

Counting cells in microscopic image using two-phase segmentation

DEEPA T P¹, JHANHAVI², SESA RAMYA³, PAVITHRA⁴, SANTHOSH KUMAR⁵

Department of Computer Science and Engineering

Acharya Institute of Technology

Bangalore, Karnataka

INDIA

deepatp@acharya.ac.in

Abstract: - Cytopathology is a branch which deals with study diseases at the microscopic cell level. Now a day, deadly disease like cancer has become a major challenge in the field of medical research. If these diseases are identified in the early stages, then they are curable. There are many methods to ease, faster and produce accurate results of this analysis. One such approach is using image processing techniques. This paper applies segmentation techniques to find important feature like cell count. This feature is further used to differentiate between normal and abnormal cells. Abnormal cells are the one which may turn into cancerous cells. Hence the early detection of such abnormal cells before they turn into cancerous cells helps to diagnose patient early. The proposed method was evaluated using cervical cells.

Key-words: - Cytopathology, Edge detection, Filter, Nuclei, Cervical cell, Cell counting, Image Segmentation

1 Introduction

Early detection of diseases like cancer helps in early diagnosis of patients. Cell samples are collected from the patient using laboratory procedure. Pathologists manually analyze these cells by placing samples under microscope. This is tiresome, time consuming, variations in results and less number of samples examined per day and costly. Hence there is a need for a method which overcome all these problems. Image processing is one such technique where cell analysis is done using techniques such as segmentation and classification. Application of these techniques is challenging because of difficulty due to variety of microscope, different staining methods, cell types, and densities in a given sample, complexity of data especially if the sample is corrupted. Hence, this paper attempts a method using image processing techniques which can overcome fewer of above mentioned problem to some extent. The conclusion drawn from this work helps to find one of the feature to know whether a given sample contain abnormal cells or not. This will further help to diagnose deadly diseases like cancer, where abnormal cells later turn into cancerous cells. Suitable image processing techniques are used in this paper to detect normal, malign, benign cells in the given microscopic image. This decision further helps pathologists to early detect chance of deadly disease.

Images used in this work are acquired using Light microscopy. For which patient should be prepared for collecting samples which are done in the

laboratory at hospitals. Cells are smeared using spatula from the patient organ (where malignancy is suspected). These cells are placed on glass slides, suitable stains are applied for clear visibility of cells. These glass slides are placed under a microscope and viewed manually by pathologists for drawing suitable conclusions. But due to human errors, fatigue, a difference in the way of analyzing, this becomes tedious and error prone process. Also, the number of samples analyzed will be limited per day. In most countries. The cost of these screening tests is high. These drawbacks can be overcome by using automated software. This paper approaches to build a methodology to design one such software.

2 Existing Work

In existing work, Pradipta Maji et al (2015), proposed an automated method for counting red blood cells present in the blood sample. This method addresses the problem of holes present in between blood cells and also overlapping characteristics of RBC's. In this method, thresholding is done using Otsu's method, then hole filling process is applied. Each blood cell is extracted and its shape is analyzed to find its circularity, overlapping, etc. Partial cells are discarded. Number of cells in overlapped area are found. This methodology can be used to count other types of blood cells.

Ji Ge et al (2014) proposed a system for automated counting and identifying fetal and

maternal RBC's on slides obtained from Kleihauer–Betke (KB) test. This test is a standard test used to quantify fetal-maternal hemorrhage in maternal care. Overlapped cells are separated using spatial color pixel classification. For accurate results, whole slide is used for supervised learning. The method was able to count over 60000 slides in 5 mins. Compared to manual testing, this method was found significantly efficient and accurate.

Marina E. Plissiti and Christophoros Nikou (2011) presented an efficient framework for training an active shape model which represents smooth shapes. The boundaries of the region of interest (ROI) are detected with the convergence of physics based deformable model. The attributes of shapes of ROI are expressed in terms of modal analysis. Proposed method is evaluated using pap images.

M. Park et al (2010) proposed a method for automated counting of microscopic image in which canny edge detector extracts the contour points using 8-neighbours connectivity. This form contour segments. The ellipse fitting algorithm is applied to each segment to fit an ellipse to the points. These contour segments are classified into 3 categories: not an ellipse, single cell and clustered cells. Not ellipse objects are ignored. Single cells are counted. Clustered cells are further segmented to identify single cells.

