White blood cells nucleus segmentation from microscopic images of strained peripheral blood film during leukemia and normal condition

Conference Paper · May 2016

DOI: 10.1109/ICIEV.2016.7760026

CITATIONS
12

READS
205

3 authors:

Rakibull Ahasan
University of Wollongong
2 PUBLICATIONS **12** CITATIONS
SEE PROFILE

Ahasan Ulla
Technische Universität Dortmund
2 PUBLICATIONS **12** CITATIONS
SEE PROFILE

A S M Bakibillah
Monash University (Malaysia)
16 PUBLICATIONS **50** CITATIONS
SEE PROFILE

White Blood Cells Nucleus Segmentation from Microscopic Images of strained peripheral blood film during Leukemia and Normal Condition

Rakibull Ahasan[1], Ahasan Ulla Ratul[1], A. S. M. Bakibillah[2]
[1]UG Student, Dept. of EEE. FE, American International University-Bangladesh (AIUB)
[2]Assistant Professor. Dept. of EEE. FE, American International University-Bangladesh (AIUB)
Banani, Dhaka 1213, Bangladesh.
ahasanrakibull1994@gmail.com

***Abstract*—Counting of White blood cells (WBC) & characterizing their nucleus can provide valuable information to the doctors in order to identify different diseases or stage of a particular disease. The manual method is a tiresome process and has a lot of inaccuracy. On the other hand, the machine (hematological analyzer) based method are very expensive. Digital image processing can be a less time consuming and cost effective method for counting and characterizing the WBC. In order to Count or characterize WBC by image processing, proper segmentation is the key challenge. In this paper we proposed an algorithm to segment WBC nucleus from microscopic images of stained peripheral blood film during leukemia and normal condition. The proposed algorithm involves different steps such as color space conversion, color thresholding, filtering, marker controlled watershed and different morphological operations. The accuracy of the result obtained is 88.57%.**

Keywords- White blood cells (WBC); segmentation; leukemi; morphological operatio; color space.

I. INTRODUCTION

Digital image processing provides us with the opportunity to extract some meaningful and useful information from an image. In case of counting WBC and characterizing their nucleus digital image processing has the advantage on required time and cost over the traditional methods. To accomplish that the key step is the proper segmentation of WBC. Previously many researchers worked on blood cell segmentation. Edge detection is a quite efficient approach for WBC and their nucleus segmentation [1][2]. Where, different color space conversion is used for segmenting three different kinds of blood cells [3][4]. Power law transformation was also used for WBC sectionalisation [5]. Boundary support vector is another interactive method suggested by Min Wang [6]. Mostafa Mohamed and Behrouz Far proposed two different segmentation algorithms for WBC nucleus. One is based on green component of the image [7] and another one is based on Gram-Schmidt vector manipulation [8]. They also introduced a dual segmentation technique for nucleus and cytoplasm, based on canny edge detection, morphological operation, chessboard distance transformation, local minima elimination and watershed transformation [9]. Layza Baldo also proposed such a dual technique where self-Dual Multiscale morphological Toggle (SMMT) used for nucleus and granulometric analysis for cytoplasm segmentation [10]. Another such algorithm which involves teager vitality operator for nucleus and structural hustler for cytoplasm segmentation [11]. One of the advanced technique for segmenting WBC nucleus and cytoplasm are Gradient Vector Flow Snake algorithm and Zack thresholding [12]. Very well known K-means clustering has imposed its bright spot on this research area. Neelam sinha used K-means clustering followed by EM-algorithm [13], Huey Nee Lim proposed traditional K-means clustering with skeleton by influence zone [14] whereas Subrajeet Mohapatra prepare a mixed breed of rough sets and K-means clustering [15]. Another advanced algorithm was also developed by using Fuzzy C-Means clustering [16]. The recent researches of this particular field also involve the application of artificial neural network (ANN). ANN uses the morphological features of the WBCs in order to classify them [17][18]. In this research, we have proposed a robust WBC nucleus segmentation technique from the microscopic image of stained peripheral blood film. The segmentation technique is able to segment nucleus during leukemia and normal condition. The main difference between leukemia and normal condition is the no. of WBC, during leukemia no. of WBC is increased due to presence of blast cells (immature WBC). This result the WBC becomes connected with each other. The proposed segmentation technique comprise RGB to L*a*b* color space conversion as RGB is perceptually nonlinear, color thresholding, average filtering to remove noise, watershed segmentation and different kinds of morphological operations.

