Original Article

Nucleus and cytoplasm segmentation in microscopic images using K-means clustering and region growing

Omid Sarrafzadeh, Alireza Mehri Dehnavi

Department of Biomedical Engineering, Faculty of Advanced Medical Technology, Isfahan University of Medical Sciences, Isfahan, Iran

Abstract

Background: Segmentation of leukocytes acts as the foundation for all automated image-based hematological disease recognition systems. Most of the time, hematologists are interested in evaluation of white blood cells only. Digital image processing techniques can help them in their analysis and diagnosis.

Materials and Methods: The main objective of this paper is to detect leukocytes from a blood smear microscopic image and segment them into their two dominant elements, nucleus and cytoplasm. The segmentation is conducted using two stages of applying K-means clustering. First, the nuclei are segmented using K-means clustering. Then, a proposed method based on region growing is applied to separate the connected nuclei. Next, the nuclei are subtracted from the original image. Finally, the cytoplasm is segmented using the second stage of K-means clustering.

Results: The results indicate that the proposed method is able to extract the nucleus and cytoplasm regions accurately and works well even though there is no significant contrast between the components in the image.

Conclusions: In this paper, a method based on K-means clustering and region growing is proposed in order to detect leukocytes from a blood smear microscopic image and segment its components, the nucleus and the cytoplasm. As region growing step of the algorithm relies on the information of edges, it will not able to separate the connected nuclei more accurately in poor edges and it requires at least a weak edge to exist between the nuclei. The nucleus and cytoplasm segments of a leukocyte can be used for feature extraction and classification which leads to automated leukemia detection.

Key Words: Blood smear microscopic image, K-means clustering, leukocyte (WBC) segmentation, region growing

Address for correspondence:
Assoc. Prof. Alireza Mehri Dehnavi, Department of Biomedical Engineering, Faculty of Advanced Medical Technology, Isfahan University of Medical Sciences, Hezar Jereeb Street, Isfahan, Iran. E-mail: mehri@med.mui.ac.ir
Received: 15.04.2013, **Accepted:** 10.01.2015

INTRODUCTION

Blood consists of three types of cells and cell fragments floating in a liquid called plasma. These cellular components are:
Red blood cells (RBCs; erythrocytes) – oxygen-carrying cells White blood cells (WBCs; leukocytes) – cells that help make up the body's immune system Platelets (thrombocytes) – fragments of cells that play an important role in formation of blood clots There are five types of leukocytes found in the blood: Basophil, eosinophil, lymphocyte, monocyte, and neutrophil. Each cell type has a specific role to play in our body's immune system.[1] WBCs play a significant role in the diagnosis of different diseases such as leukemia and

Access this article online	
Quick Response Code:	Website: www.advbiores.net
	DOI: 10.4103/2277-9175.163998

How to cite this article: Sarrafzadeh O, Dehnavi AM. Nucleus and cytoplasm segmentation in microscopic images using K-means clustering and region growing. Adv Biomed Res 2015;4:174.