3 Methodology

Based on features of cells and nucleus malignancy of cells are detected by extracting features like, roundness of nuclei, area of nuclei, distribution of nuclei are in cytoplasm (nuclei to cytoplasm ratio), diameter of nuclei and cell, the number of cells in the given sample. The proposal in this paper describes for counting number of cells in the given sample which helps to find the percentage of abnormality in that sample.

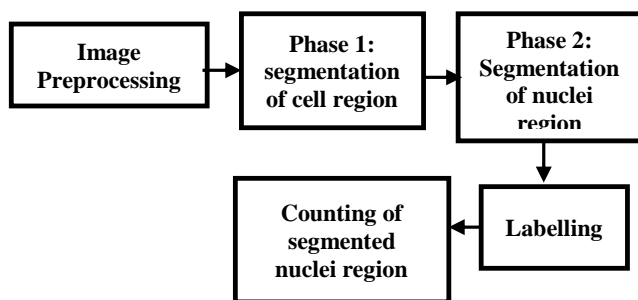


Fig. 1: Proposed Methodology

The image processing techniques like segmentation is used to obtain nucleus region, which is considered

as a region of interest (ROI) in this work as shown in Fig. 1.

$$\text{Percentage of abnormality in given sample} = \frac{\text{Total number of abnormal cells in given a sample}}{\text{Total number of cells in a given sample.}} \quad (1)$$

This equation (1) shows, it's necessary to count the actual number of cells in a given image. For this, a suitable and efficient segmentation method with pre-processing which can clearly segment nuclei area is required. This work is an approach to find one such method which is suitable for microscopic image.

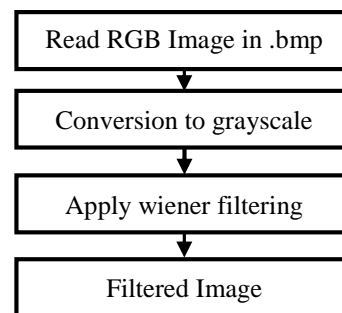


Fig. 2: Pre-processing of image

3.1 Image acquisition

Images are acquired under light microscopic by placing a glass slide sample on it. They are acquired with 40X magnification, with dimension 760 X 560pixels and stored in .bmp format. 40X magnification helps to analyze nuclei features from the given sample images. The cervical cell image was considered for testing.

3.2 Pre-processing

There exists some unwanted information like blood cells, yeast, stains (over-stained), bacteria, etc. In the given cell sample, and may also contain some noises during image acquisition. These are removed using filters and morphological operations like opening and closing. Wiener filtering algorithm as shown in equation (2) [1] with window size 25X25 is used for pre-processing. This results in image with removal of noise induced during acquisition and small objects.

$$w(u, v) = \frac{H^*(u, v)}{|H(u, v)|^2 + \frac{P_n(u, v)}{P_s(u, v)}} \quad (2)$$

Where,

$H(u, v)$ = Degradation function

$H^*(u, v)$ = Complex conjugate of degradation function

$P_n(u, v)$ = Power Spectral Density of Noise

$P_s(u, v)$ = Power Spectral Density of un-degraded image

3.3 Two-phase Segmentation

The cell separated from unwanted background information is segmented to extract regions of cells in turn region of the nucleus. This results with the image in which nuclei and cell boundaries are well identified and distinguished. This is achieved through edge detection. Each segmented area of nuclei are labelled. Number of labels specify number nucleus in turn number of cells. Based on nuclei boundary, number of cells are counted but not cell boundaries. If the cells are overlapped, the entire cell boundary is counted as one segment. The following methods are used to segment cell and nucleus region separately.

3.3.1 Phase 1: Segmentation of cells

There are two ROI in this method. One is cell region and other is nuclei region. Phase 1 deals with segmentation of cell region. Later, thresholding is applied to filtered image which converted grayscale into binary and also separate foreground from background. Then the image is complemented which results in dark areas become light and vice versa. As a result edges of cells are well distinguished from the background. Boundary tracing involves tracing exterior boundaries of the region of interest, where outermost objects are considered as parent object and traces towards its children which are innermost objects. Later, inside objects in a cell are closed using hole filling technique. Image is plotted with boundary pixels in red color. Hence first region of interest that is cells are marked.

