International Journal of Applied Information Systems (IJAIS) – ISSN : 2249-0868
Foundation of Computer Science FCS, New York, USA
Volume 10 – No.1, November 2015 – www.ijais.org

Leukemia Detection using Digital Image Processing Techniques

Himali P. Vaghela
M.Tech Student
Dept. of Electronics and Communication
Charusat University, Changa

Hardik Modi
(Member IEEE)
Faculty of Electronics and Communication
Charusat University, Changa

Manoj Pandya
PhD
Project Manager
BISAG Gandhinagar

M.B. Potdar
PhD
Project Director BISAG
Gandhinagar

ABSTRACT
This paper discusses about methods for detection of leukemia. Various image processing techniques are used for identification of red blood cell and immature white cells. Different disease like anemia, leukemia, malaria, deficiency of vitamin B12, etc. can be diagnosed accordingly. Objective is to detect the leukemia affected cells and count it. According to detection of immature blast cells, leukemia can be identified and also define that either it is chronic or acute. To detect immature cells, number of methods are used like histogram equalization, linear contrast stretching, some morphological techniques like area opening, area closing, erosion, dilation. Watershed transform, K means, histogram equalization & linear contrast stretching, and shape based features are accurate 72.2%, 72%, 73.7 % and 97.8% respectively.

Keywords
Blood disease detection, leukemia detection, k means clustering, watershed transform, histogram equalizing, and shape based features, count number of red and white cells

1. INTRODUCTION
There are different types of white cells in our body. Leukemia is nothing but cancer of blood cells in which number of white cells is increasing and those are immature cells that destroy other cells. Today laboratory test takes longer interval of time to diagnose the disease of leukemia and it is also time consuming, prone to human error and also tedious. The ratio of white blood cell in our body is 1000:1. It means that 1 white blood cell is present between 1000 red cell. So if number of white blood cells increase remarkably in large number then the person is succumbed to suffer from the leukemia. It further falls into two type: acute and chronic. If number of white cell is increasing in our body then this immature cells start destroying another cells of our body.

Here task is to detect immature cell using different image processing techniques and count total number of cells. So we need to use the technology that identifies different types of blood cells within short duration of time in emergency. Furthermore it is vital to study in detail how to differentiate different cell and recognize it as immature cell and according to it, detect the leukemia. Acute and chronic also have two types. 1) Lymphocytic and 2) myeloblastic that both are due to immature blast of lymphoid and myeloid cell respectively.

The task can be improved and performed in near-real time environment using biomedical image processing. Here formation of lymphoid and myeloid cells of series is shown in figure 1. Figure 1 has been taken from web sources [22].

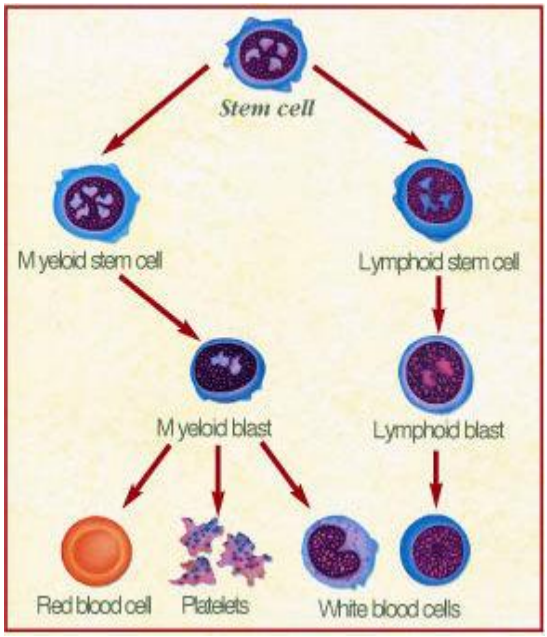


Figure 1. The Formation of Myeloid and Lymphoid Series of Cell

2. RATIONALE
Various methods have been applied to automate the task to find out leukemia cells and count it.

2.1 Watershed transform
Lim Huey Nee et al. proposed methods for segmentation of white cells based on morphological operation, gradient magnitude and watershed transform. First image acquisition techniques is used then segmentation is done to separate the blast cell and back ground. For this method, first the RGB image is converted into HSV color model and saturated component is extracted for further processing and then find the gradient magnitude for the saturation component. This is used for edge detection. Moreover, sobel, canny, prewitt operators are used for the edge detection. After extracting the white cells from the image and elimination of the background and red cells, dilation or erosion process is carried out. Then watershed transform is carried out to separate the connected cell. Thus, leukemic cell can be identified and this method gives very accurate result [2]. But the exact separation of cells cannot be done using this method. The process flow is shown in figure 2. Figure 2(a) has been taken from web sources [23].

2.2 K Means Clustering Technique

To identify the abnormalities in blood cells or to identify the lymphoblast, Mashiat Fatma and Jaya Sharma proposed the method of clustering techniques. After pre-processing of image feature extraction is done that gives useful information about image. The pre-processing techniques are only used for image enhancement. It does not give any necessary information of image. So initially acquisition process is done and then contrast enhancement is necessary to see clear image. After that RGB image is converted to HSI color model and then K-means clustering technique is applied for segmentation. Median filter is used to remove the noise from image. After feature extraction, image is classified by clustering techniques. Two types of methods can also be used.1) supervised learning and 2) unsupervised learning. Feature extraction is used to identify the white cells or the lymphoblast from image. So some features are used [3].

