Journal of Applied Sciences

ISSN 1812-5654

science alert

ANSInet
an open access publisher
http://ansinet.com

Journal of Applied Sciences 10 (11): 959-966, 2010
ISSN 1812-5654

An Automated White Blood Cell Nucleus Localization and Segmentation using Image Arithmetic and Automatic Threshold

[1]H.T. Madhloom, [1]S.A. Kareem, [2]H. Ariffin, [1]A.A. Zaidan, [1]H.O. Alanazi and [2]B.B. Zaidan
[1]Faculty of Computer Science and Information Technology, University of Malaya, Kuala Lumpur, Malaysia
[2]Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia

Abstract: The aim of this research is to automate the process of detection and classification of leukocytes using image processing techniques. White blood cell recognition and classification into various distinct subtypes is very important in clinical and laboratory tests. The nucleus features are adequate to identify the type of the cell in most of the case, the traditional morphology test which is done by a hematology expert to look at the cell under the microscope is a time consuming and tedious job, beside that the medical instrument which is used to do the test are costly and may not be exist in all the hospitals and clinics. An automatic image segmentation system can make the inspection procedure of blood smear much easier and faster and the amount of data that can be analyzed by such a clinician handle more data than they normally can handle. The most crucial step in such systems is in white blood cell segmentation. In this research we focus on white blood cell nucleus segmentation that can be used to separate the nucleus from the whole cell body by using a combination of automatic contrast stretching supported by image arithmetic operation, minimum filter and global threshold techniques. Results showed that the proposed method manages to obtain accuracy between 85-98%. The results showed that the proposed method is promising comparing to the result from the expert.

Key words: Differential blood count, image analysis, automatic cell segmentation, leukemia diagnosis, segmentation evaluation

INTRODUCTION

The principal aim of image segmentation is the extraction of significant objects presented in the image either by partitioning the image into connected semantic regions or by extraction one or many specific objects from the image (Kies and Benamrane, 2008) and one of the main concern for the matter of segmentation nowadays is medical images. Medical images are considered one of the most important tools and techniques used for many clinical diagnosis and decision making (Anoraganingrum, 1999; Markiewicz *et al*., 2005; Adollah *et al*., 2008). The importance and success of medical image processing has a profound impact. An automatic image segmentation system can make the inspection procedure of blood smear much easier and faster and the amount of data that can be analyzed by such a system is more than what the clinician can handle. Besides that, the analyses of blood cell suffer from slowness and the results may not always be the same depending on the clinician's experience. Medical instruments that are used for blood cell analyses such as hematology analyzer and flow-cytometry are quite expensive and may not exist in all the medical laboratories. This has resulted in the area of automatic image segmentation and classification gaining importance and encouraging researchers to develop a computerized automatic medical decision- support system that is cost effective, efficient, easily available and can be used by a non specialist or a rather junior clinician. The main objective of medical image processing and analysis is basically to collect useful information about the region of interest in the image, extract statistical measurements that can lead to early, accurate diagnosis, hence cost-saving and providing better monitoring and evaluation in the progress of the treatment (Gonzalez *et al*., 2003). Medical image processing aids in the discrimination between acute lymphoblast leukemia and acute myeloid leukemia. Ultimately, the main purpose of medical imaging is to better manage leukemia patients and hence improving their healthcare and quality of life.

Segmentation is the process of identifying regions of an image that have common properties while at the same time separating regions that are dissimilar and retaining maximum useful information from the image (Adollah *et al*., 2008; Gonzalez *et al*., 2003).

Blood disorder is considered among the most dangerous of diseases that can lead to death. Many of these blood diseases are related to white blood cells such as Leukemia.

Corresponding Author: Hayan T. Madhloom, Faculty of Computer Science and Information Technology, University of Malaya, 50603 Kuala Lumpur, Malaysia Tel: +060172452457

The observation of blood smear under a microscope gives invaluable qualitative and quantitative information which helps to diagnose a variety of diseases including Leukemia (Theera-Umpon, 2005). White blood cell identification and classification into several subtypes is very valuable in clinical and laboratory tests. In many cases, only the nucleus information is enough to classify a particular cell (Theera-Umpon, 2005).

