An efficient technique for white blood cells nuclei automatic segmentation

Conference Paper · October 2012

DOI: 10.1109/ICSMC.2012.6377703

CITATIONS
56

READS
1,853

3 authors, including:

Mostafa Mohamed
The University of Calgary
18 PUBLICATIONS **172** CITATIONS
SEE PROFILE

Amr Guaily
Nile University
42 PUBLICATIONS **168** CITATIONS
SEE PROFILE

2012 IEEE International Conference on Systems, Man, and Cybernetics
October 14-17, 2012, COEX, Seoul, Korea

An Efficient Technique for White Blood Cells Nuclei Automatic Segmentation

Mostafa Mohamed, Behrouz Far
Department of Electrical and Computer Engineering
University of Calgary
Calgary, Canada
{mabdelal, far}@ucalgary.ca

Amr Guaily
Engineering Mathematics and Physics Department
Cairo University
Cairo, Egypt

***Abstract*—Blood tests are of the most important and often requested clinical examinations. Manual microscopic assessment is a must do when a blood sample is suspicious of abnormality. This manual process is tedious, time consuming and subjective. Automating microscopic blood classification is desirable to help the pathologists to speed-up and enhance the results accuracy. Segmentation is the first and most important step in automatic blood cell classification. In this paper, we present an effective technique for automatic blood cell nuclei segmentation. The technique is based on gray scale contrast enhancement and filtering. Minimum segment size is implemented to remove false objects. The technique is tested on 365 blood images. The segmentation performance is quantitatively evaluated on the test set to be 79.7%. This performance is high compared to other published algorithm executed on the same dataset. Evaluation is done on each of the five normal white blood cell types to compare separate performance. The lowest segmentation accuracy is for Eosinophil with 69.3% and the highest is Monocyte with 86.3%. The MATLAB source code and the blood images dataset are published on MATLAB file exchange website for comparison and re-production.**

Keywords- Blood cell; white blood cells; WBC; Segmentation; Leucocyte; MATLAB; Code; Dataset

I. INTRODUCTION

Blood tests are of high importance for diagnosis of many diseases and also to investigate functions of body organs such as kidney, liver, thyroid, and heart. Examples of the diseases that blood tests can investigate are: cancer, HIV/AIDS, diabetes, anemia, and coronary heart disease [1] [2]. Manual microscopic examination is a must when there is a suspicion of abnormality in the blood sample. It is tedious, time consuming, and subjective. Automating the visual sample inspection is needed to help the pathologists increasing productivity and reducing costs. The automation process includes image acquisition, image processing, segmentation, feature extraction, and classification. Segmentation is considered the most important and critical step in the process as it affects the rest of the following steps [3].

This paper is focusing on the segmentation step. We propose an efficient technique for white Blood cells (WBC) nuclei automatic segmentation. In this research, the algorithm proposed by Madhloom et al. [4] is modified to account for more general situations. The proposed modification is to reduce dependence on the image initial contrast. This contrast dependence leads to the capturing of all objects that have the same gray-level as of the WBCs. To overcome this disadvantage we propose to use some constraints to eliminate the false objects.

The remainder of this paper is organized as follows. In Section 2, literature review is presented. In section 3, the proposed segmentation algorithm is explained. In Section 4, results obtained by the proposed framework are presented, and finally, conclusions are drawn in Section 5.