II. PROPOSED ALGORITHM

The main goal of the work is to segment the WBC from the microscopic image of stained peripheral blood film during leukemia and normal condition. The main difference between these two conditions is the connectivity among WBCs. During normal condition the WBCs are separated from each other and during leukemia those become connected, due to the presence of blast cell. The proposed algorithm has been developed in such a manner that the WBCs can be separated during leukemia and the separation process does not hamper during normal condition. The microscopic images of stained

978-1-5090-1269-5/16/$31.00 ©2016 IEEE

peripheral blood film has been collected from the university degliStudi di Milano, Italy, provided by Fabio Scotti. The proposed algorithm requires color space conversion, color thresholding, morphological operation etc. The entire programming is done by using MATLAB. Fig. 1 shows the proposed Algorithm to segment WBC Nucleus.

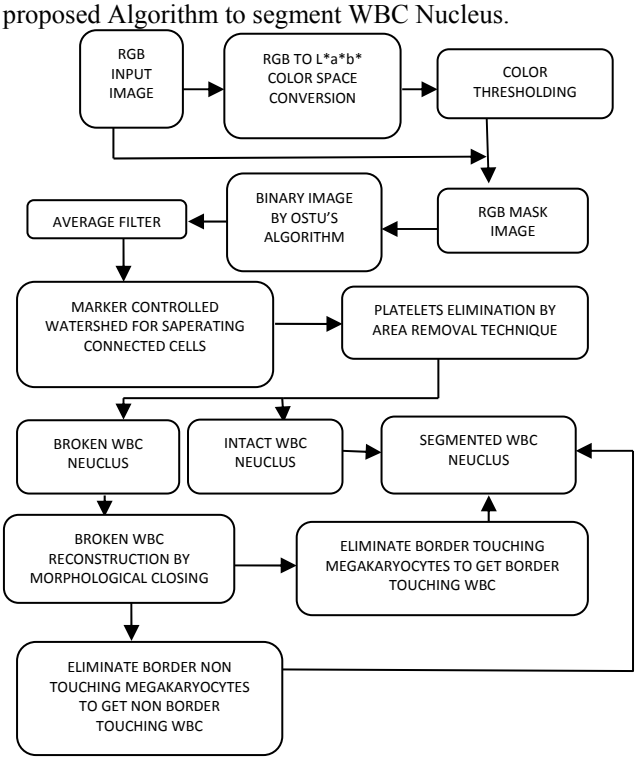

Figure 1. Block diagram of proposed algorithm.

RGB to L*a*b* color space conversion and color thresholding provides the segmented image of WBC nucleus, platelets and megakaryocytes (if present). Average filter is applied to remove any kind of random noise. The value for applying average filter is chosen as a way that it will not cause and harm to the images. The marker controlled watershed is applied to separate the connected WBC nucleus during leukemia condition. The platelets and megakaryocytes have been removed by area removal technique and morphological closing operation is applied to reconstruct the broken WBC due to marker controlled watershed function. The main goal of the research is to segment the WBC nucleus from the microscopic image of stained peripheral blood film.

To demonstrate the algorithm, some random images were chosen from the data set shown in fig. 2. Fig 2(a) and 2(b) are the images of leukemia condition. Fig. 2(c), 2(d), 2(e) and 2(f) are the images of normal condition. In normal condition WBCs are separated from each other but 2(d), 2(e) and 2(f) are the images which contain megakaryocytes. These are the immature stage of the platelets. Color is same as WBC and morphological texture is bigger than mature platelets but smaller than WBC. All the input images were in the RGB color space. The mathematical equation of the RGB chromaticity coordinates is [19] –

$$r = \frac{R}{R+G+B}; \quad g = \frac{G}{R+G+B}; \quad b = \frac{B}{R+G+B} \tag{1}$$

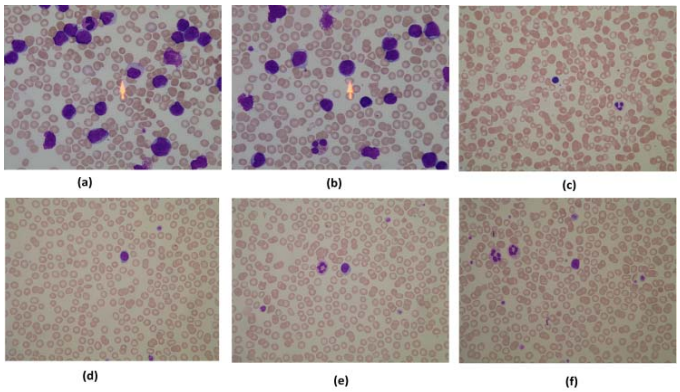

Figure 2. Input images in RGB format.