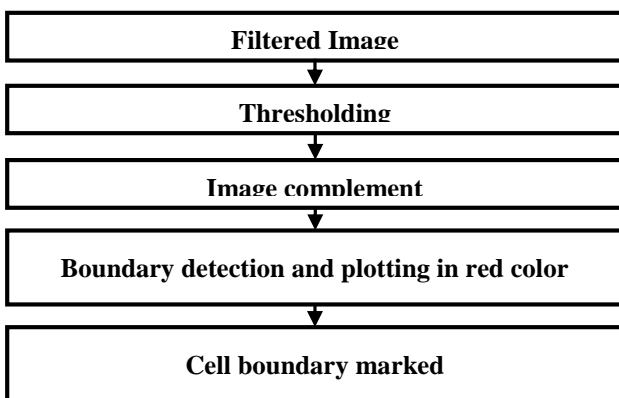


Fig. 3: Phase 1: Segmentation of Cell Region

3.3.2 Phase 2: Segmentation of nuclei

The nuclei is separated from its cytoplasmic background, i.e, segmentation of nuclei. The color space conversion helps in this case to separate the nuclei from the background. Naturally the brighter area is the cytoplasm (that has high intensity) and the darker region is the nuclei (that has low intensity). High intensity objects should be separated from the low intensity object. The high intensity is

removed once the images are converted from one color space to another, i.e. cytoplasmic area is removed. Edge detection techniques are used to segment the cells, change in the intensity of the image is represented by areas of edges.

These edges are detected using edge detection operator like Laplacian of Gaussian edge detection. The Laplacian of an image is often used for edge detection as it will highlight the regions that has rapid intensity change such as zero crossing edge detection. It is often applied to images that is smoothed with some approximation such as Gaussian smoothing filter as shown is equation (3) is used to reduce its sensitivity to noise, and hence it describes here the two different variants together. Here input is taken as a single grey level image and the output produced will be another grey level image. [3] The standard formula for Laplacian is Lap (x, y) of an image along with the pixel values Image (x, y) is given by:

$$\text{Lap}(x, y) = -\frac{1}{\pi\sigma^4} \left[1 - \frac{x^2 + y^2}{2\sigma^2} \right] e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (3)$$

3.4 Boundary marking and Labelling.

A cell boundary color is marked as red and green color is marked to nuclei through Matlab functions using be boundaries. The segmented nuclei is further should be labelled the Matlab function bwlabel that will in turn returns the number of components and labels created. Number of labels created is mapped to a number of nuclei which in turn gives counts of cells.

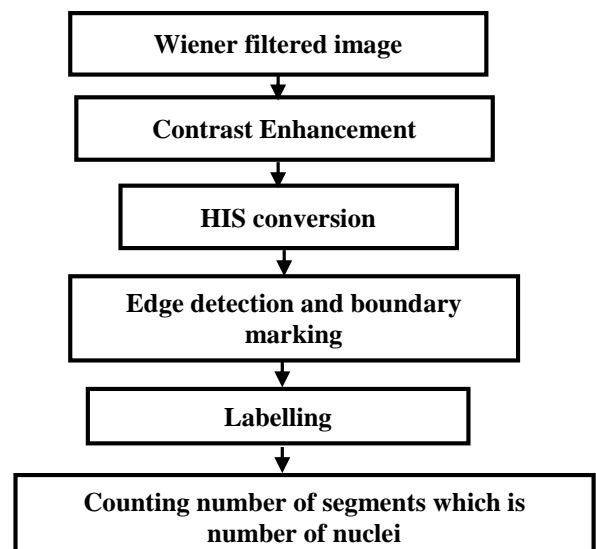


Fig. 4: Phase 2: Segmentation of Nuclei region








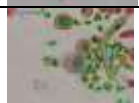
For labelling, matlab function bwlabel was used, [Label, Label_count] = bwlabel (image, conn) - (4)