II. LITERATURE REVIEW

The ultimate goal of blood cell segmentation is to extract the cells from complicated background and to segment every cell into morphological components such as nucleus, cytoplasm, and some others. There is no universal algorithm for segmentation of every medical image. Ongun et al. [5] proposed an algorithm which segment the WBCs using active contour models (snakes and balloons). Shape based and texture based features are utilized for the classification task. In their study they consider twelve classes of WBCs. Adollah et al. [6]present a comprehensive survey about segmentation methods. The main objective of their study is to develop an automated system on blood cell classification. Theerapattanakul et al. [7] propose a method of segmentation of the white blood cell by utilizing the active contour. They start by double thresholding then they scan the binary image searching for the nucleus of individual white blood cell of which the intensity exceeding the thresholding value. An initial circular shape (snake) is placed on the nucleus locations found. Finally, active contour model is used with gradient flow vector force as a force to drive the snake contour fitting the white blood cell to be extracted. F. Sadeghian et al. [8] demonstrate a framework for segmenting white blood cells. They segment the WBC to its two dominant elements: nucleus and cytoplasm. The proposed scheme has two parts: The nucleus segmentation part is based on morphological analysis, and the cytoplasm segmentation is based on pixel-intensity thresholding. The results show that the proposed method is able to yield 92% accuracy for nucleus segmentation and 78% for cytoplasm segmentation. The major limitation of their framework is that the framework has been done on sub-images to have easier implementation. Madhloom et al. [4] propose an algorithm to automate the process of detection and classification of leukocytes. Specifically, white blood cells are recognized and classified into various distinct subtypes. Markiewicz et al. [9] present an automatic system for blood cell recognition on the basis of the bone marrow images. They apply the

978-1-4673-1714-6/12/$31.00 ©2012 IEEE

morphological preprocessing of the image for individual cell extraction, generation and selection of the diagnostic features and the recognition system using Gaussian kernel Support Vector Machine (SVM). They claim to reach accuracy of 87%. Theera-Umpon [10]developed an automatic segmentation technique for microscopic bone marrow white blood cell images. They apply the fuzzy C-means (FCM) algorithm and the mathematical morphology to segment white blood cells. The FCM algorithm is applied to overly segment each cell image to form patches. Cell and nucleus smoothing and small patch removal are done by using the binary morphological operations. Angulo and Flandrin [11] present a technique to automatically detect the working area of peripheral blood smears stained with May- Grünwuald Giemsa. The optimal area is defined as the well spread part of the smear. The algorithm consists of two stages. First, an image analysis using mathematical morphology is applied for extracting the erythrocytes. Second, the number of connected components from the three kinds of particles is counted and the coefficient of spreading and the coefficient of overlapping are calculated. Cao et al. [12] present an algorithm that focuses on the detection of red blood cells in urine image. After the urine image is preprocessed by an improved Sobel operator, red blood cells are localized using Hough Transform. Features are extracted and selected by Principal Component Analysis (PCA) then classification is done with LDA (Linear Discriminant Analysis). Ramoser et al. [13] present a fully automated approach to leukocyte segmentation that is robust with respect to cell appearance and image quality. A set of features is used to describe cytoplasm and nucleus properties. Pairwise SVM classification is used to discriminate between different cell types. Evaluation on a set of 1166 images (13 classes) resulted in 95% correct segmentations and 75% to 99% correct classification (with reject option). Vromen and McCane [14]present a model based contour tracing approach to the problem of automatically segmenting a Scanning Electron Microscope image of red blood cells. They use a second order polynomial model and a simple Bayesian approach to ensure smooth boundaries. Poomcokrak and Neatpisarnvanit [15] proposed a method to detect normal red blood cells (RBCs). Neural network is used for classification. This study found the method proposed system has sensitivity 0.86, specificity 0.76 and accuracy 0.74.

III. PROPOSED ALGORITHM

The available techniques for image segmentation can be divided into two main groups as follows [16]: methods based on gray level (e.g.: thresholding, Edge based segmentation), and methods based on image texture.

In the current research, the algorithm proposed in [4] is modified to reduce noise and enhance accuracy. The main disadvantage of the algorithm proposed in [4] is that its accuracy depends momentously on the initial contrast of the image. This limitation leads to capturing of all objects with gray-levels close to that of the WBCs, as will be displayed later using a test image. To overcome this disadvantage we propose to use the nucleus minimum segment size as a constraint to eliminate the non-nucleus objects. A nucleus segment minimum size limit is chosen to be half the RBC average size. This value is selected by experiment. Also morphological opening is executed to remove small pixel groups.

A. Proposed algorithm steps

The blood image is processed as follows:

1) Convert the input image, A, to a gray scale image B.

2) Adjust the gray scale image, B, intensity values with a linear contrast stretching to get image L.

3) Enhance the contrast of the gray scale image, B, using histogram equalization to get image H.

4) Obtain the image R1=L+H.

5) Obtain the image R2=L-H.