Where, 'R', 'G' and 'B' denote red, green and blue. The input images of RGB color space is converted into L*a*b* color space image. It is a two-step operation. First the RGB color space image needs to be converted into XYZ which can be done by the following simple matrix, after normalizing each RGB value [20]-

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.412453 & 0.357580 & 0.180423 \\ 0.212671 & 0.715160 & 0.072169 \\ 0.019334 & 0.119193 & 0.950227 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{2}$$

Where, X, Y and Z denote the response of the mix of cones, the luminance and the response of S cone respectively. Then the XYZ color space image is converted into L*a*b*. The following equation shows the conversion process [20]-

$$L^* = 116 f\left(\frac{Y}{Y_n}\right) - 16 \tag{3}$$

$$a^* = 500\left[f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right)\right] \tag{4}$$

$$b^* = 200\left[f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right)\right] \tag{5}$$

Where, $f(s) = s^{1/3}$ *for* $s > 0.008856$

and $f(s) = 7.787s + \frac{16}{116}$ *for* $s \le 0.008856$

Where 'L' denoted the lightness, 'a' indicates red or green and 'b' indicates blue or yellow. After the proper conversion of color space a range of threshold value is chosen from the histogram analysis, for channel L*, a* and b*. Then the matrix of L*a*b*color space image is repeated and all the pixel value set to zero which doesn't fall in the range of chosen thresholding values for channel L*, a* and b*. This process returns an L*a*b*color space image containing only the WBC nucleus, platelets and megakaryocytes (if present). Resultant L*a*b* color space mask images are shown in the Fig. 3.

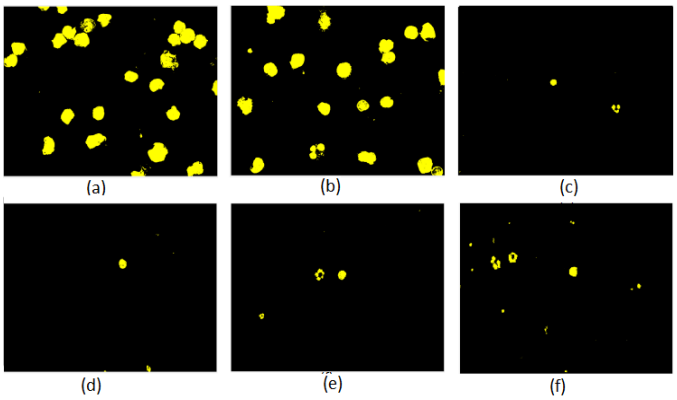

Figure 3. WBC mask image in L*a*b* color space.

Now, the matrix of RGB input image is repeated and the all the pixel values are set to zero which were zero in the L*a*b* color space mask image. Fig. 4 shows the resultant RGB color space mask image.

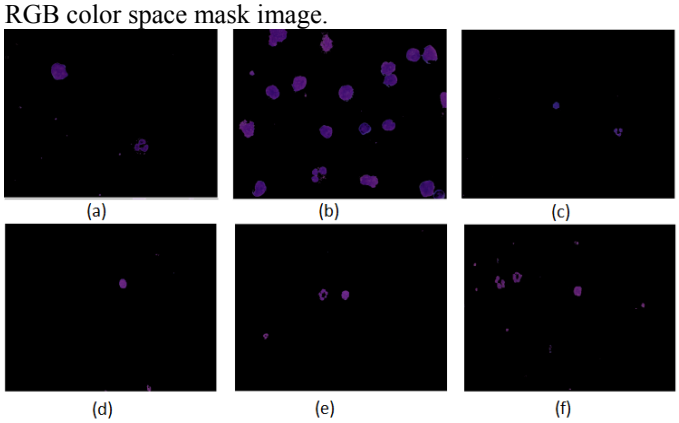

Figure 4. WBC mask image in RGB color space.

Then the RGB mask image is converted into gray scale by the following mathematical equation-

$$0.2928 * R + 0.5870 * G + 0.1140 * B \quad (6)$$

Otsu's global thresholding technique is applied to identify a global thresholding value of the gray scale image. The mathematical equation of Otsu's global thresholding technique is [21]-

$$\sigma_w^2(T) = v_1(T)\sigma_1^2(T) + v_2(T)\sigma_2^2(T) \quad (7)$$

Where, σ_w is the within class variance, σ_i is the class variance of foreground or background, T is the optimum thresholding value to binarize the image and class probabilities, v_1 and v_2, are estimated from histogram. And using the global thresholding value the gray scale image is converted into binary image by the following technique [21]-

$$I_{bin}(x,y) = \begin{cases} 1, & if \quad I_{gray}(x,y) \geq T \\ 0, & otherwise \end{cases} \quad (8)$$

An average filter is applied on the binary image to eliminate any kind of random noise. Fig. 5 shows the resultant images.