This returns a matrix called Label which contains labels for connecting objects in the image. This matrix is the same size as that of the image. Churn refers to the number of connected components. In this work 8-connected components were identified. Elements in matrix Label are greater than or equal to zero. Background pixels are labeled with the value zero. For each connected foreground pixels are labelled with 1, 2... for the number of connected components. For example, single nuclei is considered as connected component or connected object. Firstly identified nuclei will be labelled 1, secondly identified will be labelled with 2 and so on. Number variable in the LHS of equation specifies the number of labels created. This number of labels which are created in turn depends on number of nuclei present, which is nothing but cell count. The Label elements will be greater than or equal to 0. The background pixels are labeled with zero. The foreground pixels in first object are labeled 1, the pixels in second object are labelled 2 and so on. Label_count in the LHS of the equation (4) returns the number of labels created. For each object that is nuclei in this image, labels are created. So, label_count specifies the number of nuclei.

4 Results and Observations

Results of the proposed methods are as listed in the table 1. 1 which shows segmentation output. Over 100 samples of different cell counts were tested. View results are as recorded below. Images are grouped into Type 1, 2, 3 based on joining and overlapping of cells.




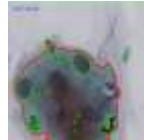
4.1 Type 1: When cells and Nuclei are disjoint

Input Image	Segmented Image	Remarks
		Nuclei and cells are well segmented and their boundary is well defined.
		
		
		

[Image Courtesy of Jens Byriel, Technical University of Denmark, Department of Automation, Lyngby, Denmark]

Table 1: Microscopic image segmented using Two-phase segmentation



4.2 Type 2: When cells are overlapped and nuclei are distinct


Segmented Image	C _{Proposed}	Remarks
	26	In this image, cells are overlapped, but nuclei are distinct, since counting is based on nuclei region, count gives the correct value
	43	In this image, cells forms cluster, nuclei regions are marked and counted
	25	In this image, cells are joined and overlapped, in this case, due to varying intensity at overlapped regions even nuclei regions are not segmented properly
	24	Here cells forms clustered and tightly coupled with each other. Hence, one single boundary is marked for entire cluster and the nuclei are both distinct and overlapped. Only distinct nuclei are detected and marked. This can be considered as future work.

[Image Courtesy of Jens Byriel, Technical University of Denmark, Department of Automation, Lyngby, Denmark]

Table 2: Segmentation results of image with overlapped cells and distinct nuclei

4.3 Type 3: When Nuclei are joined or overlapped

Segmented Image	C _{Proposed}	Remarks
	29	Here, cells are overlapped, but nuclei boundaries are joined (not overlapped). Two joined nuclei for example, is segmented as one region, hence counting as one nucleus instead of two.
	36	Here the cells are overlapped, nuclei are joined and few nuclei are overlapped. Overlapped nuclei are segmented counted as one nucleus.

	25	False detection due to percentage of joining and overlapping is more. Few such nuclei are not segmented at all.
---	----	---

[Image Courtesy of Jens Byriel, Technical University of Denmark, Department of Automation, Lyngby, Denmark]

Table 3: Segmentation results of image with overlapped or joint nuclei

Samples	C _{manu} _{al}	C _{Proposed}	TP	FP	TN	FN
Sample 1	8	12	8	0	0	4
Sample 2	9	22	9	0	2	13
Sample 3	7	11	7	0	1	4
Sample 4	4	8	4	0	1	4
Sample 5	1	7	1	0	0	6
Sample 6	2	3	2	0	3	1
Sample 7	10	24	10	1	4	14
Sample 8	4	39	4	3	6	35
Sample 9	4	25	4	4	2	21
Sample 10	1	25	1	0	2	24
Sample 11	2	3	2	0	0	1
Sample 12	9	11	9	3	4	2
Sample 13	4	9	4	2	0	5
Sample 14	3	10	3	0	0	7
Σ n=14	68	209	68	13	25	141

Table 4: Statistical values for counting number of cells for number of samples =14

Above Table 4 gives statistical results on a number of cells by proposed method and manual method. Following terminologies used-

1. TP-True positive: Nuclei region correctly identified as nuclei
2. FP-False Positive: Nuclei region, but identified as non-nuclei
3. TN-True Negative: non nuclei correctly identified as non-nuclei
4. FN-False Negative: Non nuclei, but identified as nuclei
5. C_{manual} –Number of cells counted manually
6. Proposed –Number of cells from proposed method

The measure of correctly identified and actual positive regions is called as Sensitivity as given in equation (5). In this method, actual positive region is nuclei region. Hence sensitivity here means the proportion of the actual nucleus region correctly identified.