6) Obtain the image R3=R1+R2.

7) Implement, three times, 3-by-3 minimums filter on the image R3.

8) Calculate a global threshold value using Otsu's method.

9) Convert R3 to binary image using the threshold from step 8.

10) Use morphological opening to remove small pixel groups. Use a disk structuring element with a radius of 9 pixels.

11) Connect the neighboring pixels to form objects.

12) Apply the size test to remove all objects that are less than 50% of average RBC area.

Figure 1. shows an overview of the proposed algorithm.

B. The details of the proposed algorithm

In step 1, we start by transforming the original smeared image to a gray-scale image, which presents the nuclei of the WBCs the darkest areas in the image. Then the gray scale image goes through a serious of processes to enhance its contrast. In step 4, the addition process will brighten most of the details in the image except the nuclei since by performing the image addition, all the resultant pixels exceeding the intensity value of 225 is truncated to 255. While in step 5, the subtraction process will highlight all the objects and its borders in the image including the cell nuclei. Step 6, the addition process will remove almost all the other blood components while retaining the nuclei with minimum effect of distortion on the nuclei part of the white blood cells. After enhancing the contrast, step 7 applies a minimum filter. The filter works in the same way as the median filter, however instead of changing the pixel intensity with the median intensity value, in minimum filters the pixel intensity is replaced with the minimum intensity value. This step is repeated 3 times for best filtering results which was evident by trials. In step 8, a thresholding technique (Otsu's method) is used. Using the threshold from step 8, we could convert the image R3 to its binary version. Then, in step 10, we use morphological opening to remove the small groups of pixels which can form false objects.

Morphological opening [17]: is done by applying erosion followed by dilation. Erosion: is applying a structuring element B on a binary image A as show by the equation

$$A \ominus B = \{z \in E \mid B_Z \subseteq A\}$$

Where the structuring element B_z is defined by

$$B_Z = \{b + z | b \in B\}, \forall z \in E$$

where E is an integer grid. Similarly erosion is defined by

$$A \oplus B = \{z \in E | (B^s)Z \cap A \neq \varnothing\}$$

B^s is given by

$$B^s = \{x \in E | -x \in B\}$$

So the morphological opening is defined as

$$A \circ B = (A \ominus B) \oplus B$$

Then, in step 11, the algorithm presented in [18] will be used to count and locate the nuclei of the WBCs. The last (step 12) is to check the relative size (area) of each object with respect to average RBC area. The 50% value is used as a minimum nucleus segment threshold. This value was chosen by trials which gave the best accuracy of segmentation.

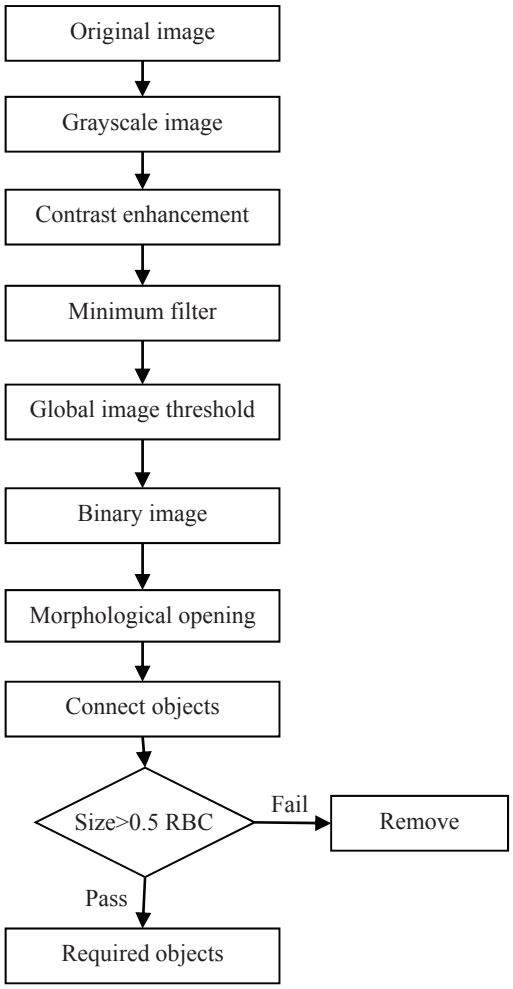

Figure 1. Proposed algorithm to segment WBCs

IV. RESULTS

A. Dataset

A normal peripheral blood sample was taken and stained using standard Gismo-Right technique. Regular light microscope was used to acquire digital images from the blood slide using a 100x objective lens. An analog CCD color camera is connected to the microscope and the video output from is fed to video grabber card to capture color images with VGA resolution (i.e. 640x480 pixels). Total of 365 images were used in this study with the following 5 cell types distributions (TABLE I.):