Blood is a red liquid (that is vital of life) composed of an isotonic fluid (plasma) in which various cells (hemocytes) are suspended. There are three major groups of these cells, all blood cells are manufactured in the bone marrow, growing from a cell known as hematopoietic stem cell (Ciesla, 2007). The red blood cells (also called Erythrocytes) that contain hemoglobin carry the oxygen to the tissues, platelets (also called thymocytes) that is responsible for the clotting of the blood, The Leukocytes (also known as white blood cell), which are cells of the immune system defending the body against both infectious disease and foreign materials (Ciesla, 2007). There are five kinds of white blood cells that form part of the immune response; each with different functions, namely, neutrophils, lymphocytes, monocytes, eosinophils and basophils and these are the white blood cells belong to two groups granulocytes or agranulocytes. Blood smear are routinely used in the clinical laboratory to give indication about the different types of diseases and help to diagnose, monitor and evaluate the patient case.

Manual classification of blood cells can be tedious and error prone and depends on the clinicians' experience and expertise. Counting and classifying blood cells is subject to the interpretation of the clinical expert. Different clinicians may give different interpretations to the same blood sample. Even the same clinician may sometimes provide different interpretations on different occasions. Thus, there is an increasing demand for an automatic system to accurately and efficiently analyses blood cells and overcome the problems associated with the manual classification system (Poon *et al*., 1992).

One of the greatest challenges to engineers and computer scientists especially in the field of biomedicine is to transform the tedious human task into a computer based process in which the system outperforms the manual system (Adollah *et al*., 2008). Until now only a few attempts of a partial/full automated system based on image processing systems have been developed and most are still at the prototype stage (Scotti, 2006).

Image segmentation is major part of any image analysis and understanding system. The main idea of image segmentation is to separate different object in the image scene by mainly dividing the image into two parts, namely, the region of interest and the background.

The segmentation process is critical because the result from this step serves as the basis for all subsequent analyses, such as the extraction of the shape of nucleus and cytoplasm). Cell segmentation is also one of the most challenging problem due to two problems, the first one is the complex biological nature of the cell and the second one is the technical problem that is caused by the staining method (Anoraganingrum, 1999). Although, the staining method is a simple procedure, there are many reasons that as a test it could easily fail or be less effective than it should be (Houwen, 2002). Furthermore, images acquired using different camera specifications can also affect the resolution of the images.

Segmentation can be categorized as supervised or unsupervised learning/classification (Yi *et al*., 2005). Different types of segmentation methods are used in white blood cell segmentation and one of the main issue in blood image segmentations is the acquisition criteria that's why varies segmentation methods developed till now and its differ from one image to other (Giridhar Akula *et al*., 2007). These are, namely, threshold-based, edge-based, region-based or clustering methods, such as, fuzzy-C mean clustering and K-mean clustering. The main concern in this paper is the segmentation of white blood cell nucleus. The goal of segmentation is to divide the white blood cell to its basic components like nucleus, cytoplasm, blobs and many others. In an automatic system the most important step is the segmentation, hence, all the other subsequent steps like feature extraction and classification are totally dependent on the accuracy and reliability of the segmentation. If the segmentation is wrong then all the subsequent steps will invariably be wrong. Threshold techniques cannot always produce meaningful results since no spatial information is used during the selection of the segmentation threshold. They are often combined with mathematical morphology operations. In the research carried out by Cseke, automatic threshold is used, based on the method proposed by Otsu (1979). In the Otsu (1979) method the thresholds are selected with a simple recursive methods derived from maximizing the interclass variance between dark, gray and bright regions. Cseke (1992) used the image mathematical morphology (opening and closing) as a final step to smooth the region of interest. The technique yields a result of a 92% accuracy.

