Comparative Analysis of Segmentation Algorithms for Leukocyte Extraction in the Acute Lymphoblastic Leukemia Images

Jyoti Rawat, A. Singh [2], H.S. Bhadauria [3], I. kumar [4]

[1,4]Ph.D Research Scholars, CSE Dept. G. B. P. E. C., Pauri, Garhwal, Uttrakhand (INDIA)

E-mail: [1]jyoty.cse@gbpec.ac.in, [4]indrajeet.cse@gbpec.ac.in

[2,3]Assistant professors, CSE Dept. G. B. P. E. C., Pauri, Garhwal, Uttrakhand (INDIA)

E-mail: [2]annapurnasingh78@gmail.com, E-mail: [3]hsb76iitr@gmail.com

Abstract: **By concerning with the health of the patients, analysis of blood cell particularly morphological structure of leukocyte in microscopic blood smear can effectively detect the important blood disorder such as the Acute Lymphoblastic Leukemia. Unfortunately, the analysis made by hematology expert is not always accurate and rapid due to the error prone modality and operator's incapability's. The presented paper shows the how to enhance the microscopic blood smear by removing the unwanted microscopic background by segmenting the cell and analyzing the existing algorithms for segmentation by performing comparative study for effectively segments and accurately measure the leukocyte characteristics in order to allow, subsequent automatic diagnosis of leukemia and other hematic diseases.**

Keywords: Blood Analysis, Leukocyte segmentation, Color L*a*b space, K-means Clustering, Marker Controlled Watershed, HSV color space, Global Thresholding.

I.   INTRODUCTION

Blood smear evaluation concerning the presence of hematologist gives important qualitative and quantitative information. From decades , the experienced hematic experts performed these operation, which basically diagnose two main analyses ,obtained like the different type of qualitative study of the morphology of the blood cells with the information of degenerative and tumoral pathologies such as Neutrophil leucocytosis, Chronic myeloid leukaemia, Chronic lymphocytic leukaemia, Prolymphocytic leukaemia, Follicular lymphoma, Hairy cell leukaemia, Multiple myeloma, Acute lymphoblastic leukaemia and other approach is quantitative as CBC (complete blood count) and differential counting the leukocytes cells.

Automated cell-counter systems like laser-based citometers are also available in medical field, but they have a drawback that they are not image- based morphological and also damage the blood samples during the blood analysis.

The major objective of the blood smear analysis is differentiating the components of blood and counting of red blood corpuscles, white blood corpuscles and platelets by observing the blood cell and also detecting various diseases like Anemia, Blood cancer and AIDS. The cell segmentation process is the separation of the cell from its complex background, and to break the cell into its morphological structures such as nucleus and cytoplasm [1, 2]. There are five types of white cell or leucocytes in the circulating blood in healthy person shown in Fig 1 [3, 4].
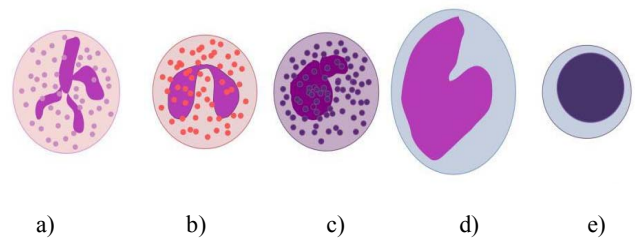

Fig. 1 A White blood cell classifications of normal blood smear.
a) Neutrophils b) Eosinophiles c) Basophiles d) Monocytes e) Lymphocytes

Image segmentation plays very crucial role in computer-aided diagnosis of medical images. The objective of image segmentation is to partition an image into non overlapping and homogeneous parts with respect to intensity and texture of that medical image.

For performing the segmentation in any image the basic step is to remove the noise from the source image(denoising) by applying the an appropriate filter such as low pass filter, vector median filter, high pass filter. After removing abrupt discontinuities (denoising), it becomes necessary to perform binary image conversion, which depends on value of the given threshold , that extract the particular objects from the background by removing outliers.

For identifying problem of any biomedical object or for extraction of an affected area clustering approach is feasible to use ,that perform the grouping, and every group is heterogeneous with each other and every element of single group is homogenous with similar characteristics and properties. K-means clustering, Genetic algorithm, Krushkal's algorithm, and fuzzy c-means algorithm are some clustering approach which have used for medical images.