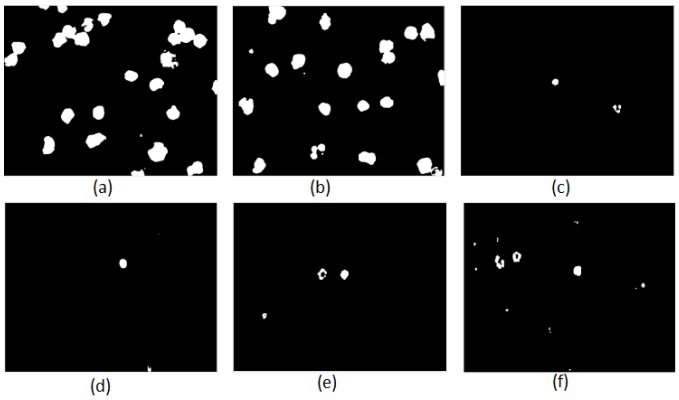

Figure 5. Images after applying average filter.

Then marker controlled watershed is applied to separate connected WBC during leukemia. The marker controlled watershed requires two images one is marker and another is mask. The marker image is created by Euclidean distance transform. This function determines the distance from every zero pixel to its nearest nonzero pixel. The mathematical equation of this operation is [2]-

$$D_{Euclidean}((x_1,y_1),(x_2,y_2)) = \sqrt{(x_1-x_2)^2 + (y_1-y_2)^2} \quad (9)$$

Where, x1 and x2 are the coordinates point in x-axis, y1 and y2 are the coordinates point in y-axis. The mask image is created by computing extended minima transform. This operation filters out small local minima and produce tiny element in the center of cells [22]. Fig. 6 shows the marker and mask image together.

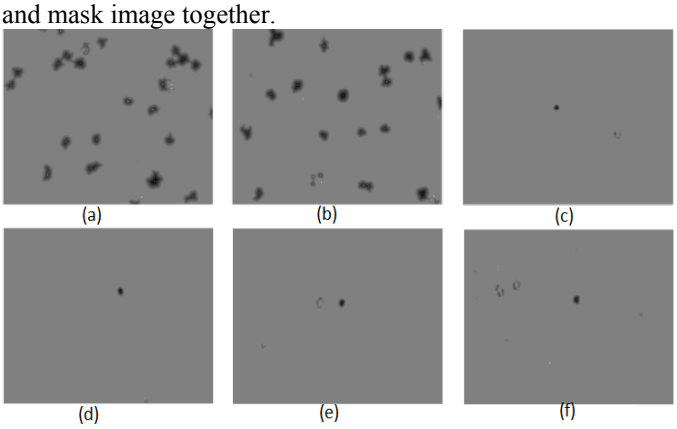

Figure 6. Marker and mask image.

A morphological reconstruction based operation is applied between marker and mask image which modify the marker image such that it has only the regional minima when the mask image is non zero. The mathematical equation of the morphological reconstruction is [23]-

$$h_{k+1} = (h_k \oplus B) \cap g \quad (10)$$

Where, g is the mask image, B is the structuring element, h_1 will be initialize as marker image and the operation will be iterate until the $h_{k+1} = h_k$. Then watershed function is applied which identify the catchment basins on the image [24]. And returns a label matrix where '0' represent the background and the watershed line, 1 represent the 1st watershed region, 2

represent the 2nd watershed region and so on. The resultant image of watershed line is shown in Fig. 7.