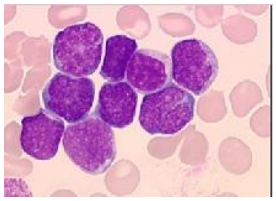


Figure 2(a) original image

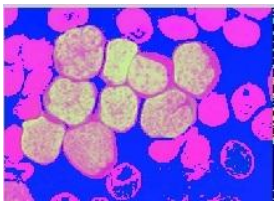


Figure 2(b) rgb2 hsv color model

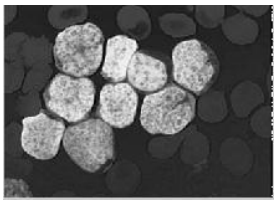


Figure 2(c) saturation component extracted

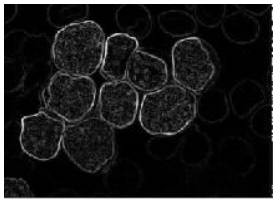


Figure 2(d) derivation of gradient magnitude

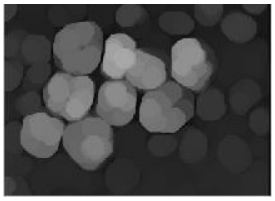


Figure 2(e) area opening: morphological operation

Figure 2(f) opening-closing :morphological operation

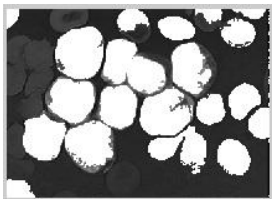


Figure 2(g) regional maxima superimposed on original image

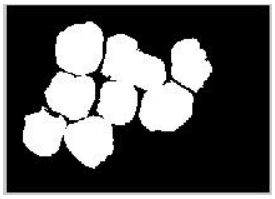


Figure 2(h) thresholded opening closing by reconstruction

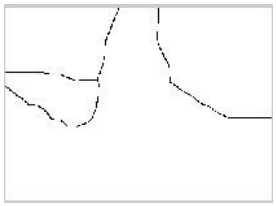


Figure 2(i) watershed ridge lines

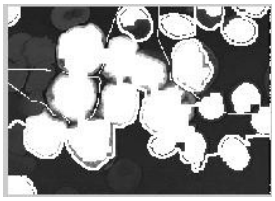


Figure 2(j) marker and object boundaries superimposed on original image

Figure 2(k) detect white cell using colored watershed label matrix

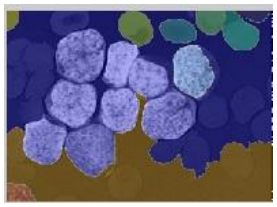


Figure 2(l) detection of white cell superimposed transparency on original image

Figure 2. Process Flow using Watershed Transform

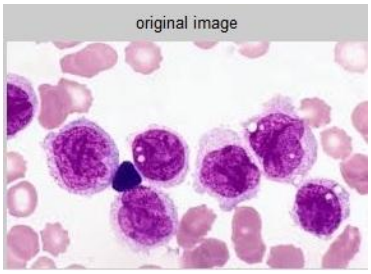


Figure 3(a) original image

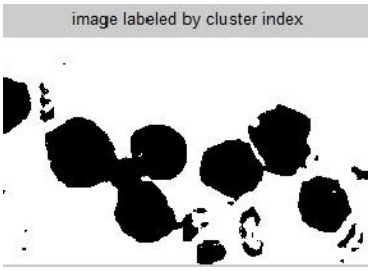


Figure 3(b) image labeled by cluster index

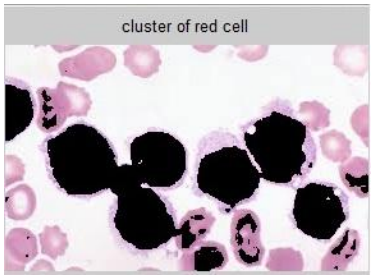


Figure 3 (c) cluster of red cells

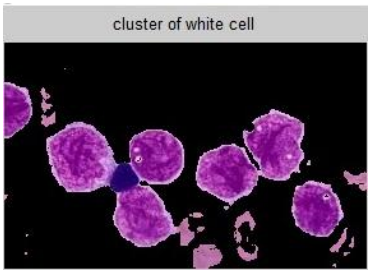


Figure 3 (d) cluster of white cells

Figure 3. Process Flow for K means Clustering

Process flow is shown in figure 3 that is used to separate white cell and red cell using K-means clustering method. Figure 3(a) has been taken from web sources [24].

2.3 Histogram Equalization and Linear Contrast Stretching

Raje, Chaitali, and Jyoti Rangole suggested method to detect white cells. To perform this operation, image is first converted from RGB to gray level and for contrast enhancement, histogram equalization process is used. Then statistical parameter like mean and standard deviation is calculated and erosion or dilation technique is used for morphological operation. Here sobel operator is used to detect the edge of cells. After that some shape based features can also be implemented if needed [6]. Process flow s is given in figure 4. Figure 4(a) has been taken from web resources [25].