TABLE I. NUMBER OF USED WHITE BLOOD CELL IMAGES IN TESTING

WBC type	*Number of images used*
Basophil	2
Eosinophil	40
Lymphocyte	33
Monocyte	19
Neutrophil	271
Total	**365**

Cell classification was done manually by expert. We are concerned about classifying the WBC images because each cell type has different shape and color as shown in Figure 2. which affects the classification accuracy as will be shown later.

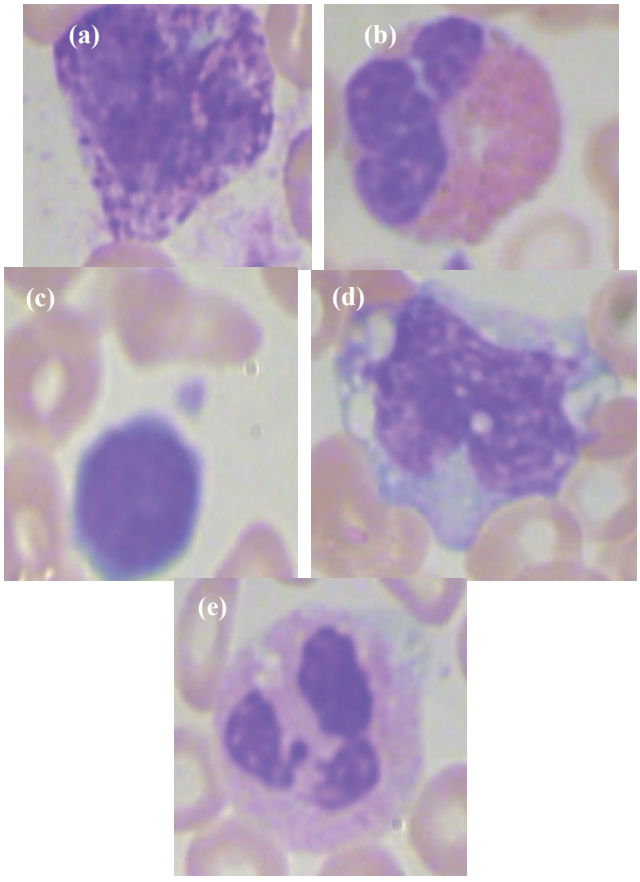

Figure 2. White blood cells (a) Basophil, (b) Eosinophil, (c) Lymphocyte, (d) Monocyte, (e) Neutrophil

B. Method evaluation

To evaluate the segmentation performance we use the performance measurement defined below

$$T_s = 100 * \frac{A_{automatic} \cap A_{expert}}{\max(A_{automatic}, A_{expert})}$$

Where $A_{automatic}$ is the automatically segmented nucleus area and A_{expert} is the area segmented manually by expert. When the automatic area is equal to the expert area then T_s = 100%. This measure is an accurate quantitative measurement of the technique where is can be applied on a large number of data set without subjectivity. For the evaluation purposes, the image dataset was manually segmented to compare the segmentation results of the proposed algorithm. TABLE II. shows the segmentation accuracy results for each cell type using both the proposed algorithm and algorithm proposed by Madhloom et al. [4]. It is clear the superiority of the proposed algorithm for each cell type and for the overall performance of 79.7% compared to 55.9%.

TABLE II. SEGMENTATION ACCURACY RESULTS COMPARISON

Technique	WBC cell type					
	Basophil	***Eosinophil***	***Lymphocyte***	***Monocyte***	***Neutrophil***	***Total***
Mad. [4]	0.643	0.448	0.589	0.574	0.570	0.559
Proposed	0.804	0.693	0.838	0.863	0.803	0.797

C. Performance evaluation for each cell type

Monocytes have the highest accuracy with 86.3%. Example of Monocyte segmentation is shown in Figure 3. The large dark nucleus is the reason for the high accuracy of segmentation for this type.