different types of infections,[2] so extracting information from them is valuable for hematologists. Leukemia is a group of diseases characterized by increased numbers of white cells in the blood and bone marrow.[1] The count and shape, lineage, and maturity level of white and red blood cells could aid in the diagnosis of diseases such as leukemia.[3] Acquiring important information by peripheral blood cell differential counting and, therefore, the microscopic review for correct patient diagnosis is an exhaustive effort and requires extremely trained or qualified experts or professionals. Automatic detection of WBCs in microscopic images as a substitute for manually locating, identifying, and counting different classes of cells is an important topic in the domain of cancer diagnosis.[4] Majority of the segmentation methods of WBCs in microscopic images that have been proposed are generally based on edge and border detection, region growing, filtering, mathematical morphology, and watershed clustering. Ritter *et al*.[5] presented a fully automatic method for segmentation and border identification of all objects that do not overlap the boundary in an image taken from a peripheral blood smear slide. Liao and Deng[6] introduced a gray level threshold based method to segment WBCs in microscopic images. Segmentation of leukocytes based on histogram analysis and measurement of distance among nuclei was done by Hamghalam *et al*.[7] The application of morphological operators has also been investigated for WBC background separation. Leyza *et al*.[8] used morphological operators and examined the scale-space properties of toggle operator to improve segmentation accuracy. A 3D structuring element based multi-scale morphology method has been proposed by Fatichah *et al*.[9] to improve the accuracy of WBC segmentation in bone marrow microscopic images. They showed that their proposed method can segment nucleus more accurately than the conventional mathematical morphology. Ongun *et al*.[10] did segmentation by morphological preprocessing followed by the snake-balloon algorithm. G′omez *et al*.[11] introduced an automatic seeded region growing algorithm called ASRG-IB1 that performs the segmentation of color (RGB) and multispectral images. However, determining the initial seed points is a challenging task in all region-based methods. Color images are very rich source of information and regions can be segmented better in terms of color as compared to grayscale images. However, selection of color space is also a vital issue in color-based clustering. Jiang *et al*.[12] proposed a WBC segmentation scheme on color space images using feature space clustering techniques, scale-space filtering for nucleus extraction, and watershed clustering for cytoplasm extraction. A two-step color image segmentation process using K-means clustering followed by EM algorithm was proposed by Sinha *et al*.[13] to segment nucleus and cytoplasm. Applying mean shift algorithm for color image segmentation of leukocyte images was presented by Comaniciu *et al*.[14] Blood cell contour detection using active contour model was first presented by Kass *et al*.[15] Thereafter, some active contour approaches for WBC segmentation have been explored.[16-18] Segmentation of leukocyte using contour propagation with distance map guiding was introduced by Srijad *et al*.[19] Their results show that active contour guided by distance mapping from a neighboring area, is able to extract nucleus and cytoplasm regions. Contour detection algorithms rely on the discontinuity of image intensities or texture at the object boundaries and are sensitive to noisy images. Some fuzzy-based algorithms are introduced by researchers. Fuzzy divergence is employed by Ghosh *et al*.[20] for threshold estimation in leukocyte images. A fuzzy-based two-stage color segmentation strategy was employed by Mohapatra *et al*.[21] for segregating leukocytes from other blood components. An unsupervised leukocyte segmentation using rough fuzzy clustering was proposed by Mohapatra *et al*.[22] They showed that the hybrid rough fuzzy c-means algorithm is robust in segmenting stained blood microscopic images. Kumar *et al*.[23] used teager energy operator for segmentation nucleus based on the edges, which are detected effectively by teager energy operator, but it requires at least a weak edge to exist between RBCs and background. For cytoplasm segmenting, they used a simple morphological method. Cseke[24] introduced multi-step segmentation scheme which implements the automatic thresholding method proposed by Otsu.[25] Since many beneficial explorations have been carried out for WBC segmentations, but majority of these methods have some defects to different extent, such as complexity of arithmetic, difficulty to ensure parameters, and so on. In this paper, we propose a new simple method which can be adapted to detect WBCs in a microscopic image and segment their components, nucleus, and cytoplasm. We also proposed a simple algorithm to separate connected nuclei based on region growing and utilizing edge information.

The outline of the paper is as follows. The next section deals with the methodology and presents the proposed algorithm. Results are discussed in the third section. The last section concludes.

MATERIALS AND METHODS

First, the nuclei are extracted from microscopic image and in the next step, the cytoplasm is segmented.

Nucleus segmentation

The algorithm for nuclei segmentation is briefly illustrated in Figure 1. The input microscopic image stores in three red, green, and blue (RGB) channels.

