to too much overlapping caused variations with the actual cell count and cell count calculated by the proposed method. But in most of the cases it was successful in finding cell count accurately.

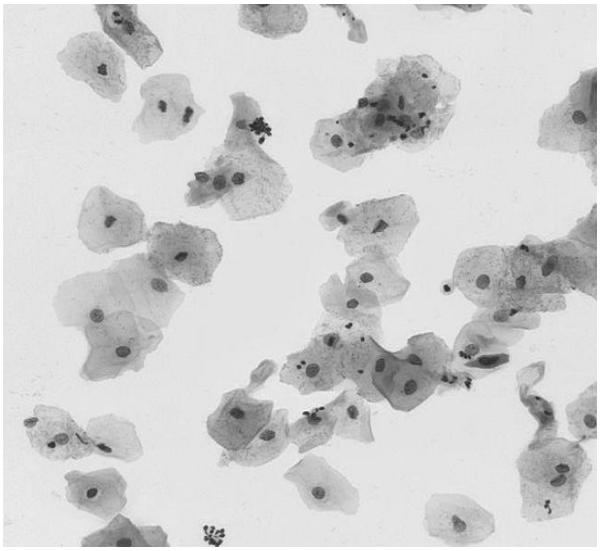

Fig. 2. Original ThinPrep Pap smear image with Overlapped cells

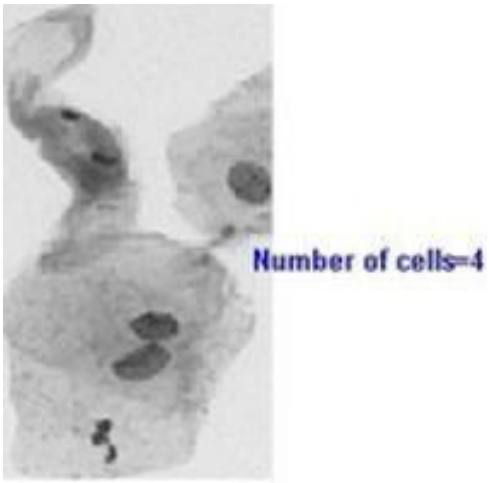

Fig. 3. Overlapped ThinPrep Pap Smear cell Cluster with cell count=4

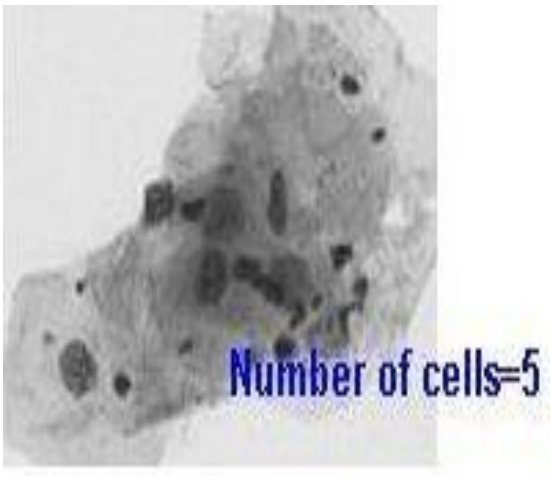

Fig. 4. Overlapped ThinPrep Pap Smear cell cluster with cell count=5

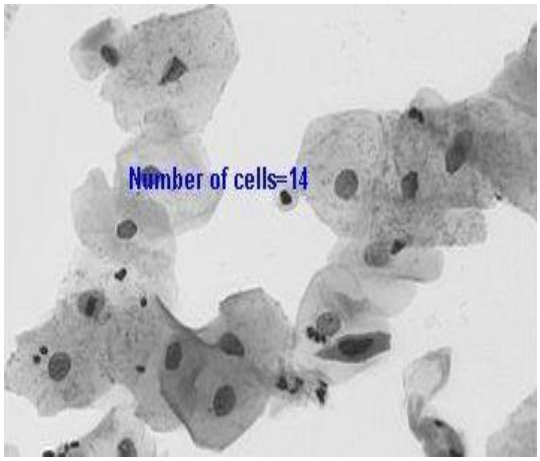

Fig. 5. Overlapped ThinPrep Pap smear Cell cluster with cell count=14

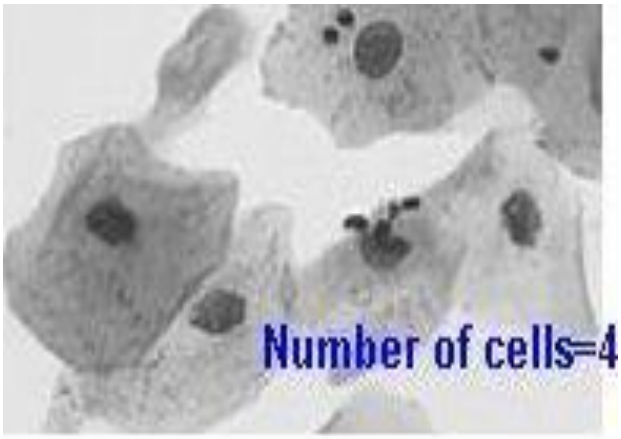

Fig. 6. Overlapped ThinPrep Pap smear Cell cluster with cell count=4

V. CONCLUSION

This paper proposes an automated method for estimating cell count in the given whole cervical cell Pap smear image. This method offers a significant performance even when cells are overlapped. The proposed method uses two stage segmentation with Bitplane slicing technique, where phase 1 separated disjoint cells and overlapped cell cluster from the background. Phase 2 segmentation separated each overlapped cells in a particular cluster. The total number of cells in a given sample is estimated as sum total number of disjoint cells and total number of cells in each overlapped cluster. This technique can be directly applied in any ThinPrep cervical Pap smear image. The proposed method estimated cell count accurately when the number of overlapped cells in a cluster is less than 5 and when cells are slightly overlapped.

The cytoplasm segmentation can be further improved by considering concavity points, which is the point of intersection at the overlapping region. This can be further improved by considering morphological features of cell like area, pixel intensity to accurately identify cell region.

The proposed method for estimating cell count can be combined with any other classification will lead to computer-based automated primary screening system for cervical cancer.

REFERENCES

1. Rana S. Hoda, M.D, FIAC and Syed A. Hoda M.D, “The Pap smear ~ Current criteria and changing concepts”, Medical University of South Carolina, Charleston, SC and New York Presbyterian Hospital-weill Medical college of Cornell University, New York, NY, Page 4.
2. Gupta, N., Bhar, V. S., Rajwanshi, A., & Suri, V. “Unsatisfactory rate in liquid-based cervical samples as compared to conventional smears: A study from tertiary care hospital”, *CytoJournal, Vol.* 13(14), 2016 http://doi.org/10.4103/1742-6413.183831
3. Happy S, Chatterjee S, Sheet D “Unsupervised Segmentation of Overlapping Cervical Cell Cytoplasm”, *arXiv preprint* , 2015, arXiv:1505.05601, Google Scholar