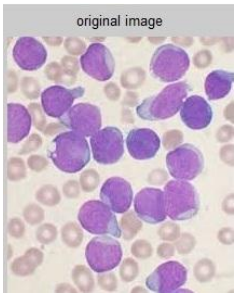


Figure 4(a) original image

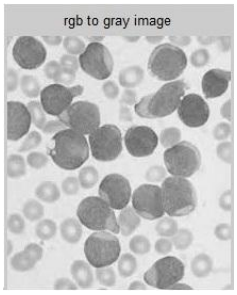


Figure 4(b) rgb2 gray

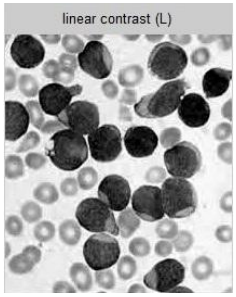


Figure 4(c) linear contrast

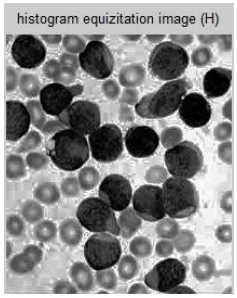


Figure 4(d) histogram equalization

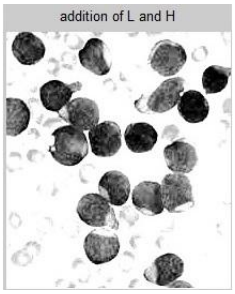


Figure 4(e) addition of 4 (c) and 4(d)

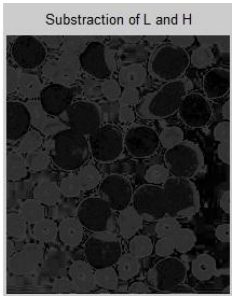


Figure4 (f) Subtraction of 4(c) and 4(d)

Figure4(g) Addition of fig, 4(e) and 4(f)

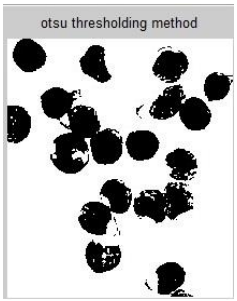


Figure 4(h) Thresholding method

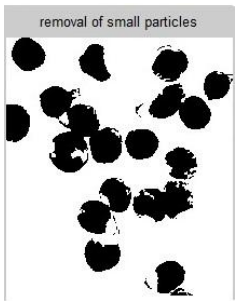


Figure 4(i) Removal of small particles

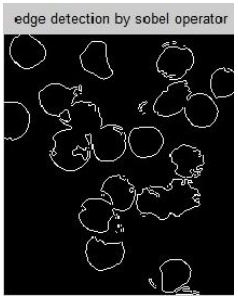


Figure 4(j) Edge detection using sobel operator

Figure 4. Process flow for histogram equalization and linear contrast stretching.

2.4 HSI Color Model based Segmentation

For segmentation of WBC and nucleus, linear contrast stretching techniques is used and for color based segmentation HSI color model is used. For further segmentation, K means clustering techniques is used. For fully segmented WBC, extract the H component and to segment nucleus S component is necessary. For further segmentation, k means clustering is used. Then unwanted object or noise from image are removed using median filter. Thus cytoplasm and nucleus can be extracted from image [15].

Process flow is shown in figure 5. Figure 5(a) has been taken from web sources [23]

3. METHODOLOGY

After getting the knowledge about all these techniques, it is concluded that shape based features are used for better result and accuracy. Shape based features is used to detect different shapes like circle, rectangle, ellipse, squares etc. Our blood cells also have different size and shapes. So to detect geo metrical shape of cells, this shape based features are very useful method to detect different type of cells and their shapes.

In section 3.1 algorithm is given for performing different image processing operation on leukemia detected images. This algorithm is useful to detect number of overlapping and non-overlapping cells or to count red and white cells. First RGB image is converted into gray level to reduced dimension of image. After that thresholdig method is used to convert the image into binary form for more accurate analysis Otsu's method is best for thresholding, After that some morphological operation like area opening, closing, erosion, dilation. Here, area opening is performed to remove connected component Dilation is techniques to add pixels to boundary of objects and erosion is used to remove the pixel on object boundaries. After detecting the boundary of object, hole filling operation is performed that is used to detect perfect cell.