One example of microscopic image of blood cells with its RGB channels is shown in Figure 2. The median filter is applied on each R, G, and B band separately to reduce noise and also not to lose edges. Since the red, green, and blue components are highly correlated, it is difficult to execute some image processing algorithms. So, the image is converted from RGB color space to L*a*b* color space. The L*a*b* color space (also known as CIE L*a*b* or CIELAB) is colorimetric, perceptually uniform, and device independent. The L*a*b* color space is an excellent decoupler of intensity (represented by lightness L*) and color (represented by a* for red minus green and b* for green minus blue).[26] The L*a*b* image of Figure 2 with its subspaces is shown in Figure 3. We used K-means clustering in order to detect nuclei which is similar to the method proposed in.[27] The colors in a* and b* subspaces are classified using K-means clustering to cluster the objects into three clusters using Euclidean distance metric. Since the nuclei are darker than the other parts in microscopic images of blood cells, we chose three clusters for K-means cluster. By this choice, one of the clusters will contain nuclei and, maybe, platelets and the other two clusters will contain cytoplasm, RBCs, and background. Figure 4 shows the result of K-means cluster for Figure 1a. For each cluster, the mean of three RGB bands is calculated and the cluster with minimum mean is considered as nucleus. Some morphological operations are performed on the nucleus cluster to eliminate probable false segments. First, the morphological open operation (an erosion followed by a dilation) is done to smoothen the contour of nuclei, break narrow isthmuses, and eliminate thin protrusions. For opening operation, a flat, disk-shaped structuring element with radius R is used. The parameter R is specified adaptively based on the *r*esolution of acquired image and *o*bjective *z*oom (*r and oz*) of the microscope device. After opening, the morphological hole filling operation is applied to fill inside of the segmented regions. Usually, in the nucleus cluster, the platelets exist. We know that the size of the platelets is very smaller than that of nuclei. So, in order to remove the platelet segments, the regions which have area lower than A_p (area of platelet) are disregarded. A_p is specified adaptively based on the (*r and oz*). The result of applying K-means clustering and morphological operations for Figure 1a is presented in Figure 5. As it is observed in Figure 5d, there are some nuclei segments which are connected to each other after segmentation. We proposed an algorithm to separate the connected nuclei based on the region growing and using information of edges. We use the fact that there is often an edge between the connected nuclei. The best color channel to extract the edge of nuclei is green channel as seen in the Figure 2c. In order to make the algorithm faster, the connected nuclei are first specified based on their area higher than a predefined area A_{WBC} (area of a single WBC). A_{WBC} is about one and a half times the area of a single WBC and is specified based on the (*r and oz*). Figure 6a-c show the binary images of the connected nuclei, green band of the connected nuclei, and edges of connected nuclei using Sobel method,[26] respectively. After detection of edges, the edges are superimposed on the binary image of the connected nuclei, as shown in Figure 6d. We utilize region growing technique in

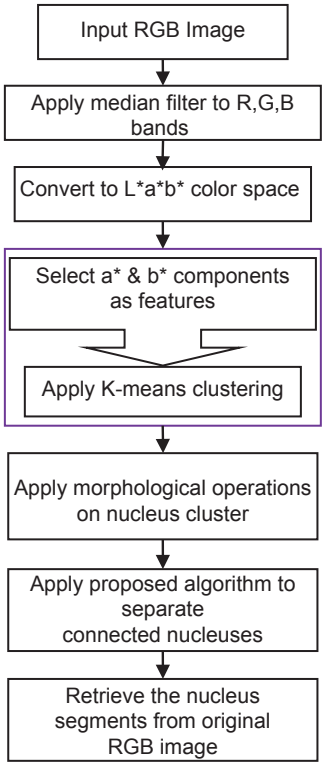

Figure 1: Steps to segment nucleuses

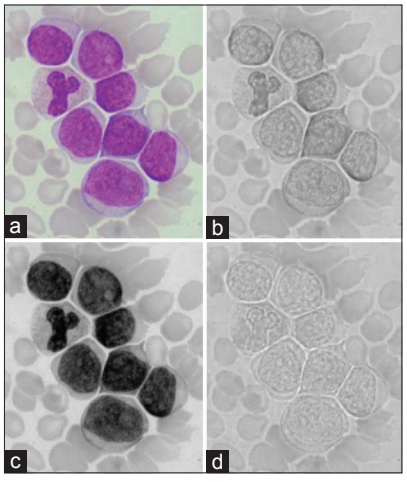

Figure 2: (a) Original RGB image, (b) red band, (c) green band, and (d) blue band