Basic steps used for segmentation as shown in Fig 2 , which provides a brief knowledge of basic image processing steps used for applying segmentation on any image for finding a desired output .
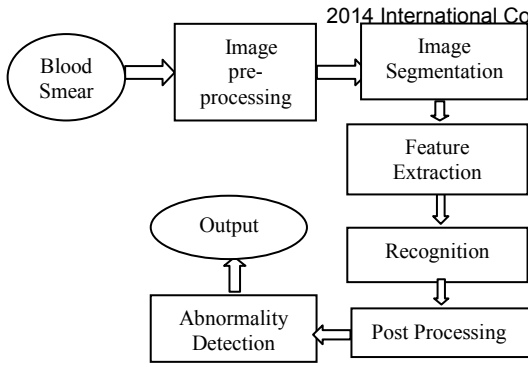
Fig. 2 Segmentation Approaches for medical image

The rest of this paper is organized as section II presents the segmentation techniques, it is followed by the related work in section III and then implementation and results in section IV, and section V explain the conclusion of the paper.

II. SEGMENTATION TECHNIQUES

Segmentation is a process of subdividing an image into the consequent parts or objects in the image, so the purpose of segmentation is to further analyze these consequent parts, once they are indentified they have to subdivided, so each of these consequent can be analyzed to extract some information so that information can be useful for high-level machine vision application, in case of hematology, information can be used for further classification and identification of hematic disease.

In segmentation there is some image that is processed and we get an output that is not an image, which is some parameter that we want to determine for further decision making process. Fig 3 shows an easy way to understand the classification of segmentation process.

Level of segmentation of an image depends on application on which we are applying the process. For example: analyze traffic pattern, here we will only find the ROI (region of interest).
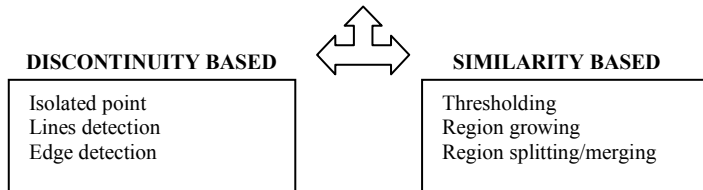

Fig. 3 Segmentation Approaches for medical image

In Discontinuity based, a subdivision of an image is carried out based on some abrupt changes in intensity level or gray level of an image [5].

Point Detection
In isolated point detection, we can use mask like this, where center is 8 and other are -1.

-1	-1	-1
-1	8	-1
-1	-1	-1

Fig. 4 Mask for Point detection

$$|R| \geq T \qquad (1)$$

R is a mask and T is non negative threshold value.

Line Detection
In Line detection, there are four masks which are used as the vertical mask, for detecting all the point which lie on a line which is vertical. Horizontal mask will help to detect horizontal points,45 degree mask will help to detect all the points in an image which is inclined on a 45 degree same as minus 45 degree [5].

So in Line detection the approach we use is applying these four masks on the image and if we take particular i^{th} mask and j^{th} mask and we find the value computed

$$|R_i| > |R_j| \ \forall j \neq i \qquad (2)$$

This says corresponding point is more likely to be associated with line in the direction of mask *i*.

-1	-1	-1
2	2	2
-1	-1	-1

-1	2	-1
-1	2	-1
-1	2	-1

Fig. 5 Horizontal & Vertical Masks

Edge Detection
Edge detection is nothing but boundary between two regions having distinct intensity levels or gray levels. There are three steps for detecting an edge:
- Image smoothing for noise reduction
- Detection of edge points
- Edge localization

In similarity based approaches, here try to group the pixels in an image which are similar in some sense [5].

Thresholding
If we have images, and every pixel is coded with 8 bits then we can decide a threshold varying from (0 to 255)128, where pixels having an intensity value greater than 128 belongs to some region and other less than 128 belongs to other region so this is a simple thresholding operation that can be used for segmentation.

$$T = T[x, y, p(x, y), f(x, y)] \qquad (3)$$

Where (x, y) is pixel location, f(x, y) is pixel intensity at location(x, y), p(x, y) is local neighborhood property centered at (x ,y).