Liao devised a method to localize white blood cell by using a simple thresholding approach to give initial labels to pixels in the blood cell image, then the labels are adjusted with a shape detection method based on large regional context information to produce meaningful result, but the algorithm is based on a priori information about the blood smear images. This scheme is only suitable if the shape of each white blood cell boundary is close to a circle (Liao and Deng, 2002). Edge detection methods

perform poorly on cell images because not all cell details are sharp, so it is difficult to get the edge information and locate the cells, but, as stated by Piuri and Scotti, edge detection methods perform well if the contrast between the background and the gray internal of the membrane of the cell is stretched by using a contrast stretching filter (Jeacocke and Lovell, 1994; Piuri and Scotti, 2004). As an attempt to automatically analyze the morphology of the white blood cell, canny edge detector is used after stretching the contrast to reconstruct the border of the membrane followed by dilation morphological operator to better connect separated points of the perimeter and to make the perimeter of the cell as a connected item (thicker by more than one pixel). For the nucleus detection the same methods is used except that a 4×4 low pass morphological filter is applied before using the canny edge detection to ensure a better homogenization of the nucleus's color. Kumar defined a new edge operator to highlight the nucleus boundary, which is very effective to segment the nuclei from cell images (Kumar *et al*., 2002). But the cytoplasm region cannot be extracted by such a method since its boundary is more complex. Kumar used a simple morphological method to segment cytoplasm from the background and Red Blood Cell (RBC), which required some assumptions that may be untenable in many cases.

Jiang *et al*. (2003) introduced novel white blood cell segmentation by joining two techniques, scale space filtering and watershed clustering which effectively avoid problems due to spatial variety and complexity in image space. In this scheme, two components of WBC, nucleus and cytoplasm, are extracted respectively using different methods. First, a sub image containing WBC is separated from the cell image. Then, scale space filtering is used to extract the nucleus region from the sub image. Later, watershed clustering in 3-D HSV (Hue, Saturation, Value) histogram is processed to extract cytoplasm region. Finally, morphological operations are performed to obtain the entire connective scheme successfully which avoids the variety and complexity in image space and can effectively extract various WBC regions from images of peripheral blood smear.

Sinha and Ramakrishnan (2003) used the K-Mean clustering method to locate the white blood cell nuclei to identify the white blood cell on the image following by a crop operation to separate the entire cell from the image scene. However, the method of cropping the entire cell in order to get the real area of the whole cell is not clearly shown. Theera-Umpon (2005) used a fuzzy C-Mean clustering to segment single cell images of white blood cells in the bone marrow into two regions, i.e., nucleus and non-nucleus. However if the number of clusters is more than 2 then the computational time increases.

MATERIALS AND METHODS

In the current research we introduce a simple method to extract the nucleus of five types of white blood cells from the other components of the image. For this experiment we use MATLAB 7.1 image processing toolbox and peripheral blood smear images for white blood cell obtained from a veterinary clinical pathology database that is available online (Lanevschi, D.A. Veterinary Clinical Pathology Image Database). All the blood smears are stained with a Wright's staining method, Fig. 1a-d show a sample of the five types of white blood cells that can appear in the peripheral blood smear.

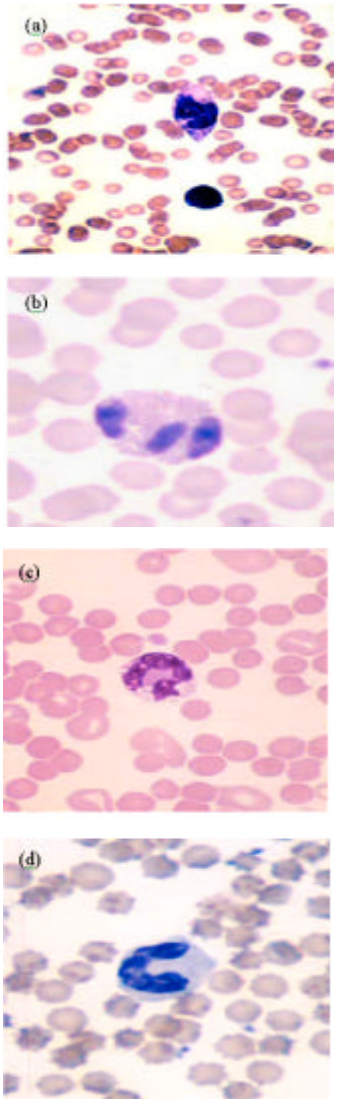

Fig. 1: White Blood Cell Samples (a) Basophils and Lymphocyte, (b) Eosinophils, (c) Monocyte and (d) Neutrophil

PROPOSED METHOD

This research was conducted at University of Malaya Faculty of computer science and also supported by University Malaya Medical center (UMMC) in Malaysia-Kuala Lumpur 2009.