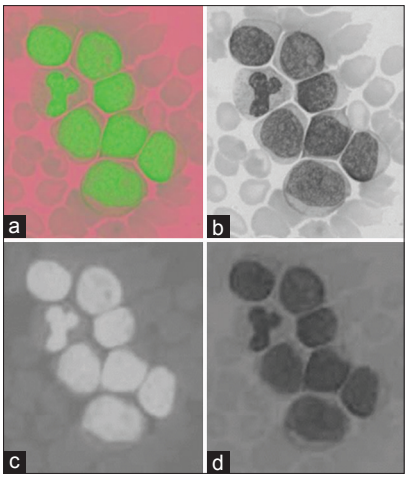

Figure 3: (a) L*a*b* color image, (b) L* band, (c) a* band, and (d) b* band

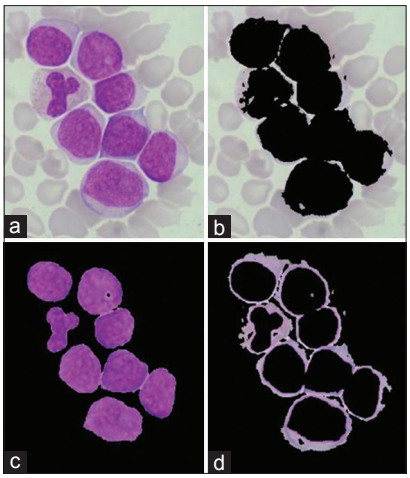

Figure 4: (a) Original RGB image, (b) cluster 1, (c) cluster 2, and (d) cluster 3

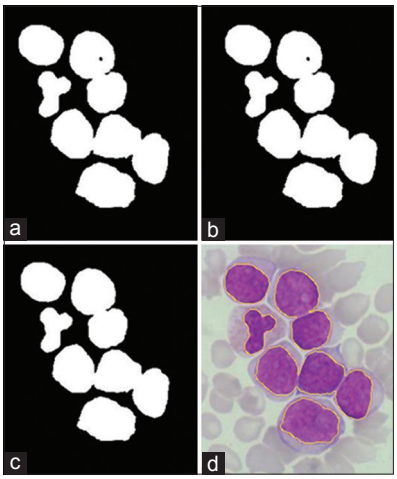

Figure 5: (a) Binary image of nuclei of K-means cluster, (b) opening operation on (a), (c) filling holes of (b), and (d) nucleus segments after applying K-means clustering and morphological operations

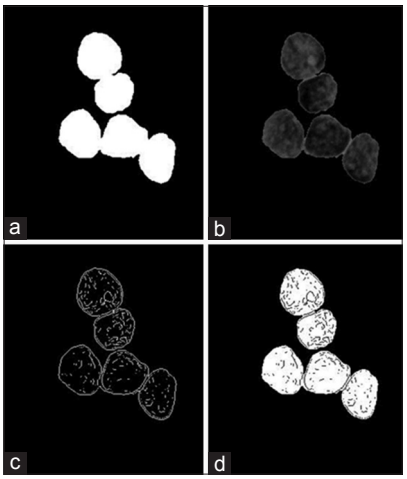

Figure 6: (a) Binary image of the connected nuclei, (b) green band of the connected nuclei, (c) edges of the connected nuclei using Sobel method, and (d) superimposed edges on the binary image of the connected nuclei

binary mode [Figure 6d] to separate the connected nuclei. The algorithm starts from a seed point inside each nucleus and grow region until no more pixels satisfy the criteria for inclusion in that region. For a central pixel (red pixel in Figure 7), the algorithm takes an 8-connected neighborhood for decision (green square in Figure 7). The black pixels in the region stand for edges [Figure 7] and if two or more pixels of eight neighbors are black, the central pixel assigned to the region and algorithm stops growing. The problem now is how to determine seed points. It is achieved by applying erosion operation with a flat, disk-shaped structuring element with a large enough radius R_1 (large radius) which is determined based on the (*r and oz*), on the binary image of the connected nuclei. Figure 8a shows the binary image of the connected nuclei and Figure 8b shows the outcome of erosion operation on Figure 8a for determining the seed points. Seed points are specified by calculating the center of mass of the regions in Figure 8b. The result of nucleus segmentation of Figure 9a after applying the proposed algorithm to separate the connected nuclei is shown in Figure 9b. Subsequently, cytoplasm segmentation of leukocytes is explained.