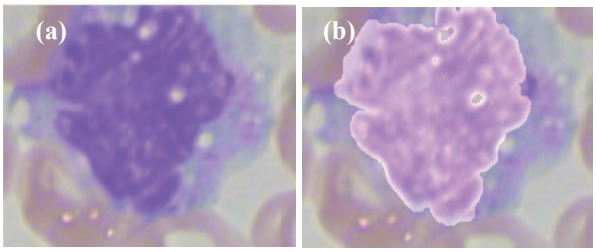

Figure 3. Monocyte segmentation, (a) Original image, (b) Segmented imaged by the proposed method

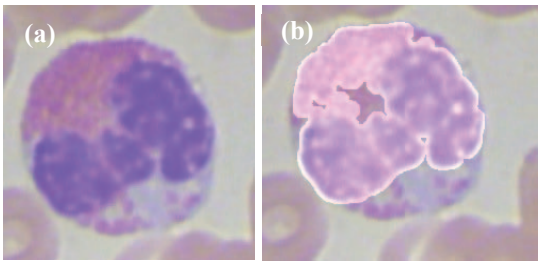

Figure 4. Eosinophil segmentation, (a) Original image, (b) Segmented imaged by the proposed method with error

Eosinophil cells have the lowest segmentation accuracy of 69.3%. This is because Eosinophil cells have dark cytoplasm enclosures which have orange color as show in Figure 4.

For the rest of cell types, the performance is around 80%. This is presented Figure 5. to Figure 7.

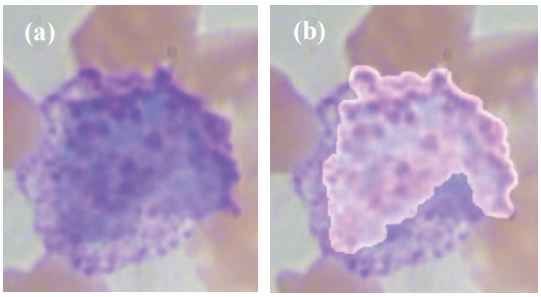

Figure 5. Basophil segmentation, (a) Original image, (b) Segmented imaged by the proposed method

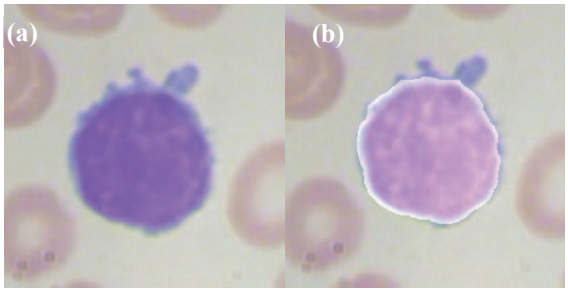

Figure 6. Lymphocyte segmentation, (a) Original image, (b) Segmented imaged by the proposed method

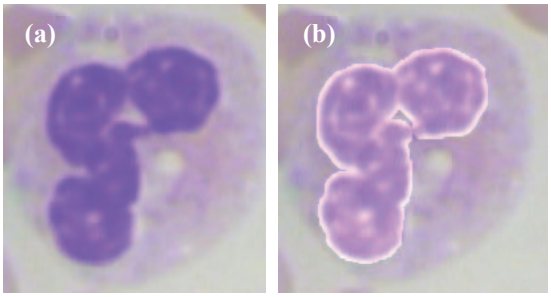

Figure 7. Neutrophil segmentation, (a) Original image, (b) Segmented imaged by the proposed method

D. Comparison with Madhloom et al. [4]

In this sub-section we are comparing our algorithm to the one proposed by Madhloom et al. [4]. The quantitative measurement we presented in TABLE II. shows an overall accuracy of 79.7% for the proposed method compared to 55.9% of Madhloom. Other way of comparison is the visual inspection using a test image shown in Figure 8. Figure 9. and Figure 10. show the segmentation mask images from Madhloom et al. and the proposed algorithm respectively. From the comparison of the tow images one can easily notice how important is the relative-size test that is used to get rid of all the non-nucleus objects. It is evident the superiority of our algorithm compared to the one proposed by Madhloom et al. Figure 11. shows the final segmented image generated by our algorithm.