Sensitivity or True Positive Rate (TPR) [22] =

$$\frac{\text{Nuclei region correctly identified as nuclei}}{(\text{Nuclei region correctly identified as nuclei} + \text{Non nuclei but identified as nuclei})} \tag{5}$$

The measure of correctly identified negatives which are actually present is called as Specificity as given in equation (6). In this image, it may be non-nuclei region or nuclei which are partly out of bounds of giving image. If the number of such region exists and not counted four cells, then it is the specificity of the proposed method.

Specificity or True Negative Rate [22] =

$$\frac{\text{Non nuclei correctly identified as non-nuclei} + \text{Nuclei region but identified as non-nuclei}}{\text{Non nuclei correctly identified as non-nuclei} + \text{Nuclei region but identified as non-nuclei}} \tag{6}$$

Therefore, for 14 samples, Sensitivity=30% and specificity is 65.7%. in 14 samples, measure of correctly identifying nucleus region is 30% and measure of correctly identifying non-nuclei region is 65.7%. Rest of 4.3% regions are ignored due to nuclei touching boundary or false objects. Similarly, total 100 images were tested and results were recorded. Out of which there were 60 images of type 1, 23 images of type 2, 17 images of type 3 were tested. The actual number of nuclei found by the proposed method,

$$C_{\text{nuclei_proposed}} = C_{\text{proposed}} - \text{FN}. \tag{7}$$

Difference between count using manual and proposed method,i.e,

$$\text{Error} = C_{\text{manual}} - \text{FN} \tag{8}$$

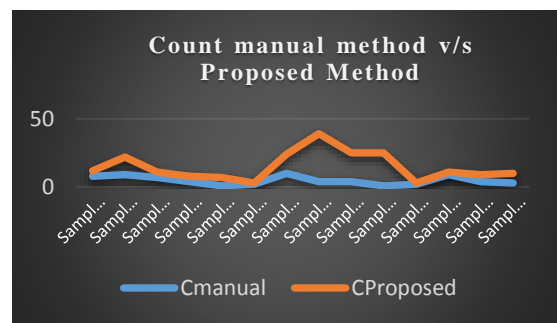


Fig. 5: Comparison of count of nuclei obtained from manual and proposed method

Above Fig. 5 shows graph of manual count versus count by proposed method for each sample. In few samples difference between count from manual and

proposed method is too high. This is due to false negatives.

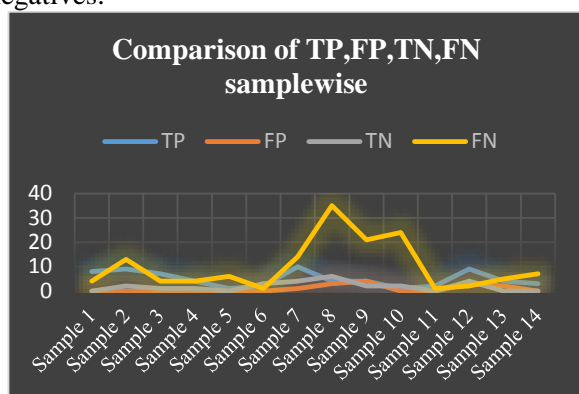


Fig. 6: comparison of TP, FP, TN and FN sample-wise

Above Fig. 6 shows graph of comparison between TP, FP, TN and FN for each sample. It shows that proposed method in the few samples identified non-nuclei as nuclei object. Whereas this can be improved by removing non-nuclei regions using more efficient filters and using feature extraction. This will be considered in next work which is a continuation of this work.