Cytoplasm segmentation

First, the nuclei segmented in the previous stage are subtracted from the original image. Figure 10a shows the original image without the nuclei. Then, the subtracted image is converted to L*a*b* color space. The colors in a* and b* subspaces are classified using K-means clustering into three clusters using Euclidean distance metric. Since the nuclei are omitted, one of the clusters contains cytoplasm and the other two clusters contain RBCs and background. Each cluster is retrieved from the original image and the mean of red band for each

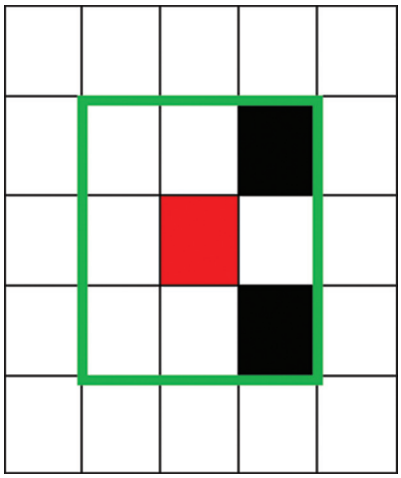

Figure 7: Central pixel (red pixel) and its 8-connected neighbor (green square) for decision in region growing algorithm

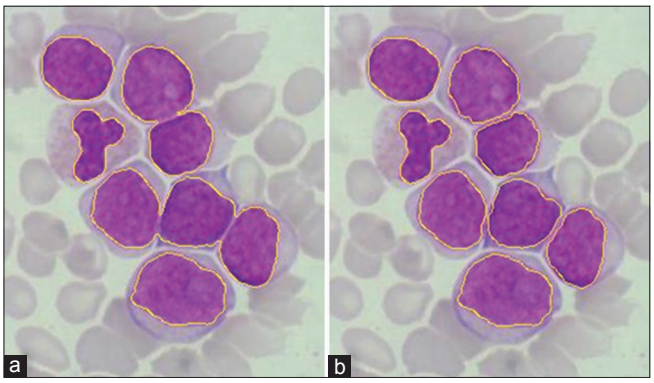

Figure 9: (a) Nucleus segments after K-means and (b) separation of the connected nuclei using the proposed method

cluster is calculated. The cluster with minimum mean of red band is considered as the cytoplasm cluster. Figure 10b shows the cytoplasm cluster obtained by K-means clustering for Figure 2a. Figure 10c shows the binary image of the cytoplasm and by doing some morphological operations, the cytoplasm is modified. The final segmentation of nucleus and cytoplasm of Figure 2a is shown in Figure 10d.

RESULTS AND DISCUSSION

Figure 11 shows some blood smear microscopic images of leukemia. Figure 11a shows the original RGB image of blood cells, Figure 11b shows the nucleus segments after applying K-means clustering, Figure 11c shows the result of applying proposed method to separate the connected nuclei, and Figure 11d shows the final segmentation of nuclei (white) and cytoplasm (gray). Top image in Figure 11 is a type of leukemia, acute myelogenous leukemia type 2 [AML with maturation (AML-M2)]. As it is seen, despite the low contrast between the cytoplasm, RBCs,

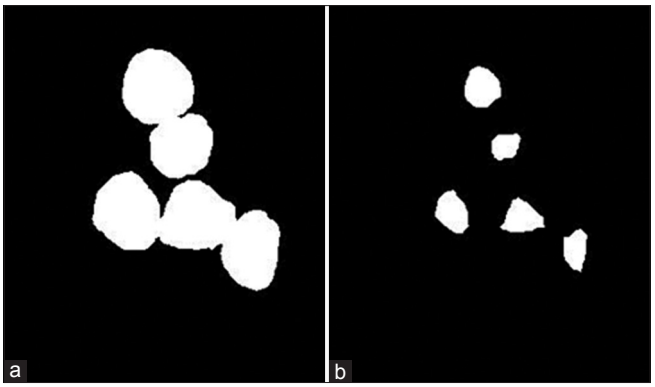

Figure 8: (a) Binary image of the connected nuclei and (b) result of erosion of (a)