In the beginning all the images are converted into grayscale images so that the nucleus part of the cell will appear as the darkest part of the image. The localization of the white blood cell nucleus is based on automatic contrast stretching, histogram equalization and image arithmetic. Figure 2 shows the proposed steps of the white blood nucleus segmentation scheme, after converting the original image which it is in RGB color format to a grayscale image, all the subsequent steps will work on the grayscale image. First one copy of the image will be enhanced with a linear contrast stretching (in this research this is referred to as L) and another copy will be enhanced with histogram equalization (in this research this is referred to as H). The result from L is added to the resultant image from H and subsequently called R1 (Eq. 1). By performing the image addition, all the resultant pixels exceeding the intensity value of 225 is truncated to 255 which brighten most of the details in the image except the nucleus.

$$R1(i, j) = L(i, j) + H(i, j) \quad (1)$$

The result R1 (Fig. 3a) is then subtracted from the histogram equalized image (H) to form R2 (Eq. 2). This operation highlights all the objects and its borders in the image including the cell nucleus as shown in Fig. 3b.

$$R2(i, j) = R1(i, j) - H(i, j) \quad (2)$$

The last arithmetic operation is to add both of the results R1 and R2 together to produce R3 (Eq. 3).

$$R3(i, j) = R1(i, j) + R2(i, j) \quad (3)$$

As shown in Fig. 4, this operation removes almost all the other blood components while retaining the nucleus with minimum affect of distortion on the nucleus part of the white blood cell.

After this operation is completed a global threshold using Otsu's method is used as long as the resulting histogram becomes bimodal as shown in Fig. 5.

If we use the automatic threshold at this stage it will lead to miss-segmentation of some part of the nucleus due to the affect of distortion after the last arithmetic operation. To avoid this problem, a [3 by 3] minimum filter is used to increase the intensity value making the nucleus part darker so it can be fully detected using a threshold. This is shown in Fig. 6 and 7, respectively.