Thresholding can be classified with these terms as, If T is a function of f(x, y) only– global thresholding, If T is a function of both f(x, y) and local properties p(x, y) – local thresholding, If T depends on the coordinates (x, y) – dynamic/adaptive thresholding [5].

Region Growing based approach
Suppose we start with any particular pixels of an image, then we group all other pixel which is connected to this particular pixel that means all adjacent pixel of that pixel and which are similar in an intensity value. So that starts from particular pixel than try to grow the region based on connectivity or based on adjacency and similarity [5].

$$R = R_1, R_2, R_3, R_4 .... R_n \qquad (4)$$

Region R, is subdivided into sub regions such that

$$R = \bigcup_{i=1}^{n} R_i \qquad (5)$$

Also governed by following set of rules:
   a) R_i is a connected set, i=1, 2,…..n.
   b) $R_i \cap R_j$= Ø for all i and j, i ≠ j
   c) $Q(R_i)$ = True for i= 1,2,…n.
   d) $Q(R_i \cup R_j)$ = False for adjacent regions, R_i and R_j.
where $Q(R_k)$ is a logical predicate [5].

Region splitting and Merging

   Under this process, first split the image into no. of different components following some criteria and after splitting the image then we will merge the component which is adjacent and similar in some sense. So first split image into multiple smaller images and then merge the image into big one as Fig 6 [5].
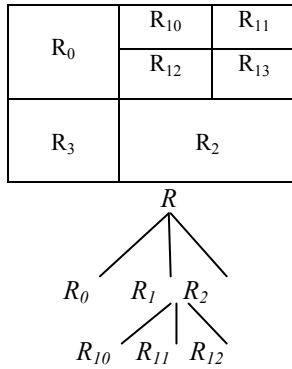

Fig. 6 Region Splitting & Merging (Quad Tree)

III.   RELATED WORK

   Kan Jiang et al. [6] design a segmentation algorithm that use scale-space filtering and watershed clustering. Firstly he extracted a sub-image of blood smear image. Then, nucleus parts of white cell from sub image are segmented by scale-space filtering and for segmenting the cytoplasm a watershed clustering is used. Finally the entire connective white cell region is obtained by using morphological operations.

   Ahmad El Allaoui et al. [7] Watershed transform is a powerful technique for rapid detection of both edges and regions. The major problem of the watershed transform is over-segmentation. Indeed, this algorithm is sensitive to any local minimum in the image, and tends to define the lines of the watershed transform where each local minimum gives rise to a region. To avoid this problem, powerful tools are adapted for different problems which have been proposed in the literature.

   Madhloom et al. [8] have given an idea about segmentation of leukocyte by focusing on nucleus. By using a combination of automatic contrast stretching the nuclei can be separated from the whole cell background by applying some image arithmetic operation with passing minimum filter and global threshold technique. Applying minimum filter prior, applying the threshold to the source image is only strength of this method. Minimum filter works as the same way as median filter, the pixel intensity is replaced with the minimum intensity value.

   A.S. Abdul Nasir et al. [9] present the approach for computer vision; color image segmentation is becoming more popular due to its good impact in most medical analysis. Its an important task to segment the white blood cell (WBC) where the leukocytes composition describes important diagnostic information of a patient. To obtain a fully segmented abnormal WBC and nucleus of acute leukemia images applied the combination between linear contrast technique and color segmentation based on HSI (Hue, Saturation, Intensity) color space. The segmentation process is eased by the unsupervised segmentation technique by k-means clustering algorithm. By using the combination of linear contrast technique and segmentation based on H component image, fully leukocyte which consists of cytoplasm and nuclei regions are eased to segment. Meanwhile, the combinations between linear contrast technique and HSI color space based segmentation have result which has a better effect on improving the accuracy of leukocytes segmentation.

   Nasrul Humaimi et al. [10] defines the method in order to perform color based segmentation on medical images by using International Commission on Illumination L*a*b* (CIELAB) color space that is able to identify the fast and error prone blood cell segmentation of both red and white blood cells.