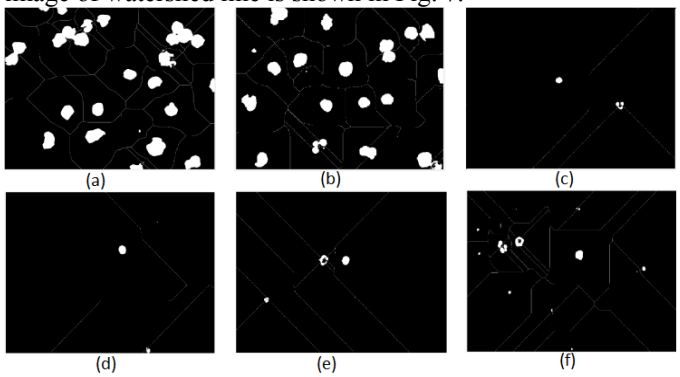

Figure 7. Watershed lines separate the connected WBC nucleus

The connected objects are separated by setting all the pixel values of binary image to zero which is on the label matrix of the watershed function. But this technique breaks down the nucleus of neutrophil, eosinophil and monocyte. For proper segmentation, reconstruction of the broken nucleus is essential. But before that the platelets and isolated pixels are removed by area removal technique. These are very small in size compare to WBC, so a thresholding value of area is chosen and all the objects are deleted which has less area than the chosen thresholding value. Result, all the platelets and isolated pixels are eliminated from the image. But as we know the megakaryocytes are bigger than the mature platelets so the images still contains the megakaryocytes. Resultant image is shown in Fig. 8.

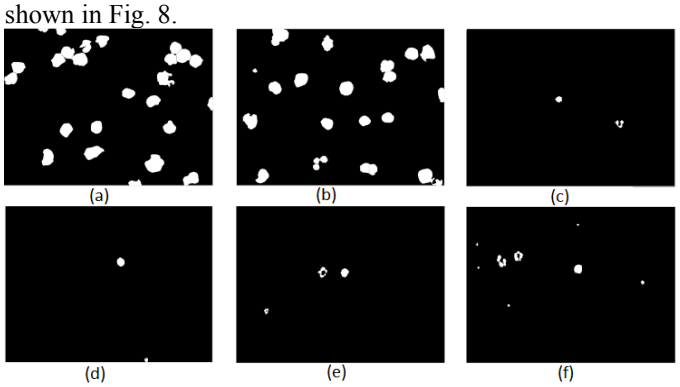

Figure 8. Images after eliminating Platelets and isolated pixels.

To reconstruct the broken WBC we need to distinguish between intact WBC (doesn't break down by applying marker controlled watershed) and broken WBC. The broken parts of the WBC are smaller in size so a thresholding value of area is chosen to delete all the broken part of WBC. The resultant image is shown in the Fig. 9.

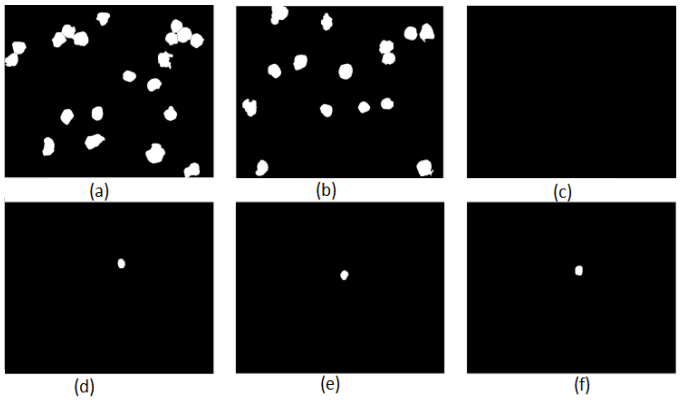

Figure 9. WBC that wasn't broken due to watershed

To return the broken WBC set based subtraction operation is done between Fig. 8 and Fig. 9. In order to reconstruct them morphological closing operation is applied with a disk shape structuring element of radius 5. The mathematical equation of morphological closing is [23]-

$$A \bullet B = (A \oplus B) \Theta B \tag{11}$$

Where, 'A' is the image and 'B' is the structuring element. The resultant image is shown in Fig. 10.

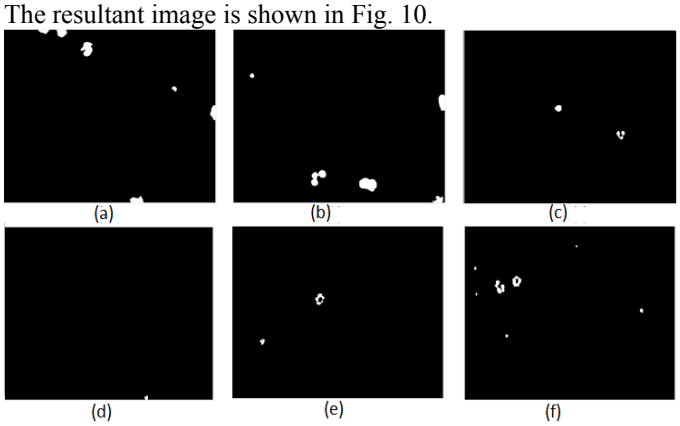

Figure 10. Broken WBC is reconstructed.