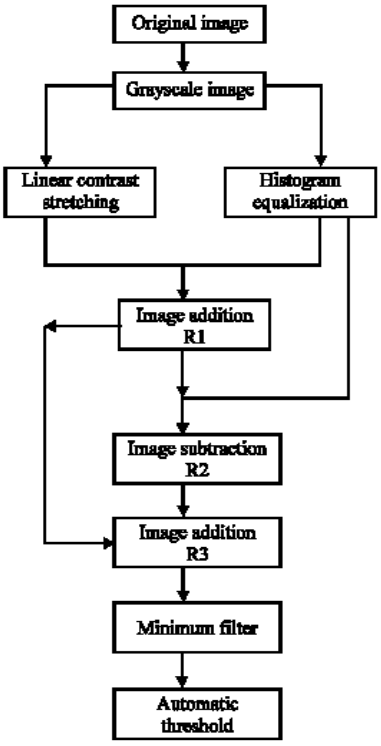

Fig. 2: The structure of steps composing the proposed method

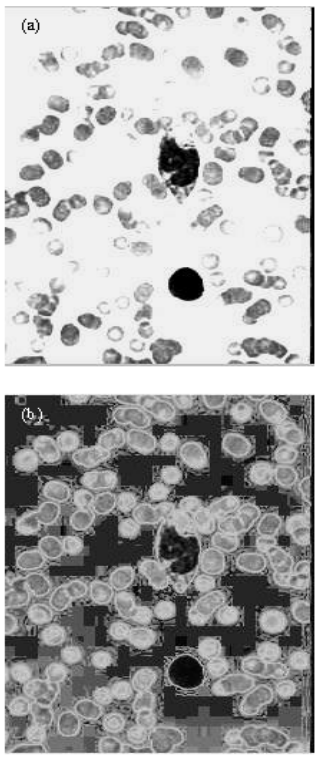

Fig. 3: (a) R1 and (b) R2

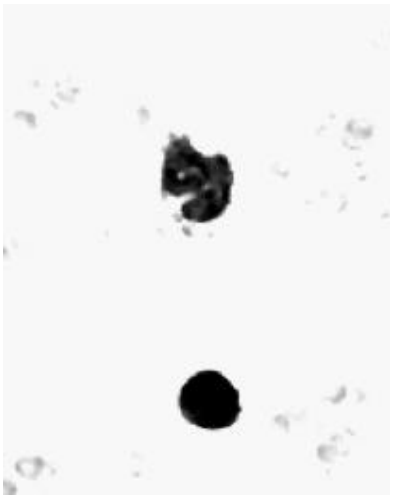

Fig. 4: Result R3 on image (a) from Fig. 1

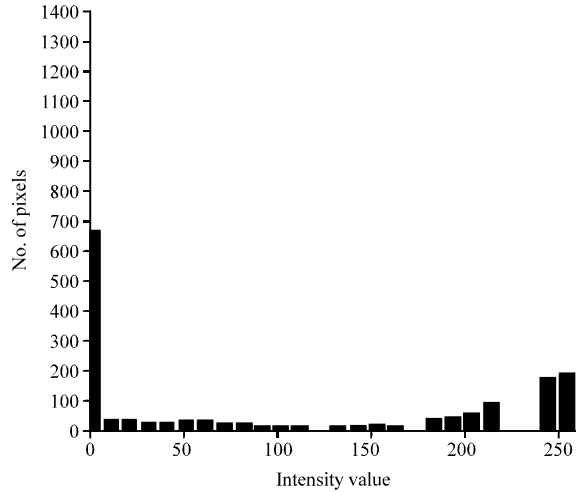

Fig. 5: Histogram of Image in Fig. 4 before apply minimum filter

The last step in the nucleus segmentation is to threshold the image using automatic global threshold. The result is shown in Fig. 7.

RESULTS AND DISCUSSION

The experiment shows how automatic contrast stretching such as, linear contrast stretching and histogram equalization can be combined together with the use of image arithmetic operations to remove most of the details in the image like red blood cells and other parts of the white blood cell and retain only the nucleus of the white blood cell as shown in Fig. 3 and 4. It also shows how minimum filter can improve the segmentation result

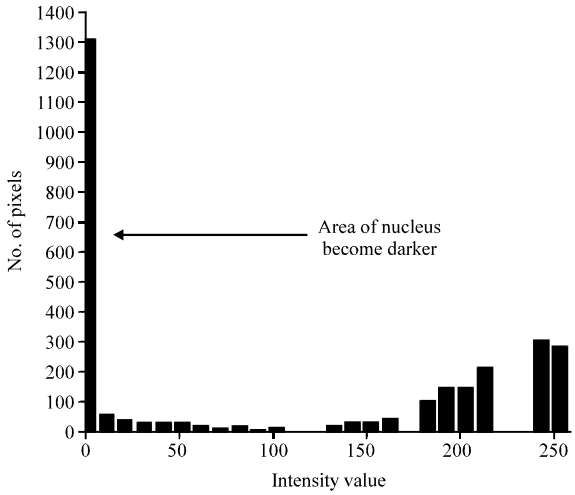

Fig. 6: Histogram of image in Fig. 3 after apply Minimum filter

Fig. 7: Result R3 after applying Minimum filter on image in Fig. 2

by making the nucleus area darker so it can be easily detected by the automatic threshold as shown in Fig. 5-7. Figure 5 and 7 show how the there is a prominent difference in the histogram before and after applying the minimum filter Fig. 8a-d show the final result of the segmentation of the nucleus after the threshold where all the other blood components are remove and only the white blood cell nucleus is maintained. Normally and in most of the image segmentation research, the initial way to validate the segmentation result is through the visual inspection of the resulted image Fig. 9a-d show how the perimeter of the resulted nucleus is superimposed on the original image. It shows how the boundary of the created

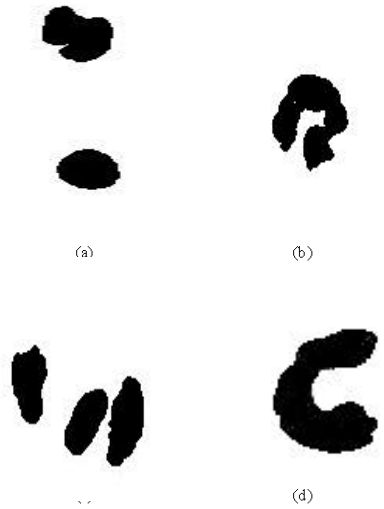

Fig. 8: Final result of nucleus segmentation (a) Basophils and lymphocyte, (b) Eosinophils, (c) Monocyte and (d) Neutrophil