   F. Boraytek et al. [11] proposed a segmentation technique for the images of microscopic blood which mainly contain RBC, WBC, and Parasites. This method is based on mathematical morphology operator and area granulometry is used to estimate cell area. To locate the cell center circle Radon transformation is applied to the labeled regions. Thresholding is used for the removing false objects from the background and binary image is produced as a result.

$$DBLT[t_1 \le t_2 \le t_3 \le t_4](I) = R^{\delta}_{T_{[t_1,t_4]}(I)}[T_{[t_2,t_3]}(I)] \qquad (6)$$

Where T is the threshold, R^{δ} is the morphological reconstruction operator, and [t1,t4],[t2,t3] are the wide and narrow threshold intervals respectively. The result is obtained after applying marker controlled watershed to the output from regional maxima. This method is also applied to similar blob object segmentation by adapting RBC characteristic for the new blob object.

   Mostafa Mohamed et al. [1] proposed the optimum segmentation , the optimum threshold of the image is the threshold which identify more high intensity edges and fewer low intensity edges than any other threshold. The threshold selection algorithm can be applied recursively to select additional thresholds by ignoring any edges which have already been detected by previously selected thresholds. The immediate spatial domain of each pixel is utilized by relaxation component to update both the label and the feature measurement at the original pixel value.

Dorini et al. [12] proposed an efficient method of watershed transform based on connectivity to segment the other structuring element like cytoplasm and background. To perform this operation an image forest transform is used.

N. Otsu's [13] method is a nonparametric and unsupervised method of automatic threshold selection for image. To maximize the separability of the output in gray intensity, discriminant criterion is used to select an optimal threshold. To make the steps simpler, only the zero[th] and the first-order cumulative moments of the gray-level histogram are utilized.

Seyed Hamid Rezatofighia et al. [14] method based on Gram–Schmidt orthogonalization with a snake algorithm with a capability to segment nuclei and cytoplasm of the leukocyte cells. Then, next step for segmentation was to extract the features from the segmented regions. Then most discriminative features are selected using a Sequential Forward Selection (SFS) algorithm. To evaluate this, performances of two classifiers, Artificial Neural Network (ANN) and Support Vector Machine (SVM), are compared.

The Gram–Schmidt procedure orthogonalizes a set of vectors in an inner product space, most commonly the Euclidean space (R_n). It translates a finite, linearly independent set S={v1...vn}to an orthogonal set. A projection operator for this relation is defined as:

$$ proj_u^v = \frac{\langle u,v \rangle}{\langle u,u \rangle} u = \langle u,v \rangle \frac{u}{\langle u,v \rangle} \qquad (7) $$

Here, $\langle u,v \rangle$ implies the inner product of vectors u and v. According to this, the Gram–Schmidt orthogonalization procedure is as:

$$ u_1 = v_1 \ , \qquad e_1 = \frac{u_1}{\| u_1 \|} \qquad (8) $$

$$ u_2 = v_2 - proj_{u_1}^{v_2}, \qquad e_2 = \frac{u_2}{\| u_2 \|} \qquad (9) $$

$$ u_3 = v_3 - proj_{u_1}^{v_2} - proj_{u_2}^{v_3} , e_3 = \frac{u_3}{\| u_3 \|} \qquad (10) $$

$$ w_k = v_k - \sum_{j=1}^{k-1} proj_{u_j}^{v_k} \ , \qquad e_k = \frac{u_k}{\| u_k \|} \qquad (11) $$

The sequence u_1, ..., u_k, is the required system of orthogonal vectors and the normalized vectorse1,...,e_k form an orthonormal set.
Using this method, a vector w_k can be attained for the set of Vectors S= {v_1...v_n}such that it has maximal projection on to v_k and is orthogonal to the other vectors in the set:

$$ w_k = v_k - \sum_{j=1}^{k-1} proj_{u_j}^{v_k} \qquad (12) $$

Therefore, the inner product of the set S in w_k is:

$$ \begin{cases} \langle v_j w_k \rangle = 0, j \in 1,....nandj \neq k \\ \langle v_k w_k \rangle = K, K \neq 0 \end{cases} \qquad (13) $$

In this method, each feature vector is considered as a vector in order to use a Gram–Schmidt process for segmentation of color images.

IV. IMPLEMENTATION AND RESULTS

In this section we evaluate and asses the performance of the different algorithms for Leukocytes segmentation.

A. Dataset

In our experiment, the images which are used for segmentation are from dataset have been captured with an optical laboratory microscope .All images have resolution of 2592×1944 and are taken with different magnifications of the microscope ranging (300 to 500).The ALL-IDB database has two distinct versions as ALL-IDB1and ALL-IDB2[15-19].