The images still contains the megakaryocytes. As the megakaryocytes are smaller in size compare to WBC. So those can be eliminated from the image by area removal technique. But it is not possible to eliminate them by choosing a single thresholding value because that will eliminate some of the border touching WBC. To overcome this problem first the image (Fig. 10) is segmented into two sub images. One contains all the border touching cells another contains all the border non touching cells. Than two different thresholding values of area is chosen to eliminate the border touching and border non touching megakaryocytes. Fig. 11 and Fig. 12 shows the resultant image after eliminating border touching and border non touching megakaryocytes.

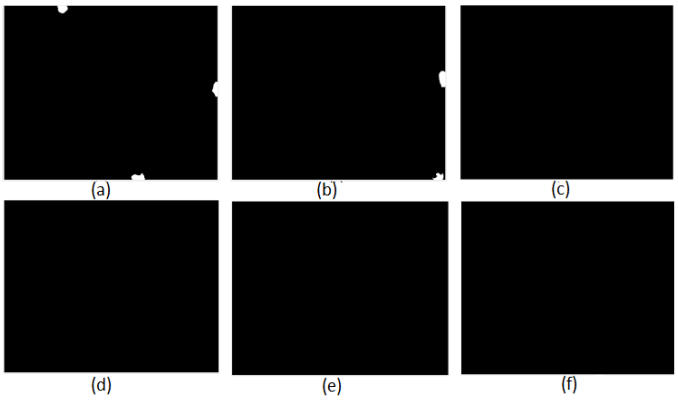

Figure 11. Images after eliminating border touching megakaryocytes

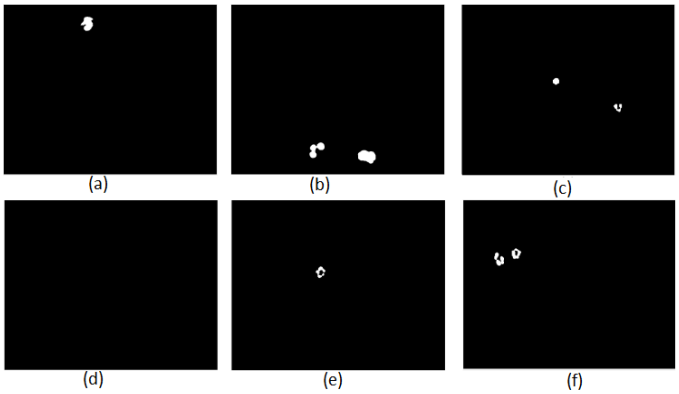

Figure 12. Images after eliminating border non touching megakaryocytes

All the WBCs of input image are present within Fig. 9, Fig. 11 and Fig. 12. So combination of these three images produces an identical binary image of the input image containing only the WBCs. Resultant image is shown in Fig. 13.

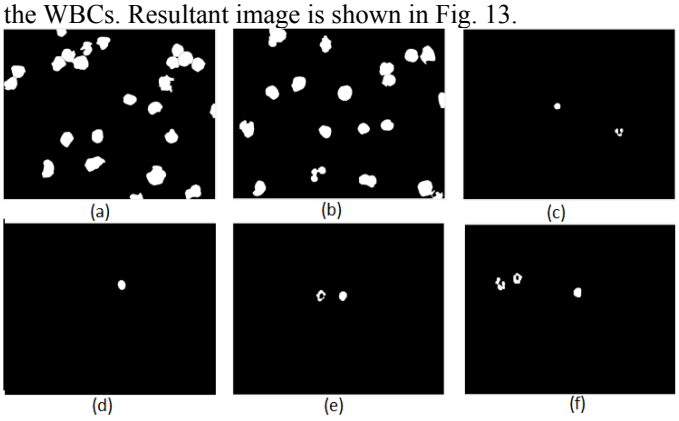

Figure 13. OR operations between Fig. no. 9, 12 and 13.

III. Results and Discussion

Applying the proposed algorithm on the 70 images of collected dataset, the accuracy of the algorithm was justified. Fig, 14 shows the step by step simulation of a normal condition image. Out of 51 normal condition images, 46 showed proper segmentation. So the rate of success of normal condition images is 90.2% for current dataset.