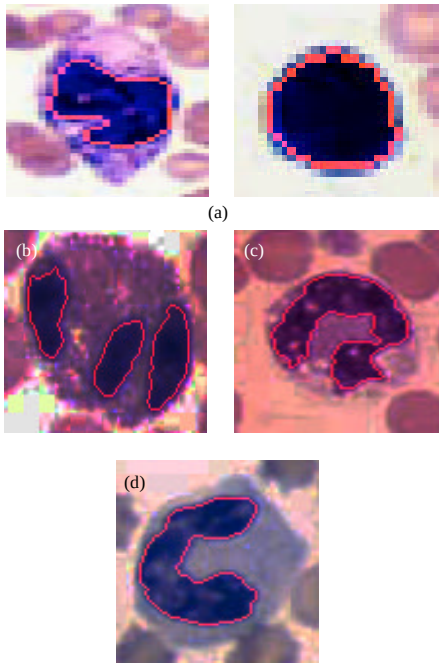

Fig 9: Final Result of created nucleus segments superimposing on the original image (a) Basophils and lymphocyte, (b) Eosinophils, (c) Monocyte and (d) Neutrophil

segment is a perfect fit of the boundary localization of the nucleus of the original grayscale image without any false tracking of the nucleus perimeter.

The method has been applied to the five types of white blood cells and it has successfully isolated the nucleus part of the cell from the other components of the

Fig. 10: Segmented nucleus of eosinophils (a) without using of minimum filter and (b) with using of minimum filter

image; as an initiative evaluation and test for the results of this method, it has been visually inspected by a hematology expert and stated that the results is encouraging.

The strength of this method is in the use of minimum filter before applying the threshold to the image. Figure 10 shows the segmented nucleus of Eosinophils before (Fig. 10a) and after (Fig. 10b) the use of minimum filter. As shown in Fig. 10a, the outline of the nucleus is not completed and the threshold produce some gaps in the nucleus which lead to miss-segmentation for the reason that the intensity value of the miss-segmented pixels are higher than the threshold value. In order to make the nucleus have a more natural look, minimum filter is applied. Figure 10b shows how the outline and the inner part of the nucleus are more prominent and it is ready for feature extraction. Minimum filter work in the same way as the well-known 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.

Compare with some other white blood cell segmentation methods from the literature, the proposed algorithms made some improvement on the initial localization and segmentations of the white blood cell nucleus, like Theera-Umpon N (Theera-Umpon, 2005), also tried to solve the same problem of white blood cell nucleus segmentation but with an unsupervised method which it is a fuzzy c-mean clustering, one of the disadvantages of this method is the identification of the numbers of clusters. The overall segmentation accuracy obtained by Theera-Umpon N was 89.90 whereas the proposed method gives maximum accuracy of 98.81. Cseke (1992) used a simple localization method to localize the white blood cell based on the green component of the original RGB image, the main problem with this method is how to identifies the threshold value that will remove all the blood components and retain only the nucleus, Cseke (1992) managed to obtain approximately 0.92 accuracy in the overall segmentation process. The same method for