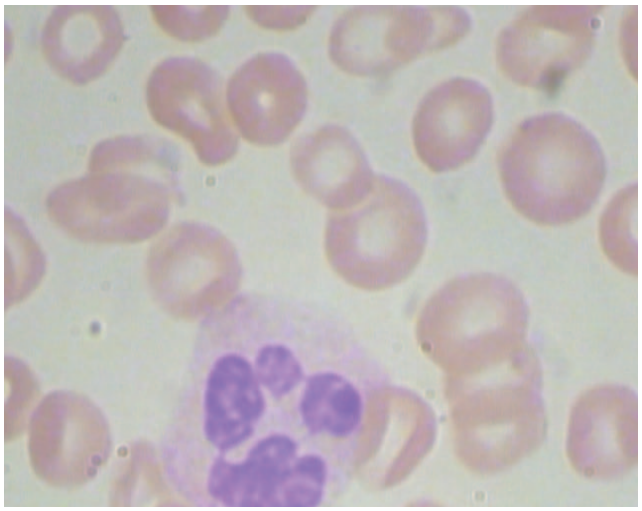

Figure 8. Original test image with one WBC (Neutrophil)

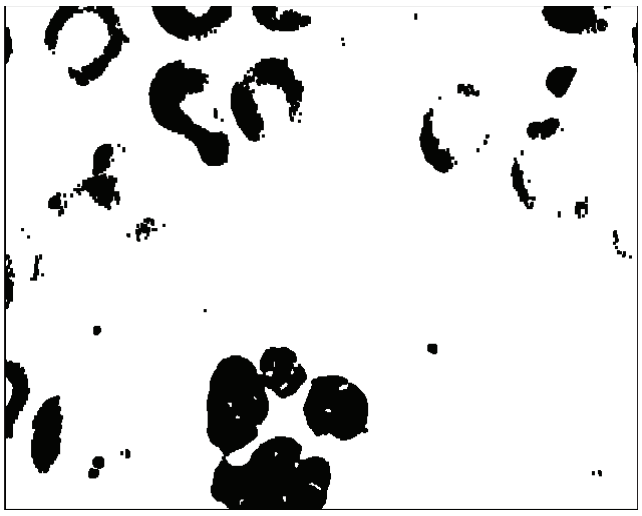

Figure 9. Segmented nucleus using Madhloom et al. algorithm [4]