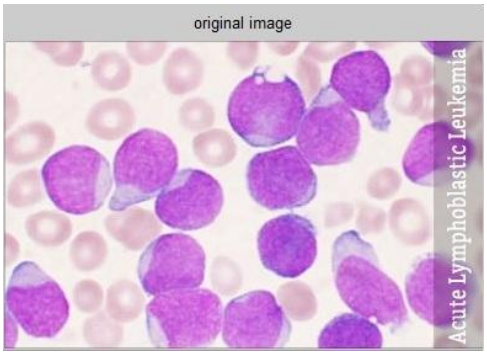


Figure 5(a) original image

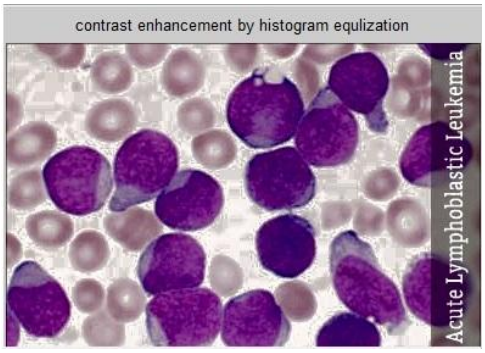


Figure 5(b) contrast enhance by histogram equalization

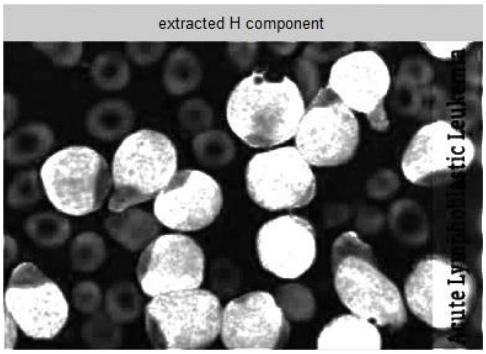


Figure 5 (e) extracted H component

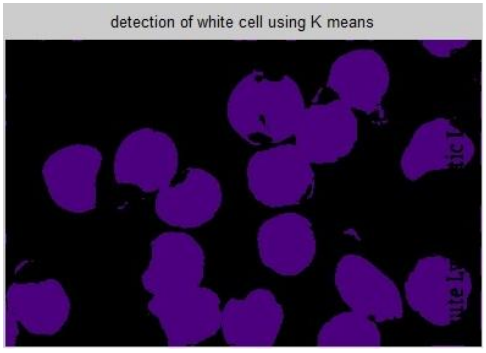
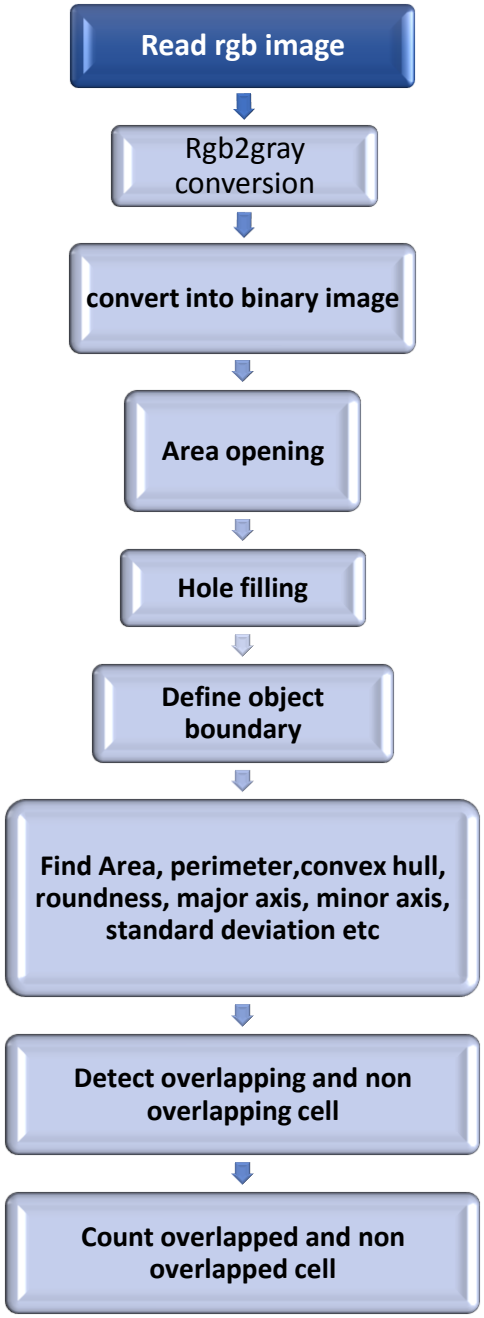


Figure 5(d) detection of white cell using K means

Figure 5. Process Flow of HSI Color Model based Segmentation

3.1 Algorithm to Count the Cells



After that hole filling operation, boundary is detected around the cells. To detect object boundaries many operators are used like sobel, prewitt, canny, etc. after define the boundary of object, many shape based features, like major axis, minor axis, area, perimeter, standard deviation, radius, roundness can be found. Radius, roundness, standard deviation can be found by given formula:

Radius = $\frac{Major\ axis+Minor\ axis}{4}$

Roundness = $\frac{4*\pi*area}{perimeter^2}$

Standard deviation=
$\sqrt{(Major\ axis - X)^2 + (Minor\ axis - X)^2}$

Where X= $\frac{Major\ axis+Minor\ axis}{2}$ (Mean)

According to value of major axis and minor axis, number of over lapping cells and non-overlapping cells can be detected. According to detection of cells, number of cell can be counted.

Figure 6(a) shows original image, figure 6(b) is gray image of original image. Figure. 6(c) is binary image of original image and after that area opening is performed in figure 6(d).

Hole filling operation is defined in figure 6(e). After that boundary is defined for each object that shown in figure 6(f). Figure 6(g) is defined overlapped cell with the value of major axis. Figure 6 (h) and figure 6(i) shows radius and roundness of cell. Figure 6 (j), figure 6(k) and figure 6(l) represent features like centroid, standard deviation and bounding box of cells respectively. This algorithm is performed on MATLAB platform. In MATLAB, 'region props' properties are used to find area, roundness, centroid, major axis, minor axis, standard deviation, etc.