TABLE 1 NUMBER OF TESTING IMAGES USED IN EVALUATION

Image Dataset	No. of images
ALL-IDB1	109
ALL-IDB2	260
Total	369

B. Performance Evaluation

To evaluate the performance of algorithms, there are two types of images we are considering, one is ALL (Acute Lymphoblastic Leukemia) images of blood cell and second one is ROI selective image of a healthy person's blood cell. Original images are shown in Fig 7, and the segmented result can be shown in Fig 8 to Fig 14. There are some parameters which are using for measurement of accuracy as:

$$ accuracy = \frac{N_{automatic} \cap N_{hematic\_expert}}{\max(N_{auttomatic}, N_{hematic\_expert})} *100 \qquad (14) $$

$N_{automatic}$ = No. of images accurately segmented by algorithms
$N_{hematic\_expert}$ = No. of images manually segmented by experts
Performance can be shown in Fig 15.

TABLE 2 EVALUATIONS OF DIFFERENT SEGMENTATION ALGORITHMS

Segmentation Method	Accuracy for ALL Infected blood smear image (%)	Accuracy for Healthy blood smear image (%)	Average Accuracy (%)
Global Threshold Method[13]	73.39	73.07	73.23
Marker controlled Watershed	64.22	63.46	63.84
L*a*b* Color Space	79.82	80.76	80.29
Madhloom [8]	77.98	76.92	77.45
K-Means Clustering	82.56	84.62	83.59
HSV Color Space	70.64	71.15	70.89

C. *Experimental Results:*
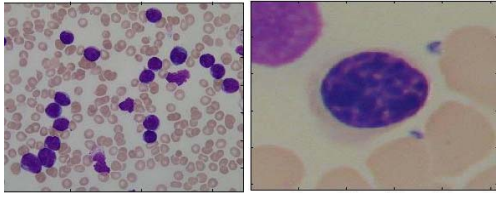
a)   ALL (Acute Lymphoblastic   b)   Healthy blood cell
     Leukemia) infected blood cell

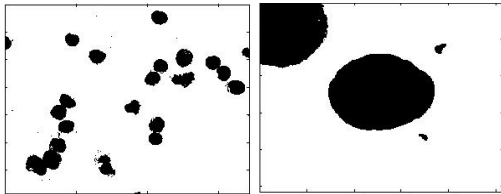Fig. 7 Giesma stained color Images (Image Courtesy [19])

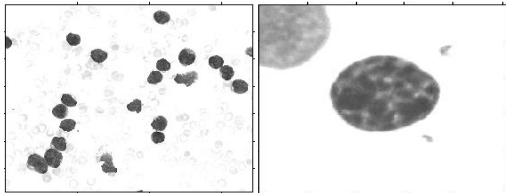Fig. 8 Segmented by Global Threshold Method [13]

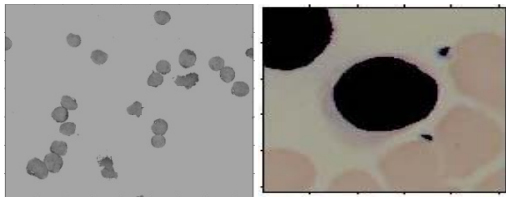Fig. 9 Segmented by Marker controlled Watershed Method

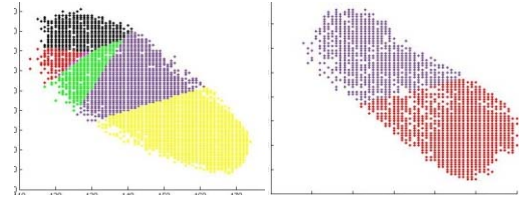Fig. 10 Color-Based Segmentation Using the L*a*b* Color Space

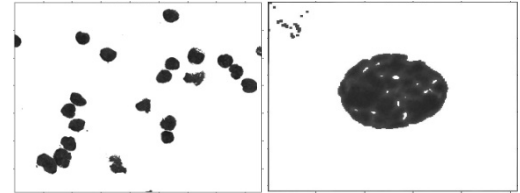Fig. 11 Scatterplot of segmented pixel in 'a*b' space(L*a*b)