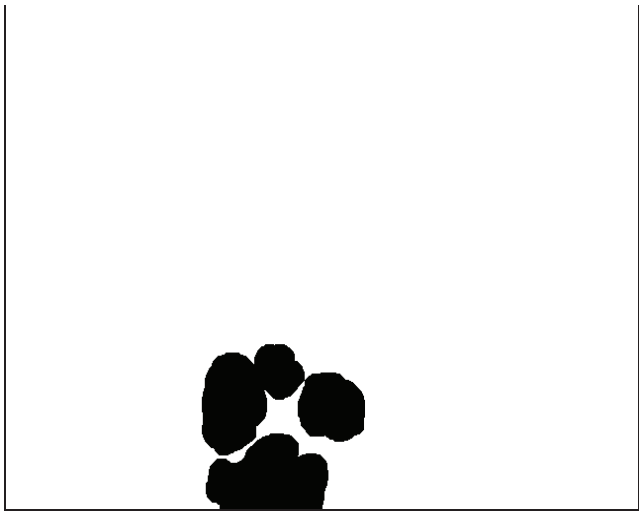

Figure 10. Segmented nucleus image using proposed algorithm

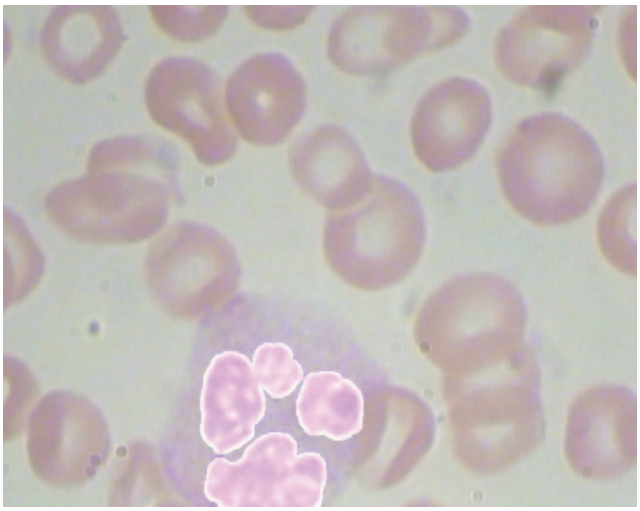

Figure 11. Original test image after applying the nucleus mask from the proposed algorithm

E. Limitations

As inferred by the overall segmentation accuracy of 79.7%, the proposed algorithm has some limitations. The main reasons for the segmentation inaccuracy are the non-homogeneous lighting and over staining. These lead to poor contrast and inexact segmentation. This can be seen in Figure 12.

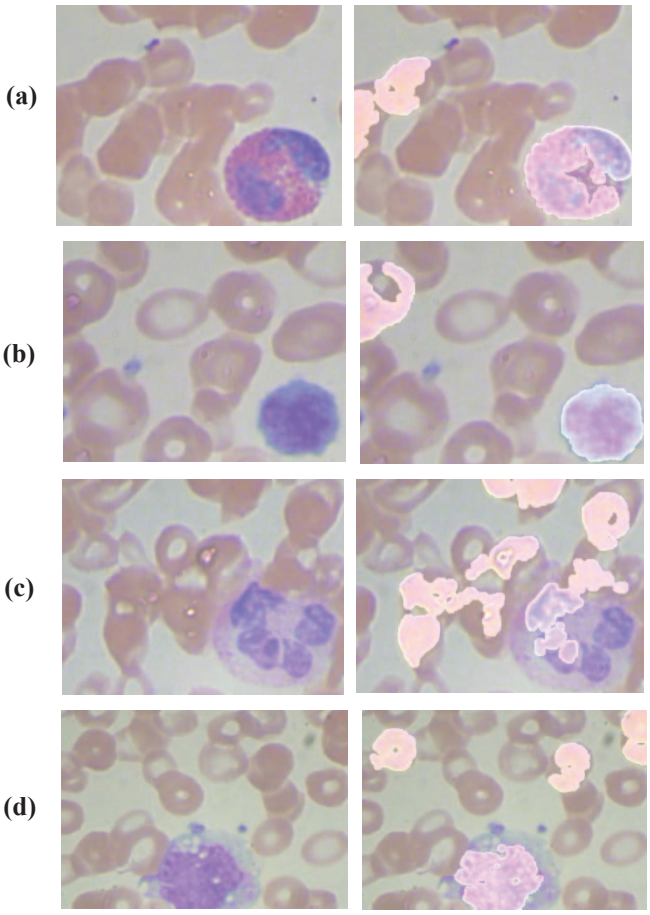

Figure 12. Examples of inaccurate segmentation in (a) Eosinophil, (b) Lymphocyte, (c) Neutrophil, (d) Monocyte.

V. Discussions, Conclusions and Perspectives

In this research work, we presented a new white blood cell segmentation algorithm. The algorithm is a modification of the one proposed by Madhloom et al. [4]. We introduced the minimum size constraint of 50% RBC size to reject small false objects. We tested both algorithms on a test set of 365 WBC images. The test results of the segmentation accuracy show a total similarity of 79.7% for the proposed method compared to 55.9% for Madhloom et al. The average segmentation accuracy for each WBC cell type is 79.7%. Monocytes have higher segmentation accuracy due to its large size; this makes it easier to segment. Eosinophil cells have lower segmentation accuracy of 69.3% due to their dark cytoplasm color which is close to the nucleus color. The total segmentation accuracy still needs more improvement. The MATLAB source code and the blood images dataset are published on MATLAB file exchange website for comparison and re-production [19].