Here the number of actual objects or cells (red cells+ white cells) are 92. Here, number of non over lapping and overlapping cells are detected is 68 and 11 respectively. So the cells which are overlapped on each other are calculated as (X*2). Where X is number of overlapped cells. So (11*2=22). Thus, according to algorithm total number of cells are 90. This result is very accurate and nearly equal to actual result. Because actual count is 92. So shape based features are more accurate and gives 97.8 % accuracy. Hence proved.

3.2 Experimental Results

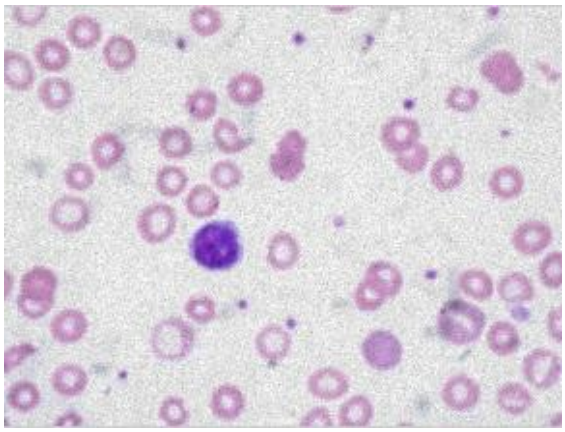


Figure 6 (a) original image

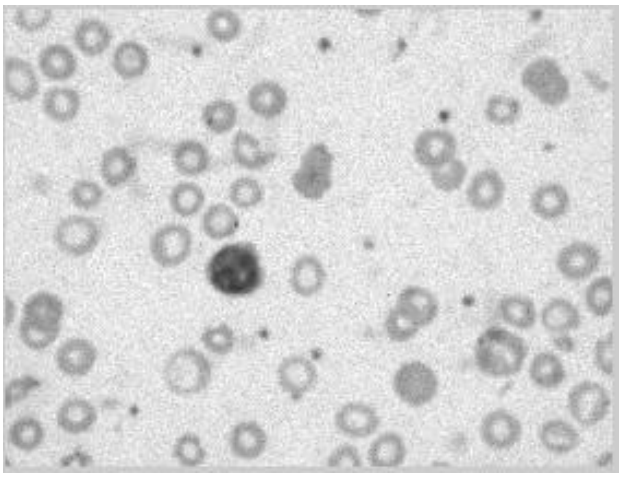


Figure 6(b) rgb2gray

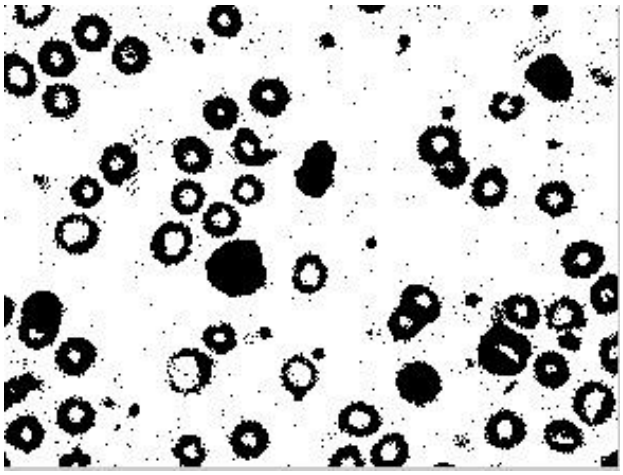


Figure 6(c) thresholding

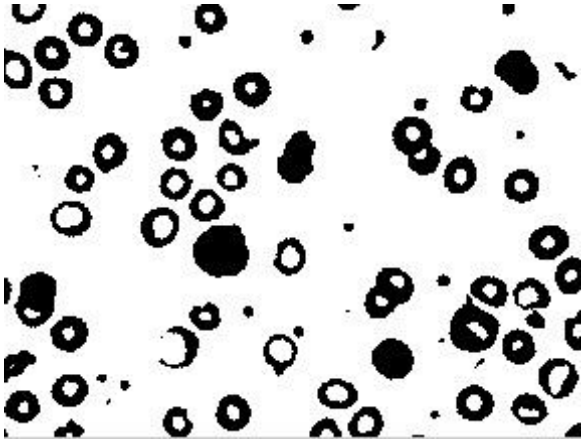


Figure 6(d) area opening

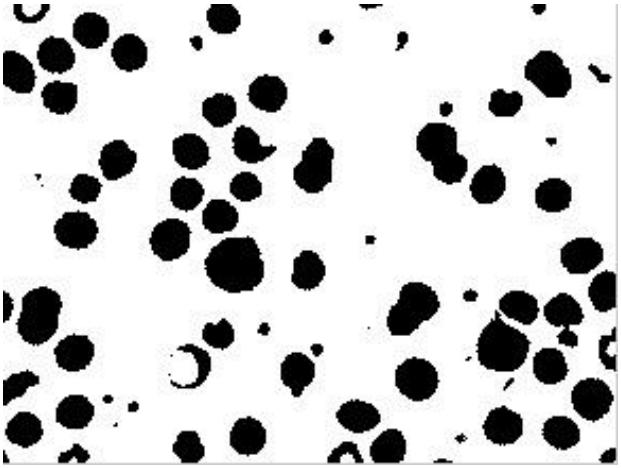


Figure 6 (e) hole filling

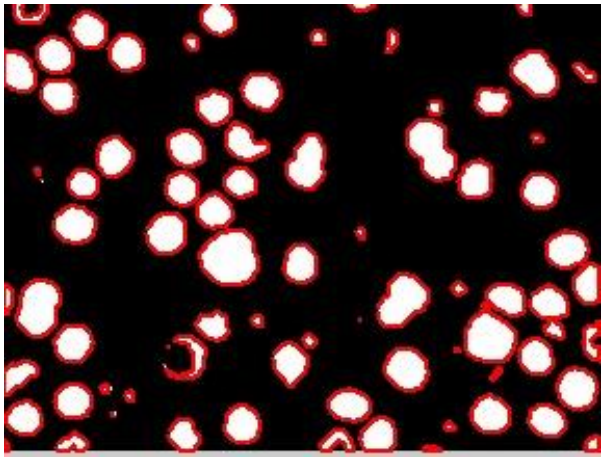


Figure 6 (f) detected boundary

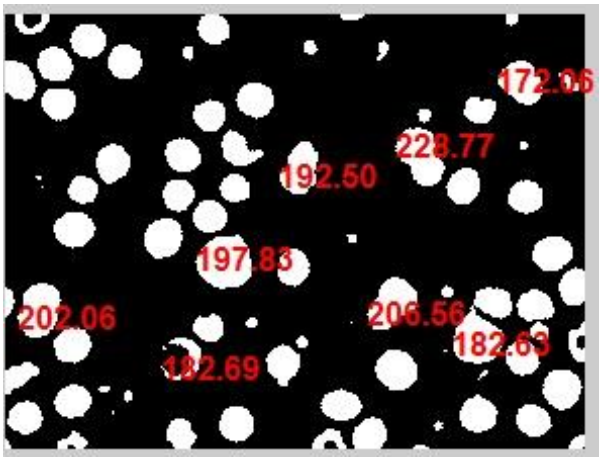


Figure 6(g) Detected overlapping cell

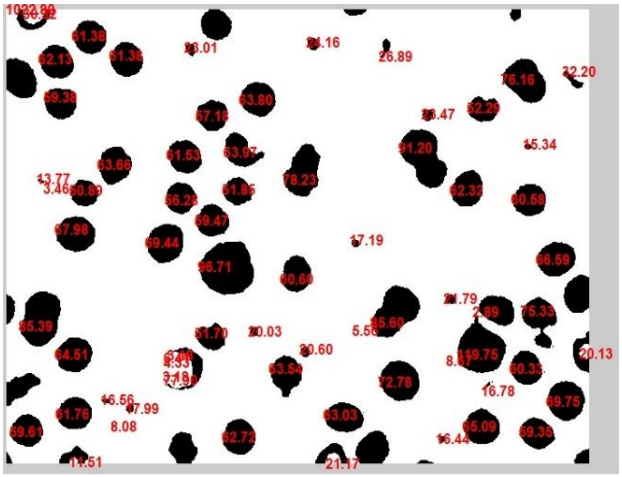


Figure 6 (h) radius of cell

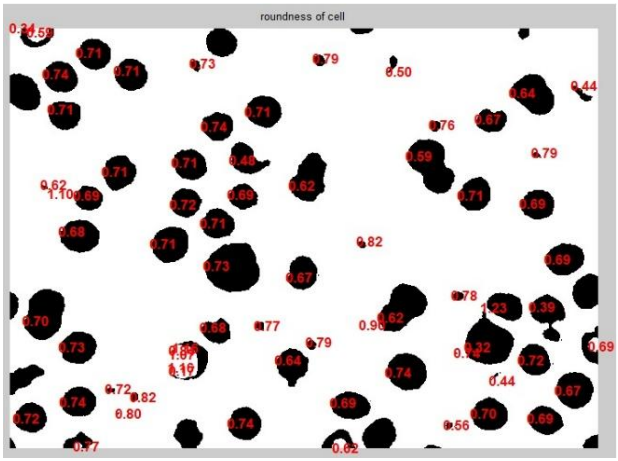


Figure 6(i) roundness of cell

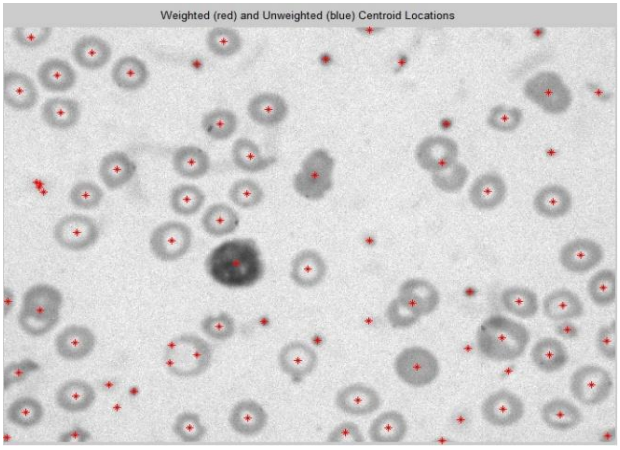


Figure 6(j) centroid of cell

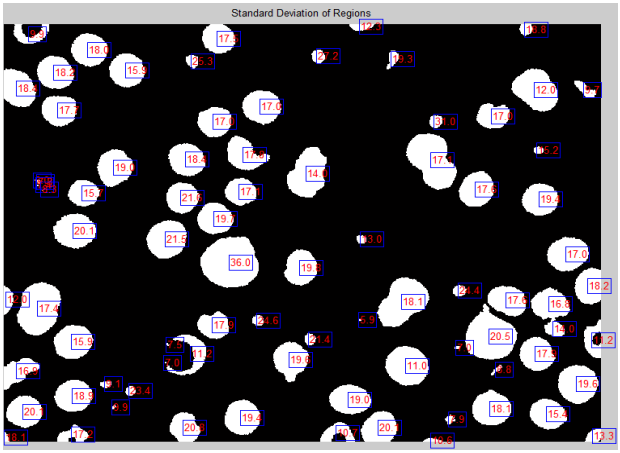


Figure 6(k) standard deviation

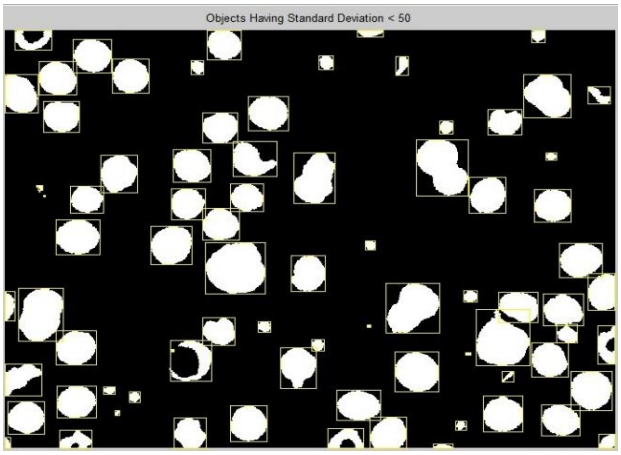


Figure 6(l) bounding box

Figure 6 Process Flow of Shape based Features