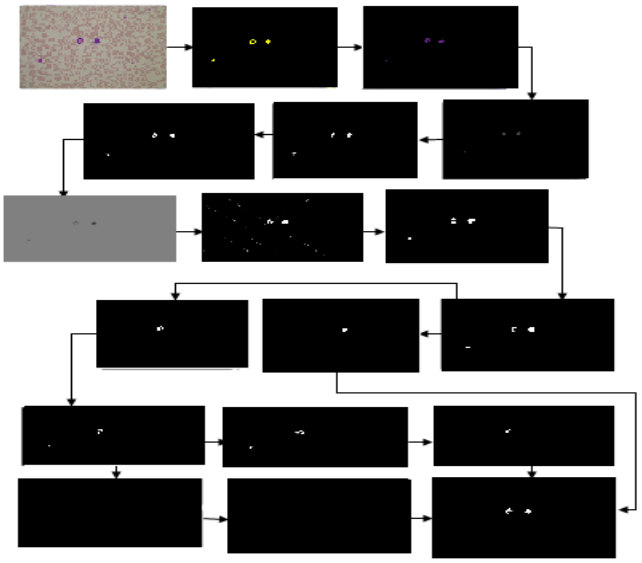

Figure 14. Segmentation during normal conditions.

For the images of leukemia condition, the success rate was calculated as 82.4%, which are 16 proper segmented images out of 19. Fig. 15 demonstrates the step by step segmentation process during leukemia condition.

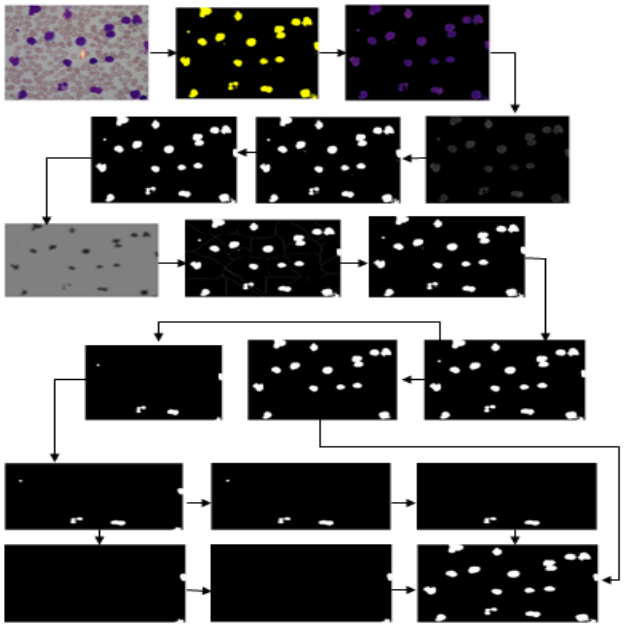

Figure 15. Segmentation during Leukemia

However, the overall success rate can vary. Good quality images of the dataset have shown a greater success on this occasion.

IV. Conclusion

Although, the proposed algorithm showed overall 88.57% accuracy considering both the condition using available dataset, the use of high quality images is suggested. The segmented image can be used for further nucleus characterization and classify the members of WBC family. There is also opportunity to improve the segmentation technique during different diseases. The further development is planned to publish it as an interactive application that will use the algorithm with slight modifications.

Acknowledgment

The authors are would like to thank Fabio Scotti for providing the dataset and Dr. Abdur Razzaque to verify the result by visual inspection.

References

[1] Jha, K.K.; Das, B.K.; Dutta, H.S., "Detection of abnormal blood cells on the basis of nucleus shape and counting of WBC," in Green Computing Communication and Electrical Engineering (ICGCCEE), 2014 International Conference on, vol., no., pp.1-5, 6-8 March 2014.

[2] Hari, J.; Prasad, A.S.; Rao, S.K., "Separation and counting of blood cells using geometrical features and distance transformed watershed," in Devices, Circuits and Systems (ICDCS), 2014 2nd International Conference on , vol., no., pp.1-5, 6-8 March 2014

[3] Tulsani, H.; Gupta, R.; Kapoor, R., "An improved methodology for blood cell counting," in Multimedia, Signal Processing and Communication Technologies (IMPACT), 2013 International Conference on, vol., no., pp.88-92, 23-25 Nov. 2013

[4] Sharif, J.M.; Miswan, M.F.; Ngadi, M.A.; Salam, M.S.H.; Mahadi bin Abdul Jamil, M., "Red blood cell segmentation using masking and watershed algorithm: A preliminary study," in Biomedical Engineering (ICoBE), 2012 International Conference on, vol., no., pp.258-262, 27-28 Feb. 2012

[5] Gopinath, L.; Sivasubramaniam, K.; Raguramajan, M.; Raajan, N.R., "Analysis of leukocyte blood cell for medical diagnosis using sectionalisation," in Advances in Electrical Engineering (ICAEE), 2014 International Conference on, vol., no., pp.1-4, 9-11 Jan. 2014

[6] Min Wang; Rong Chu, "A novel white blood cell detection method based on boundary support vectors," in Systems, Man and Cybernetics, 2009. SMC 2009. IEEE International Conference on, vol., no., pp.2595-2598, 11-14 Oct. 2009

[7] Mohamed, M.; Far, B., "An enhanced threshold based technique for white blood cells nuclei automatic segmentation," in e-Health Networking, Applications and Services (Healthcom), 2012 IEEE 14th International Conference on, vol., no., pp.202-207, 10-13 Oct. 2012.