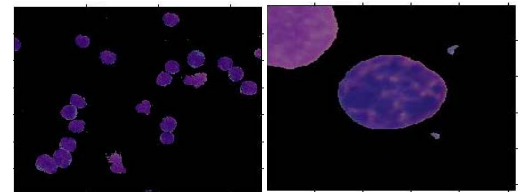Fig. 12 Segmentation by Madhloom et al. Method [8]

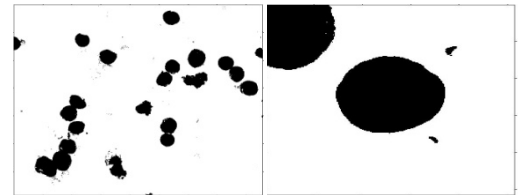Fig. 13 Segmented by Color-Based Segmentation Using K-Means Clustering

Fig. 14 Segmented by Color-Based Segmentation Using HSV Color Space
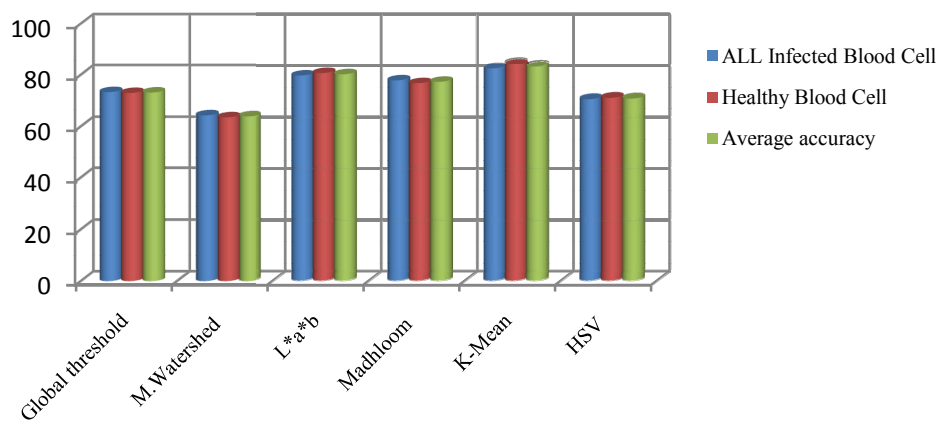

Fig. 15 Performance Comparison of Different Segmentation Methods for both types of Images.

V. CONCLUSION

In this work we have analyzed the performance of different existing segmentation algorithms for leukocytes nuclei segmentation of microscopic blood smear. Different types of Acute Lymphoblastic Leukemia Images (ALL) have been successfully isolated and extracted the nucleus part from other structure of the leukocyte cell. Result has been examined and statistically approved by the hematologist and also compared manually according to ratio of correct pixels of the nuclei in cell.The highest accuracy for nuclei segmentation is given by K-mean clustering for ALL Infected blood smear is 82.56% and 84.62% for healthy blood smear, and the lowest is 64.22% by Marker controlled watershed for ALL infected blood cell and 63.84% for healthy blood vessel.

REFERENCE

[1] Mostafa Mohamed, Behrouz Far, "*An Enhanced Threshold Based Technique for White Blood Cells Nuclei Automatic Segmentation*" Published in 2012 IEEE 14th International Conference on e-Health Networking, Applications and Services (Healthcom) 978-1-4577-2040-6/12/2012 IEEE.

[2] Navin D. Jambhekar, "*Red Blood Cells Classification using Image Processing*" Department of Computer Science published in Science Research Reporter 1(3): 151-154, Nov. 2011 ISSN: 2249-7846.

[3] Barbara J. Bain *"A beginner's guide to blood cells,"* 2nd Edition, Department of Hematology ,St Mary's Hospital Campus, Imperial College, London, 2004 ISBN 1-4051-2175-0.

[4] J. M. Sharif, M. F. Miswan, M. A. Ngadi, Md Sah Hj Salam, Muhammad Mahadi bin Abdul Jamil, "*Red Blood Cell Segmentation Using Masking and Watershed Algorithm: A Preliminary Study*", Published in 2012 International Conference on Biomedical Engineering Penang, Malaysia,27-28 February 2012, 978-4577-1991-2/12/2011 IEEE.

[5] R.G. Gonzalez, R.E. Woodsand S.L. Eddins, *"Digital Image Using MATLAB Processing"* , Pearson Education, Inc. NJ, chap. 9, pp. 334-377, chap. 10, pp. 378-425, 2004.