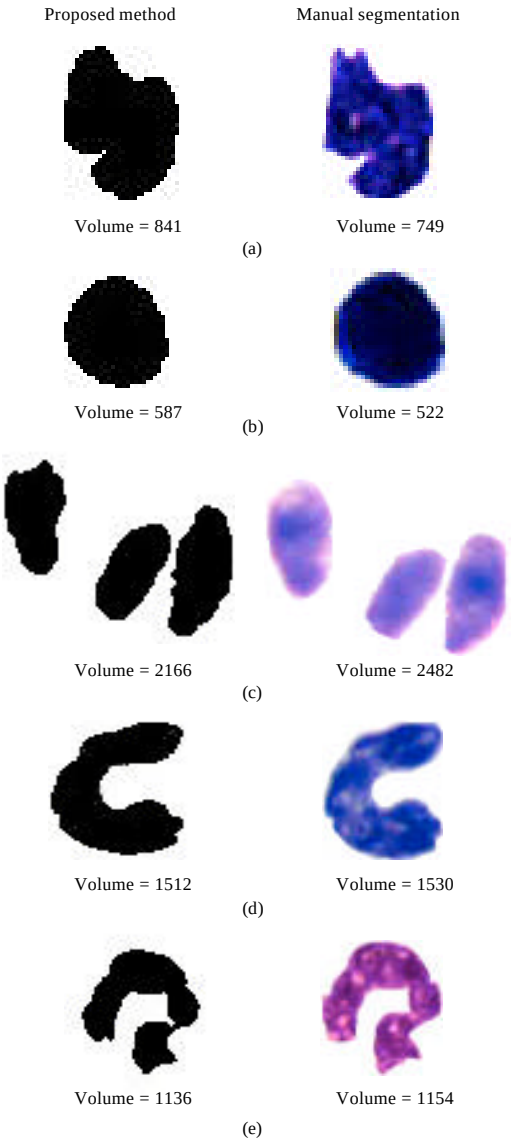

Fig. 11: Accuracy rate of the segmentation, (a) accuracy rate = 87.717, (b) accuracy rate = 87.548, (c) accuracy rate = 854109, (d) accuracy rate = 98.81 and (e) accuracy rate = 98.41,

white blood cell localization is used by Jiang *et al.* (2003) as they stated that nucleus pixels can be easily separated from others in the image histogram of green layer. But an automatic threshold value may not he easily obtained because nucleus cannot form an obvious cluster in histogram due to its relatively small area in the whole cell Image, that's why our proposed method will prepare the image for the next step which is threshold the image in order to obtain only the white blood cell nucleus. In addition Liao and Deng (2002) devised a method to localize white blood cell to give initial label to some pixels in the blood cell, but he adjust the label with a shape detection method, the disadvantage of this method is that it is only appropriate if the shape of each white blood cell boundary is close to a circle, while the proposed method it will localize the white blood cell nucleus no matter what shape it's taken

For the same purpose Piuri and Scotti (2004) proposed an algorithm for nucleus localization which is based on contrast stretching and opening operation to extract a sub image of white blood cell from the original, the structure element size is based on a priori knowledge about the cell diameter, the opening operation will not remove all the background component hence the threshold may remove some part of the nucleus like in Fig. 10a, oppositely present proposed method will remove almost all the components and highlight the nucleus too before applying the threshold to the image.

To validate the proposed method statistically, the global quantitative method is used, in this method the nucleus of each cell is segmented manually by a hematologist and then it is compared with the proposed method result, the evaluation is done according to the volume and proportion of correct pixels of the nucleus that is resulted from the manual and automatic method to get the error rate so it can be used as an enhancement factor in the future work. Figure 11 shows the final result of the segmentation evaluation and how it is compared to the result from manual segmentation made by expert.

CONCLUSIONS

A prominent facet of automatic blood cell recognition and analysis is the ability to find all the nucleated cells in the image scene and remove all the remaining components. The main issue for such system is to perfectly localize the white blood cells in the image and distinguish it from the other parts of the image. The nucleus information can give valuable information about whether the cell is a blast or normal cell. Until now most of the researches that have been done were about cropping the while blood cell image manually and then do the process on it.

This study introduces a method for white blood cell nucleus localization and segmentation as a first step towards a fully automatic system for leukemia diagnosis and classification using peripheral blood microscope image. White blood cell segmentation is the key procedure in the automatic leukemia diagnosis system.

The main purpose of white blood cell segmentation is to accurately crop a sub-image of white blood cell from the image scene and this can be done using the nucleus

because it is the darker part in the blood cell image. Then, the cropped sub-image is segmented into its morphological parts, such as, cytoplasm, nucleus etc. in order to get the features from these region of interest and feed them to a classifier system to make a reasonable decision about the categories of different leukemic cells.

The variation in the accuracy rate between the proposed result and the manual segmentation is due to the type of the cell and also the method that is used to prepare the slides beside that the manual segmentation is highly depend on the hematologist capabilities. The best segmentation result is manifest from the Monocyte and Neutrophil image where the boundaries of the nucleus and the cytoplasm are very distinct. Due to the small size of the dataset, further work needs to be done on a larger dataset of white blood cells.