Acknowledgment

We like to thank all who helped us to do this work specially: Amid Abdullah from Calgary laboratory services; Dr. Yaser Hasan and Dr. Mohamed Albashar from school of medicine, Cairo University; Heba Mohamed; Farrah Alami; and Mahtab Moshirpour. They all helped us in understanding and interpreting the biological and clinical parts of the research.

References

[1] "Blood Diseases," *The British Medical Journal,* vol. 2, no. 3907, pp. 998 - 999, 1935.

[2] "What Are Blood Tests?," National Heart, Lung, and Blood Institute (NHLBI), [Online]. Available: http://www.nhlbi.nih.gov/health/health-topics/topics/bdt/. [Accessed 2 May 2012].

[3] J. Wu, P. Zeng, Y. Zhou and C. Olivier, "A novel color image segmentation method and its application to white blood cell image analysis," in *8th international Con. on Signal Processing*, 2006.

[4] 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 Image Arithmetic and Automatic Threshold," *Journal of Applied Sciences,* vol. 10, no. 11, pp. 959-966, 2010.

[5] G. Ongun, U. Halici, K. Leblebicioglu and V. Atala, "An automated differential blood count system," in *Int. Conf. of the IEEE Engineering in Medicine and Biology Society*, 2001, Istanbul.

[6] R. Adollah, M. Y. Mashor, N. F. Mohd Nasir and H. Rosli, "Blood Cell Image Segmentation: A Review," in *IFMBE*, Kuala Lumpur, 2008.

[7] J. Theerapattanakul, J. Plodpai and C. Pintavirooj, "An efficient method for segmentation step of automated white blood cell classifications," in *TENCON*, Bangkok, 2004.

[8] F. Sadeghian, Z. Seman, A. Ramli and B. H. Abdul-Kahar, "A Framework for White Blood Cell Segmentation in Microscopic Blood Images Using Digital Image Processing," *Biological Procedures Online,* pp. 196-206, 2009.

[9] T. Markiewicz, S. Osowski and B. Mariańska, "White Blood Cell Automatic Counting System Based on Support Vector Machine," *Lecture Notes in Computer Science - ICANNGA,* pp. 318-326, 2007.

[10] N. Theera-Umpon, "White Blood Cell Segmentation and Classification in Microscopic Bone Marrow Images," *Lecture Notes in Computer Science - FSKD,* p. 787–796, 2005.

[11] J. Angulo and G. Flandrin, "Automated detection of working area of peripheral blood smears using mathematical morphology," *Analytical Cellular Pathology,* pp. 37-49, 1 January 2003.

[12] G. Cao, C. Zhong, L. Li and J. Dong, "Detection of Red Blood Cell in Urine Micrograph," in *ICBBE*, 2009.

[13] H. Ramoser, V. Laurain, H. Bischof and R. Ecker, "Leukocyte segmentation and classification in blood-smear images," in *IEEE Engineering in Medicine and Biology*, Shanghai, 2005.

[14] J. Vromen and B. McCane, "Red Blood Cell Segmentation from SEM Images," in *Image and Vision Computing (IVCNZ)*, New Zealand, 2009.

[15] J. Poomcokrak and C. Neatpisarnvanit, "Red Blood Cells Extraction and Counting," in *International Symposium on Biomedical Engineering (ISBME)*, 2008.

[16] N. Sharma and L. M. Aggarwal, "Automated medical image segmentation techniques," *Journal of Medical Physics,* vol. 35, no. 1, pp. 3-14, 2010.

[17] S. E. Umbaugh, Computer Imaging: Digital Image Analysis and Processing, CRC Press, 2005.

[18] R. M. Haralick and L. d. Shapiro, "Computer and Robot Vision," in *Computer and Information Science* , Addison-Wesley, 1992, pp. 28-48.

[19] "Source code and blood image dataset," Mathworks File Exchange website, [Online]. Available: http://www.mathworks.com/matlabcentral/fileexchange/36634 .

[20] D. Tagliasacchi and G. Carbon, "The blood cells," April 1997. [Online]. Available: http://www.funsci.com/fun3_en/blood/blood.htm. [Accessed 2 May 2012].