[6] Kan Jiang, Qih'gmln Liao, Sheng-Yang Dai, "*A Novel White Blood Cell Segmentation Scheme Using Scale-Space Filtering And Watershed Clustering* " Published in Proceedings of the Second International Conference on Machine Learning and Cybernetics, xi, 2-5 November 2003.

[7] Ahmad El Allaoui, M'barek Nasri, "*Medical Image Segmentation By Marker Controlled Watershed And Mathematical Morphology*" published in The International Journal of Multimedia & Its Applications (IJMA) Vol.4, No.3, June 2012.

[8] H.T. Madhloom, 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 Arithmetic and Automated Threshold"*, Journal of Applied Science, Vol.10 No. 11,pp.959-966,2010.

[9] A.S.Abdul Nasir, M.Y.Mashor, H.Rosline "*Unsupervised Colour Segmentation of White Blood Cell for Acute Leukaemia Images*" Published in 978-1-61284-896-9/11/2011 IEEE.

[10] Nasrul Humaimi Mahmood, Poon Che Lim, Siti Madihah Mazalan, and Mohd Azhar Abdul Razak, "*Blood Cells Extraction Using Color Based Segmentation Technique*", published in Int. J. LifeSc. Bt & Pharm. Res. 2013 ISSN 2250-3137 Vol. 2, No. 2, April 2013© 2013 IJLBPR.

[11] F. BorayTek, Andrew G. Dempster and Izzet Kale, "*Blood Cell Segmentation Using Minimum Area Watershed And Circle Radon Transformations*" University of Westminster Applied DSP and VLSI Research Group published in C. Ronse et al.(eds.), Mathematical Morphology : 40 Years On, 441–454.

[12] Dorini, L.B., Minetto, R., Leite, N.J., *"White blood cell segmentation using morphological operators and scale-space analysis"* In: Proceedings of the XXth Brazilian Symposium on Computer Graphics and Image Processing , pp.294–304, 2007.

[13] N. Otsu, *"A threshold selection method From gray-level histograms"* transactions on systems, man, and cybernetics, vol. Smc-9, no. 1, January 1979, IEEE, pp 62-66.

[14] Seyed Hamid Rezatofighia, Hamid Soltanian-Zadeh*,"Automatic recognition of five types of white blood cells in peripheral blood"* published in Computerized Medical Imaging and Graphics35, 333–343,Control and Intelligent Processing Center of Excellence, Elsevier(2011).

[15] R. Donida Labati, V. Piuri, and F. Scotti*, "All-IDB website,"* University of Milan, Department of Information Technologies, Online available at http://www.dti.unimi.it/fscotti/all.

[16] V. Piuri, F. Scotti, "*Morphological classification of blood leucocytes by microscope images*", in Proc. of the IEEE International Conference on Computational Intelligence for Measurement Systems and Applications ,Boston, MA, USA, pp. 103 - 108, July, 2004.

[17] F. Scotti, "*Automatic morphological analysis for acute leukemia identification in peripheral blood microscope images*", in Proc. of the IEEE International Conference on Computational Intelligence for Measurement Systems and Applications, Giardini Naxos - Taormina, Italy, pp. 96 - 101, July, 2005.

[18] F. Scotti, "*Robust Segmentation and Measurements Techniques of White Cells in Blood Microscope Images*", in Proc. of the IEEE Instrumentation and Measurement Technology Conference (IMTC 2006), Sorrento, Italy, pp. 43 -48, April, 2006.

[19] R. Donida Labati, V. Piuri, F. Scotti, "*ALL-IDB: the Acute Lymphoblastic Leukemia Image Database for Image Processing*", in Proc. of the IEEE International Conference on Image Processing (ICIP 2011), Brussels, Belgium, pp. 2045 - 2048, September 11 - 14, 2011.

[20] Sanaullah Khan, Aamir Khan, Faisal Saleh Khattak and Arslan Naseem, "*An Accurate and Cost Effective Approach to Blood Cell Count*" Published in International Journal of Computer Applications (0975 – 8887) Volume 50 – No.1, July 2012.

[21] Ralf Kohler , "*A Segmentation System Based on Thresholding*" Published in Computer Graphics Image Processing 15, 319-338 (1981), 0146-664X/81 / 040319-20502.00/0.