REFERENCES

Adollah, R., N.F.M. Nasir, H. Rosline, H. Mahsin and H. Adilah, 2008. Blood cell image segmentation: A review. Proceedings of the 4th Kuala Lumpur International Conference on Biomedical Engineering, July 30, Malaysia, Kuala Lumpur, pp: 141-144.

Anoraganingrum, D., 1999. Cell segmentation with median filter and mathematical morphology operation. Proceedings of the IEEE International Conference on Image Analysis and Processing, Sept. 27-29, Venice, Italy, pp: 1043-1046.

Ciesla, B., 2007. Hematology in Practice. F.A. Davis Co., Philadelphia.

Cseke, I., 1992. A fast segmentation scheme for white blood cell images. Proceedings of the 11th IAPR International Conference on Pattern Recognition Image, Speech and Signal Analysis, Aug. 30-Sept. 1, Netherlands, pp: 530-533.

Giridhar Akula, V.S., P. Chandrasekhar Reddy, V. Rajaraman and A.V. Sasidhar, 2007. An enhanced procedure for image segmentation and smoothing. J. Applied Sci., 7: 1452-1455.

Gonzalez, R.C., E.R. Woods and L.S. Eddins, 2003. Digital Image Processing using MATLAB. 2nd Edn., Prentice Hall, USA., ISBN-10: 0130085197, pp: 624.

Houwen, B., 2002. Blood film preparation and staining procedures. Clin. Lab. Med., 22: 1-14.

Jeacocke, M.B. and B.C. Lovell, 1994. A Multi-resolution algorithm for cytological image segmentation. Proceedings of the 2nd Australian and New Zealand Conference in Intelligent Information Systems, Nov. 29-Dec. 2, USA., pp: 322-326.

Jiang, K., Q.X. Jiang and Y. Xiong, 2003. A novel white blood cell segmentation scheme using scale-space filtering and watershed clustering. Mach. Learning Cybernetics, 5: 2820-2825.

Kies, K. and N. Benamrane, 2008. Medical image segmentation using enhanced hoshen-kopelman algorithm. J. Applied Sci., 8: 2474-2479.

Kumar, B.R., D.K. Joseph and T.V. Sreenivas, 2002. Teager energy based blood cell segmentation. Proceedings of the 14th International Conference on Digital Signal Processing, July 1-3, Santorini, Greece, pp: 619-622.

Liao, Q. and Y. Deng, 2002. An accurate segmentation method for white blood cell images. Proceedings of the 2002 IEEE International Symposium on Biomedical Imaging, July 7-10, USA., pp: 245-248.

Markiewicz, T., S. Osowski, B. Marianska and L. Moszczynski, 2005. Automatic recognition of the blood cells of myelogenous leukemia using SVM. Proceedings of the International Joint Conference on Neural Networks, July 31-Aug. 4, Montreal, Canada, pp: 2496-2501.

Otsu, N., 1979. A threshold selection method from grey level histogram. IEEE Trans. Syst. Man Cybern., 9: 62-66.

Piuri, V. and F. Scotti, 2004. Morphological classification of blood leucocytes by microscope images. Proceedings of the International Conference in Computational Intelligence for Measurement Systems and Applications, July 14-16, Boston, MA. USA., pp: 103-108.

Poon, S.S., R.K. Ward and B. Palcic, 1992. Automated image detection and segmentation blood smears. Cytometry, 13: 766-774.

Scotti, F., 2006. Robust segmentation and measurements techniques of white cells in blood microscope images. Proceedings of the IEEE Conference in Instrumentation and Measurement Technology, April 24-27, IMTC Sorrento, pp: 43-48.

Sinha, N. and A.G. Ramakrishnan, 2003. Automation of differential blood count. Proceedings of the Conference on Convergent Technologies for Asia-Pacific Region, Oct. 15-17, Banglore, India, pp: 547-551.

Theera-Umpon, N., 2005. Patch-based white blood cell nucleus segmentation using fuzzy clustering. ECTI Transa. Electrical Eng. Electr. Commun., 3: 15-19.

Yi, F., P. Chen and L. Li, 2005. White blood cell image segmentation using on-line trained neural network. Proceeding of the 27th Annual IEEE Conference Engineering in Medicine and Biology, Sept. 17-18, Shanghai, China, pp: 6476-6479.