4. J. Su, X. Xu, Y. He, J. Song, "Automatic Detection of Cervical Cancer Cells by a Two-Level Cascade Classification System", *Anal. Cell. Pathol. (Amst).*, vol. 2016, pp. 9535027, 2016.
5. M. E. Plissiti, C. Nikou and A. Charchanti, "Automated Detection of Cell Nuclei in Pap Smear Images Using Morphological Reconstruction and Clustering," in *IEEE Transactions on Information Technology in Biomedicine*, vol. 15, no. 2, pp. 233-241, March 2011. doi: 10.1109/TITB.2010.2087030
6. Y. Jusman, S.C. Ng, N.A.A. Osman, "Intelligent screening systems for cervical cancer", *The Scientific World Journal*, 2014.
7. P. Sukumar and Dr. R. K. Gnanamurthy, “Computer Aided Detection of Cervical Cancer Using Pap smear Images Based on Hybrid Classifier”, International Journal of Applied Engineering Research ISSN 0973-4562 Volume 10, Number 8 (2015) pp. 21021-21032.

AUTHORS PROFILE

Ms. Deepa. T. P, working as Assistant Professor at Jain University. She is pursuing Ph.D. in Medical Image Processing from VIT University, Vellore. She is doing research on primary screening system for cervical cancer. She did B.E. in Computer Science and Engineering from VTU in 2003 and M. Tech in Computer Science and Engineering from VTU in 2006. She has more than 11 years of teaching experience in various reputed engineering colleges and university. She has guided various industry, abroad universities collaborative projects. She has published papers in various peer-reviewed impact factor and Scopus indexed journals. She also presented papers in various state, national and international level conferences. She was a resource person in various workshops, short-term courses, campus connect programs like Infosys, Wipro, CISCO certification programs and teacher’s training workshop. She was part of many industry-academia connect program and MOU for the abroad university. She is a reviewer for the open access peer-reviewed journal and national conference. Her main areas of interests are Digital Image Processing, Brain Signal Processing, Electronic Waste Management and Surveillance, Wireless Sensor Network.

Dr A. Nagaraja Rao has Completed M.Sc in Computer Science from SV University, M.Tech. in Computer Science and Engineering from JNTU Hyderabad and Ph.D. in Computer Science from University of Mysore. He has more than 15 years of teaching and research experience. He is a life member of CSI, ISTE professional bodies. At present he is working with VIT, Vellore as an Associate Professor in School of Computer Science and Engineering (SCOPE). He has published more than 40 papers in international journals and Conferences. His areas of interest include Image Processing, Medical Image processing, Data Analytics and Pattern Recognition.