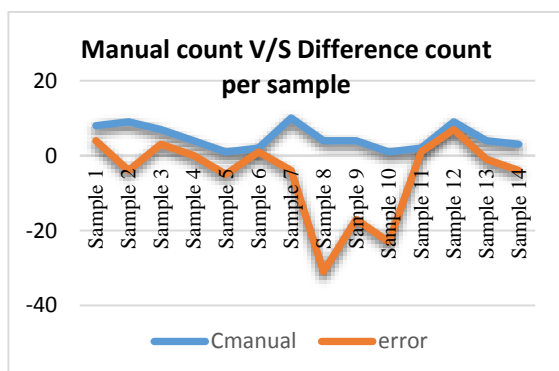


Fig. 7: Graph of count manual v/s difference count per sample

In fig 7 is a graph of manual count and difference between manual count and proposed count (error) occurred from proposed method per sample, absence of error bar specifies, there is no error. The positive error value specifies over-counting by the proposed method. The negative error value specifies undercutting by the proposed method.

5 Conclusion

In the image, nuclei boundary is successfully marked by green color and cell boundary by red color. Those cells touching boundary are ignored because they are partly visible and only part of nuclei are visible. Each nucleus region is successfully labelled. And number of labels represented the

number of nuclei in turn number of cells. Compared to manual counting, this proposed method found to 98% correct, with a difference of one or two cells. But this algorithm failed to count two joined and overlapped nuclei as one nucleus region, as a result, there was high false negative value for some sample. For example, if there are two joint nuclei, it was counted as one. These issues need to be addressed further and considered as a future enhancement of this proposed method.

Acknowledgement

Authors are very much thankful to Chairman, Acharya Institutes. Dr.H.D.Maheshappa, Principal, A.I.T,Bangalore. Dr.K.G.Mohan, HOD, CSE, A.I.T, Dean-R&D, Acharya IT, Bangalore. Dr.Nagaraja Rao.A., Associate professor, VIT University, Vellore. Jens Byriel, Technical University of Denmark, Dept. of Automation, Denmark for providing access to download required dataset. Dr.T.G.Basavaraju, Professor and Head Department of Computer Science Engineering Government SKSJ Technological Institute, Bangalore. Prof.Chandrakala, Assistant Professor, Dept. of ISE, Acharya IT, Bangalore. Late. Prof. Nagaraj, Sr. Faculty; Sr.Consultant at CSR - Centre of Excellence Mysore, Mr.Chandrashekhara Gokhale.V., Deputy Manager, Special Equipments, BIAL, Bangalore and all our colleagues, family and friends who have directly or indirectly supported this work.

References:

- [1] Suresh Kumar¹, Papendra Kumar², Manoj Gupta³, Ashok Kumar Nagawat⁴, Performance Comparison Of Median And Wiener Filter In Image De-Noising, International Journal Of Computer Applications (0975 – 8887),Volume 12– No.4, November 2010
- [2] Pinaki Pratim Acharjya, Ritaban Das & Dibyendu Ghoshal, Study And Comparison Of Different Edge Detectors For Image Segmentation, Global Journal Of Computer Science And Technology Graphics & Vision Volume 12 Issue 13 Version 1.0 Year 2012
- [3] Raman Maini & Dr. Himanshu Aggarwal, Study And Comparison Of Various Image Edge Detection Techniques, International Journal Of Image Processing (Ijip), Volume (3):Issue 1
- [4] Pradipta Maji, Ankita Mandal ,Madhura Ganguly, Sanjoy Saha, An Automated Method