4. CONCLUSION

Various methods have been discussed and evaluated. Merits and demerits of different methods are derived from experimental results. It is concluded that proposed method: shape based features finding is more accurate than other methods for counting leukemic cells and it also gives highest accuracy 97.8 % as shown in table below. To detect different types of geometrical shape of cells like basophils, eosinophil, lymphocytes, monocytes etc. shape based features are used and according to count of immature cells, disease can be diagnosed. These are the limitations of watershed transform, K means clustering, and histogram equalizing.

Method	Merit/demerits	Accuracy
Watershed transform	Merit: Easy method for detection of white cell Demerits: It cannot give accurate result and cannot implement on each and every image	72.2 %
K means clustering	Merits: It is used for clustering and separate the data based on value of K. Demerits: It does not give classification with labeled data and also not applicable on incremental data.	72 %
Edge detection using histogram equalizing method and linear contrast stretching	Merit: This is very useful method to detect white cell and for contrast enhancement. Demerit: It is hard to define boundary of overlapping cell.	73.7 %
Shape based features	Merit: Very easy for the detection of white and overlapping cell and shape of cell Demerit: This is based on statistics so can get approximate result.	97.8 %

5. REFERENCES

[1] Mohapatra, Subrajeet, Sushanta Shekhar Samanta, Dipti Patra, and Sanghamitra Satpathi. "Fuzzy based blood image segmentation for automated leukemia detection." In *Devices and Communications (ICDeCom), 2011 International Conference on*, pp. 1-5. IEEE, 2011.

[2] Lim, Huey Nee, Mohd Yusoff Mashor, and Rosline Hassan. "White blood cell segmentation for acute leukemia bone marrow images." In *Biomedical Engineering (ICoBE), 2012 International Conference on*, pp. 357-361. IEEE, 2012.

[3] Fatma, Mashiat, and Jaibir Sharma. "Identification and classification of acute leukemia using neural network." In *Medical Imaging, m-Health and Emerging Communication Systems (MedCom), 2014 International Conference on*, pp. 142-145. IEEE, 2014.

[4] Madhloom, H. T., S. A. Kareem, H. Ariffin, A. A. Zaidan, H. O. Alanazi, and B. B. Zaidan. "An automated white blood cell nucleus localization and segmentation using image arithmetic and automatic threshold." (2010).

[5] Halim, NH Abd, M. Y. Mashor, A. S. Abdul Nasir, N. R. Mokhtar, and H. Rosline. "Nucleus segmentation technique for acute leukemia." In *Signal Processing and its Applications (CSPA), 2011 IEEE 7th International Colloquium on*, pp. 192-197. IEEE, 2011.

[6] Raje, Chaitali, and Jyoti Rangole. "Detection of Leukemia in microscopic images using image processing." In *Communications and Signal Processing (ICCSP), 2014 International Conference on*, pp. 255-259. IEEE, 2014.