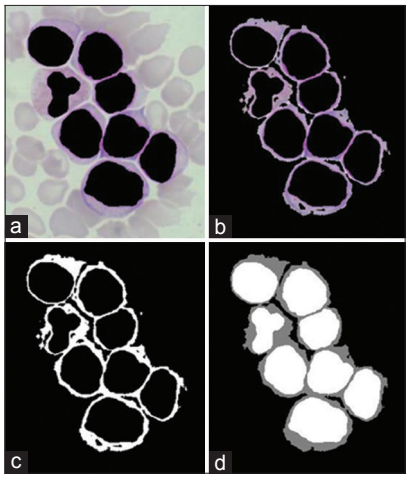

Figure 10: (a) Original image without nuclei, (b) cytoplasm cluster after applying K-means clustering on (a), (c) binary image of cytoplasm, and (d) nuclei (white) and cytoplasm (gray) of Figure 2a

and background, the proposed method segments the cytoplasm and nuclei accurately. Also, the proposed method separates the connected nuclei as well as possible. Bottom image in Figure 11 is a type of myeloblast leukemia, and it is observed from the image that nuclei and cytoplasm are detected promisingly and the connected nuclei are made discrete by the proposed algorithm. For the proposed method to work precisely, it required at least a weak edge to exist between the connected nucleuses.

CONCLUSIONS

In this paper, we proposed a method based on K-means clustering and region growing in order to detect leukocytes from a blood smear microscopic image and segment its components; nucleus and cytoplasm. It is shown that the proposed method works well even though there is no significant contrast between the components in the image. As a step of the algorithm relies on the information of edges, it will not able to separate the connected

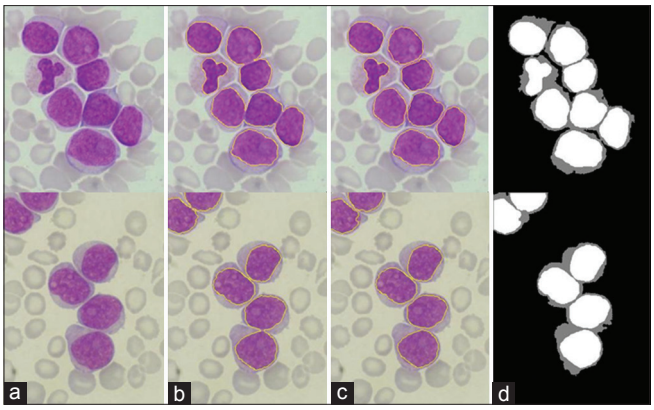

Figure 11: (a) Original RGB image, (b) nucleus segments after applying K-means clustering, (c) separation of connected nuclei using the proposed method, and (d) nuclei (white) and cytoplasm (gray) segments

nuclei more accurately, but it will be able to segment the nuclei exactly.