- For Counting And Characterizing Red Blood Cells Using Mathematical Morphology, 978-1-4799-7458-0/15/\$31.00 _C 2015 Ieee
- [5] Ji Ge, *Member, Ieee*, Zheng Gong, *Member, Ieee*, Jun Chen, Jun Liu, *Student Member, Ieee*, John Nguyen, Zongyi Yang, Chen Wang, And Yu Sun*, *Senior Member, Ieee*, *A System For Counting Fetal And Maternal Red Blood Cells*, Ieee Transactions On Biomedical Engineering, Vol. 61, No. 12, December 2014
- [6] Jasna Lojk1, Luka Šajn2, Uroš Čibej2, Mojca Pavlin1, Automatic Cell Counter For Cell Viability Estimation, Mipro 2014, 26-30 May 2014, Opatija, Croatia
- [7] J.Hari, A.Sai Prasad, S.Koteswara Rao, Separation And Counting Of Blood Cells Using Geometrical Features And Distance Transformed Watershed, 2nd International Conference On Devices, Circuits And Systems (Icdcs),, 978-1-4799-1356-5/14/ ©2014 Ieee2014
- [8] Watcharin Tangsuksant□, Chuchart Pintavirooj, Somchart Taertulakarn□, Somsri Daochai, Development Algorithm To Count Blood Cells In Urine Sediment Using Ann And Hough Transform, The 2013 Biomedical Engineering International Conference (Bmeicon-2013), 2013
- [9] Siti Madihah Mazalan, Nasrul Humaimi Mahmood, Mohd Azhar Abdul Razak. Automated Red Blood Cells Counting In Peripheral Blood Smear Image Using Circular Hough Transform, First International Conference On Artificial Intelligence, Modelling & Simulation, 2013
- [10] Mostafa Mohamed, Behrouz Far, An Enhanced Threshold Based Technique For White Blood Cells Nuclei Automatic Segmentation, Ieee 14th International Conference On E-Health Networking, Applications And Services (Healthcom), 2012
- [11] Marina E. Plissiti, Christophoros Nikou, Cell Nuclei Segmentation By Learning A Physically Based Deformable Model, 978-1-4577-0274-7/11/ ©2011 Ieee
- [12] Weiyan Hou, Zhengwei Duan, Xiaodan Liu, A Template-Covering Based Algorithm To Count The Bundled Steel Bars, 4th International Congress On Image And Signal Processing, 2011
- [13] M. Park, *Member, Ieee*, J.S. Jin, Y. Peng, P. Summons, D. Yu, Y. Cui, S. Luo, F. Wang, L. Santos And M. Xu, Automatic Cell Segmentation In Microscopic Color Images Using Ellipse Fitting And Watershed, The 2010 Ieeeicme International Conference On Complex Medical Engineering July 1 3-1 5, 2010, Australia
- [14] Dina Thomas #1, Rajesh Bordawekar #2, Charu C. Aggarwal #2, Philip S. Yu, On Efficient Query Processing Of Stream Counts On The Cell Processor, 1084-4627/09 © 2009 Ieee, Doi 10.1109/Icde.2009.35
- [15] Di Fan, Maoyong Cao, Changzhi Lv, Wall-Adherent Cells Segmentation Based On Cross-Entropy And Watershed Transform, 978-1-4244-1724-7/08 ©2008 Ieee
- [16] Kehong Yuan, Jiying Zou, Shu Feng, Chaijie Duan, Shanglian Bao, Qiansheng Cheng, Roundness Curve For Classification Of Cell Phase On Microscopic Image, Proceedings Of The 5th International Conference On Information Technology And Application In Biomedicine, In Conjunction With The 2nd International Symposium & Summer School On Biomedical And Health Engineering Shenzhen, China, May 30-31, 2008
- [17] Vinh Thong Ta, Olivier Lezoray, Abderrahim Elmoataz, Graph Based Semi And Unsupervised Classification And Segmentation Of Microscopic Images, Ieee International Symposium On Signal Processing And Information Technology, 2007
- [18] Fatemeh Zamani Reza Safabakhsh, An Unsupervised Gvf Snake Approach For White Blood Cell Segmentation Based On Nucleus, 0-7803-9737-1/06 ©2006 Ieee
- [19] #P.S. Hiremath And Humnabad Iranna Y., Automated Cell Nuclei Segmentation And Classification Of Squamous Cell Carcinoma From Microscopic Images Of Esophagus Tissue, 1-4244-0716-8/06 ©2006 Ieee.
- [20] E. Montseny P. Sobrevilla S. Romani, A Fuzzy Approach To White Blood Cells Segmentation In Color Bone Marrow Images, 0-7803-8353-2/04 2004 Ieee 173
- [21] Hazem Refai, Ph.D., Lirn Li, Electrical And Computer Engineering Department, University Of Oklahoma, Tulsa T. Kent Teague, Ph.D., R. Naukom, Department Of Surgery, University Of Oklahoma College Of Medicine, Tulsa, Automatic Count Of Hepatocytes In Microscopic Images, 0-7803-7750-8/03 2003 Ieee
- [22] Chang Sm, Matchar Db, Smetana Gw, Et Al., Editors. Methods Guide For Medical Test Reviews [Internet]. Rockville (Md): Agency For Healthcare Research And Quality (Us); 2012 Jun.