[7] Mohapatra, Subrajeet, and Dipti Patra."Automated leukemia detection using hausdorff dimension in blood microscopic images." In *Emerging Trends in Robotics and Communication Technologies (INTERACT), 2010 International Conference on*, pp. 64-68. IEEE, 2010.

[8] Berge, Heidi, Dale Taylor, Sriram Krishnan, and Tania S. Douglas. "Improved red blood cell counting in thin blood smears." In *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on*, pp. 204-207. IEEE, 2011.

[9] Mazalan, Siti Madihah, Nurul H. Mahmood, and Mohd Azhar Abdul Razak. "Automated Red Blood Cells Counting in Peripheral Blood Smear Image Using Circular Hough Transform." In *Artificial Intelligence, Modelling and Simulation (AIMS), 2013 1st International Conference on*, pp. 320-324. IEEE, 2013.

[10] Akrimi, Jameela Ali, Azizah Suliman, Loay E. George, and Abdul Rahim Ahmad. "Classification red blood cells using support vector machine." In *Information Technology and Multimedia (ICIMU), 2014 International Conference on*, pp. 265-269. IEEE, 2014.

[11] Mohapatra, Saurav, Dipti Patra, Sudhakar Kumar, and Siddhartha Satpathi. "Kernel induced rough c-means clustering for lymphocyte image segmentation." In *Intelligent Human Computer Interaction (IHCI), 2012 4th International Conference on*, pp. 1-6. IEEE, 2012.

[12] Ge, Jia, Z. Gong, Jiann-Jong Chen, Jiangchuan Liu, John Nguyen, Z. Y. Yang, Chingyue Wang, and Yue Sun. "A system for automated counting of fetal and maternal red blood cells in clinical KB test." In Robotics and Automation (ICRA), 2014 IEEE International Conference on, pp. 1706-1711. IEEE, 2014.

[13] Supardi, N. Z., M. Y. Mashor, N. H. Harun, F. A. Bakri, and R. Hassan. "Classification of blasts in acute leukemia blood samples using k-nearest neighbour." In Signal Processing and its Applications (CSPA), 2012 IEEE 8th International Colloquium on, pp. 461-465. IEEE, 2012.

[14] Mohammed, Emad, Mostaja MA Mohamed, Christopher Naugler, and Behrouz H. Far. "Chronic lymphocytic leukemia cell segmentation from microscopic blood images using watershed algorithm and optimal thresholding." In Electrical and Computer Engineering (CCECE), 2013 26th Annual IEEE Canadian Conference on, pp. 1-5. IEEE, 2013.

[15] Abdul Nasir, A. S., M. Y. Mashor, and H. Rosline. "Unsupervised colour segmentation of white blood cell for acute leukaemia images." In Imaging Systems and Techniques (IST), 2011 IEEE International Conference on, pp. 142-145. IEEE, 2011.

[16] Rawat, Jyoti, A. Singh, H. S. Bhadauria, and I. Kumar. "Comparative analysis of segmentation algorithms for leukocyte extraction in the acute Lymphoblastic Leukemia images." In Parallel, Distributed and Grid Computing (PDGC), 2014 International Conference on, pp. 245-250. IEEE, 2014.

[17] Aimi Salihah, A. N., Mohd Yusoff Mashor, Nor Hazlyna Harun, and H. Rosline. "Colour image enhancement techniques for acute leukaemia blood cell morphological features." In Systems Man and Cybernetics (SMC), 2010 IEEE International Conference on, pp. 3677-3682. IEEE, 2010

[18] Das, Biplab Kanti, Krishna Kumar Jha, and Himadri Sekhar Dutta. "A New Approach for Segmentation and Identification of Disease Affected Blood Cells." In Intelligent Computing Applications (ICICA), 2014 International Conference on, pp. 208-212. IEEE, 2014..

[19] Mohapatra, Subrajeet, and Dipti Patra. "Automated cell nucleus segmentation and acute leukemia detection in blood microscopic images." In Systems in Medicine and Biology (ICSMB), 2010 International Conference on, pp. 49-54. IEEE, 2010.

[20] Madhloom, Hayan T., Sameem Abdul Kareem, and Hany Ariffin. "A Robust Feature Extraction and Selection Method for the Recognition of Lymphocytes versus Acute Lymphoblastic Leukemia." In Advanced Computer Science Applications and Technologies (ACSAT), 2012 International Conference on, pp. 330-335. IEEE, 2012.

[21] Putzu, Lorenzo, and Cecilia Di Ruberto. "White blood cells identification and counting from microscopic blood images." World Academy of Science, Engineering and Technology 7, no. 1 (2013): 363-370.

[22] (2015) Figure 1 website. [Online] Available at: www.ufrgs.br

[23] (2015) Figure 2(a) website. [Online] Available at www.doctortipster.com

[24] (2015) Figure 3(a) website. [Online] Available at: www.pathologystudent.com

[25] (2015) Figure 4(a) website.[Online] Available at: www.medicalxpress.com

[26] (2015) Figure 5(a) website.[Online] Available at: www.prezi.com

[27] (2015) Figure 6(a) website.[Online] Available at: www.commons.wikimedia.org