REFERENCES

1. Hoffbrand AV, Moss PAH, Pettit JE, editors. Essential Haematology. 5th ed. Oxford: Blackwell Publishing; 2006.
2. Guyton C, Hall JE. Textbook of Medical Physiology. 11th ed. Philadelphia, Pennsylvania: Elsevier Inc.; 2006.
3. Markiewicz T, Moszczynski. Analysis of Features for Blood Cell Recognition", VI International Workshop Computational Problems of Electrical Engineering. Poland. 2004.
4. Fatichah C, Tangel ML, Widyanto MR, Dong F, Hirota K. Parameter optimization of local fuzzy patterns based on fuzzy contrast measure for white blood cell texture feature extraction. J Adv Comput Intell Intell Inf 2012;16:412-9.
5. Ritter N, Cooper J. Segmentation and border identification of cells in images of peripheral blood smear slides. In Proceedings of the Thirtieth Australasian Conference on Computer Science 2007;62:161-9.
6. Liao Q, Deng Y. An accurate segmentation method for white blood cell images. Proceedings of IEEE International Symposium on Biomed Imaging; 2002. p. 245-58.
7. Hamghalam M, Ayatollahi A. Automatic Counting of Leukocytes in Giemsa-Stained Images of Peripheral Blood smear. IEEE International Conference on Digital Image Processing; 2009. p. 13-6.
8. Dorini LB, Minetto R, Leite NJ. White blood cell segmentation using morphological operators and scale-space analysis. In SIBGRAPI '07: Proceedings of the XX Brazilian Symposium on Computer Graphics and Image Processing, 2007. p. 294-304.
9. Fatichah C. 3D Structuring Element based Multiscale Morphology for Bone Marrow White Blood Cell Segmentation. Tokyo, Japan: International Symposium on Intelligent Systems (iFAN2010); 2010.
10. Ongun SG. Feature extraction and classification of blood cells for an automated differential blood count system: Neural Networks, Proceedings. IJCNN '01. International Joint Conference on, Vol. 4: 2461-6, 2001.
11. G'omez O, Gonz'alez JA, Morales EF. Image Segmentation Using Automatic Seeded Region Growing and Instance-Based Learning. Proceeding of 12th Ibero American Congress on Pattern Recognition, CIARP 2007.
12. Jiang K, Liao QM, Dai SY. A novel white blood cell segmentation scheme using scale-space filtering and watershed clustering. Machine Learning and Cybernetics, International Conference on, Vol. 5. 2820-5, 2003.
13. Sinha, Ramakrishnan A. Automation of differential blood count. in Proceedings of the Conference on Convergent Technologies for Asia-Pacific Region, pp. 547-51, 547–551, 2003.
14. Comaniciu D, Meer P. Cell Image segmentation for diagnostic pathology. New York, NY, USA: Springer; 2001.
15. Kass M, Witkins A, Terzopoulos D. Snakes: Active contour models. in Proceedings of the 1st International Conference on Computer Vision, 1987. p. 259-68.
16. Yang L, Meer P, Foran DJ. Unsupervised segmentation based on robust estimation and color active contour models. IEEE Trans Inf Technol Biomed. 2005;9:475-86.
17. Eom S, Kim S, Shin V, Ahn B. Leukocyte Segmentation in Blood Smear Images Using Region-Based Active Contours. LNCS 4179. Berlin, Heidelberg: Springer-Verlag; 2006. p. 867-76.
18. Zamani F, Safabakhsh R. An unsupervised GVF Snake approach for white blood cell segmentation based on nucleus. ICSP 2006 Proceedings, 2006.
19. Srijad C. Segmentation of leukocyte using contour propagation with distance map guiding. The 3rd International Symposium on Biomedical Engineering (ISBME 2008), 2008.
20. Ghosh M, Das D, Chakraborty C, Ray AK. Automated leukocyte recognition using fuzzy divergence. Micron 2010;41:840-6.
21. Mohapatra S, Samanta SS, Patra D, Satpathi S. Fuzzy based blood image segmentation for automated leukemia detection. Devices and Communications (ICDeCom), 2011 International Conference on, p. 1-5, 2011.
22. Mohapatra S, Patra D, Kumar K. Unsupervised leukocyte image segmentation using rough fuzzy clustering. International Scholarly Research Network (ISRN) Artificial Intelligence. Vol 2012, Article ID 923946, 12 pages, 2012.
23. Kumar BR, Joseph DK, Sreenivasc TV. Teager energy based blood cell segmentation. Digital Signal Processing. DSP 2002 14th International Conference on, vol. 2. 2002. p. 619-22.
24. Cseke I. A fast segmentation scheme for white blood cell images. In: Pattern Recognition. Conference C: Image, Speech and Signal Analysis, Proceedings. 11th IAPR International Conference on, Vol. 3. 530-3. 1992.
25. Otsu N. A threshold selection method from gray-level histograms. Automatica 1975;11:285-96.
26. Gonzalez RC, Richard E. Woods. Digital Image Processing. 3rd ed. United States: Prentice Hall; 2007.
27. Chitade AZ, Katiyar SK. Color based image segmentation using K-means clustering. Int J Eng Sci Technol 2010;2:5319-25.

Source of Support: This work is supported by Department of Biomedical Engineering, Faculty of Advanced Medical Technology, Isfahan University of Medical Sciences, Isfahan, Iran, **Conflict of Interest:** The authors declare that they have no conflict of interest.