[8] Mohamed, M.M.A.; Far, B.,"A Fast Technique for White Blood Cells Nuclei Automatic Segmentation Based on Gram-Schmidt Orthogonalization,"in Tools with Artificial Intelligence (ICTAI), 2012 IEEE 24th International Conference on, vol.1, no., pp.947-952, 7-9 Nov. 2012.

[9] Mohammed, E.A.; Mohamed, M.M.A.; Naugler, C.; Far, B.H.,"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, Regina, SK, 2013, pp.1-5.

[10] Dorini, L.B.; Minetto, R.; Leite, N.J., "Semiautomatic White Blood Cell Segmentation Based on Multiscale Analysis," in Biomedical and Health Informatics, IEEE Journal of, vol.17, no.1, pp.250-256, Jan. 2013

[11] Kumar, B.R.; Joseph, D.K.; Sreenivas, T.V., "Teager energy based blood cell segmentation," in Digital Signal Processing, 2002. DSP 2002. 2002 14th International Conference on, vol.2, no., pp. 619-622 vol.2, 2002

[12] P.S.Hiremath, ParashuramBannigidad and SaiGeeta,"Automated Identification and Classification of White Blood Cells (Leukocytes) in Digital Microscopic Images,"in IJCA, Special Issue on, RTIPPR (2):59–63, 2010. Published By Foundation of Computer Science.

[13] Sinha, N.; Ramakrishnan, A.G., "Automation of differential blood count," in TENCON 2003. Conference on Convergent Technologies for the Asia-Pacific Region, vol.2, no., pp.547-551 Vol.2, 15-17 Oct. 2003.

[14] Huey Nee Lim; Mashor, M.Y.; Supardi, N.Z.; Hassan, R.,"Color and morphological based techniques on white blood cells segmentation,"in Biomedical Engineering (ICoBE), 2015 2nd International Conference on ,vol., no., pp.1-5, 30-31 March 2015.

[15] Mohapatra, S.; Patra, D.; Kumar, K., "Blood microscopic image segmentation using rough sets,"in Image Information Processing (ICIIP), 2011 International Conference on ,vol., no., pp.1-6, 3-5 Nov. 2011.

[16] K.S. Kim; P.K. Kim; J.J.; Y.C. Park, "Analyzing Blood Cell Image to Distinguish Its Abnormalities, "in Proceeding of the eighth ACM international conference on Multimedia, 2011.

[17] Q. Zheng; B. K. Milthorpe; A. S. Jones, "Direct neural network application for automated cell recognition," in Cytometry, vol. 57A issue 1, p.p. 1-9, 2004.

[18] F. Scotti, "Automatic morphological analysis for acute leukemia identification in peripheral blood microscope images," CIMSA. 2005 IEEE International Conference on Computational Intelligence for Measurement Systems and Applications, 2005, pp. 96-101.

[19] R.W.G. Hunt; M.R. Pointer, "RGB Colorimetry," in Measuring Color, 4th ed, A John Wiley & Sons, Ltd.,, 2011, PO19 8SQ, United Kingdom ch.11, sec. 11.4, pp. 231–240.

[20] Jim Schwiegerling, "Color," in Visual and Ophthalmic Optics, series ed., USA: SPIE Press, 2004, pp. 66–73.

[21] Gautam, A.; Bhadauria, H., "Classification of white blood cells based on morphological features," in Advances in Computing, Communications and Informatics (ICACCI, 2014 International Conference on, vol., no., pp.2363-2368, 24-27 Sept. 2014.

[22] Soille, P. , "Morphological Image Analysis: Principles and Applications", Springer-Verlag, 1999, pp. 170-171

[23] Rafael C. Gonxalez, Richard E. Woods and Steven L., "Morphological image processing", in Digital image processing using Matlab, 2nd ed. New Jersey, USA: Pearson Prentice Hall Pearson Education, 2009, ch. 10, pp. 486-534.

[24] Mayer, Fernand, "Topographic distance and watershed lines", in Signal Processing, Vol. 38, July 1994, pp. 113-125.