Research Journal of Applied Sciences, Engineering and Technology 7(14): 2987-2993, 2014
DOI:10.19026/rjaset.7.631
ISSN: 2040-7459; e-ISSN: 2040-7467

Submitted: November 28, 2013 Accepted: December 17, 2013 Published: April 12, 2014

Research Article

Automatic Color Nuclei Segmentation of Leukocytes for Acute Leukemia

Naveed Abbas and Dzulkifli Mohamad
Faculty of Computing, ViCube Research Group, Universiti Technologi Malaysia, Malaysia

Abstract: In this study, we used an efficient and simple technique for the automatic segmentation of nuclei of Leukocytes, which is not only accurate in the results but also fast as compared to other published algorithms. Leukemia is a blood cancer and has several types but all types begin from the cells in the bone marrow. Hematology, the study of blood is an important step for the diagnosis of various diseases especially for Leukemia. Manual Hematology is done by experts for any blood disorder, but is time consuming, high mental and physical labor is required, highly subjective and erroneous. The diagnosis of any diseases required high accuracy but Leukemia is a life critical disease and has zero tolerance for the errors. Automatic diagnosis of microscopic images through digital image processing techniques is the need of the day. Segmentation is one of the most important and challenging technique in the automatic diagnosis of Leukemia. The developed method is tested on 380 images and the accuracy of various types of Leukocytes (WBCs) was found out qualitatively. Also the developed technique is quantitatively evolved for speed performance. The segmentation accuracy of the developed technique is 96.5% while the efficiency reduces the computation time by 50% as compared to other published techniques. The algorithm along with the data set is published on MATLAB file exchange for an evaluation.

Keywords: Blood cells, dataset, hematology, leukemia, leukocytes, segmentation, WBCs

INTRODUCTION

Clinical examination of many diseases like Blood Cancer, HIV/AIDS, Malaria, Diabetes and Anemia etc, requested the blood tests called Hematology. Leukemia, the blood cancer is a fatal disease starts in the bone marrow, has the responsibility of formation of blood cells. The bone marrow forms abnormal leukocytes i.e., White Blood Cells (WBCs) called leukemia cells and are present in the patients of Leukemia. These cells are formed very fast and definitely results as abundance in blood which caused anemia, bleeding and infections (Abbas and Mohamad, 2013) There are four main types of Leukemia; the diagnosis involves blood order tests and keen examination needs the experts to check the nuclei of the leukocytes through manual microscopic methods in order to know about the type of Leukemia (Nordqvist, 2009). As manual hematology is time consuming job also involves the jeopardy of errors.

Hence, the automatic diagnosis through Digital Image processing techniques is the need of the day. Segmentation (grouping the image's pixels to its constituent object in the image) in this diagnosis is considered to be the foundation and critical step for the letter steps of features extraction and classification of Leukemia cells. The main focus of this study is on automatic segmentation of nuclei of leukocytes and no any enhancement or modification of the developed methods has been made but a very simple and novel method has been introduced in the area.

The current topic is rich in literature but due to the use of In-vitro slides the main goals of the Leukemia cell segmentation in the real environment will be highly opposite to the published work. The images from which the segmentation of the Leukemia cells has been made have a very diverse and complicated background having Red Blood Cells (May be either clumped or not), platelets or any other artifacts. At the present there is no universal Nuclei segmentation technique. In the study of Mohamed and Far (2012b), did intensity based segmentation through Gram-Schmidt Orthogonalzation for improving the speed from its earlier version i.e., Mohamed and Far (2012a) and achieve 85.5% accuracy and 16% faster performance than Rezatofighi *et al*. (2009) and Ramoser *et al*. (2006). Ongun *et al*. (2001) segment the White Blood Cells and for features of shape and texture used morphological structures in the process of identification and classification by considering twelve classes of Leukocytes. Tabrizi *et al*. (2010) used contour based segmentation approach (snake contour tracing) for segmentation of Leukocytes and further for identification PCA, LVO and SVM are employed.

Rezatofighi *et al*. (2009) in the same way as Mohamed and Far (2012b) did intensity based segmentation through Gram-Schmidt Orthogonalzation

Corresponding Author: Naveed Abbas Faulty of Computing, ViCube Research Group, Universiti Technologi Malaysia, Malaysia

but instead of color, binary segmentation has been done the rest of the work has no difference.

Dorini *et al*. (2007) did contour based binary segmentation of the leukocytes using watershed transform. In the work of Theera-Umpon and Dhompongsa (2007) did binary segmentation on granulometric moments and further the Bayes classifier and neural network are employed for classification process. Ramoser *et al*. (2006) presented an automated segmentation method for leukocytes segmentation. For classification of cell types SVM is used. Evaluation has been made on a set of 1166 images having 13 types and resulted in 95% correct segmentations.

MATERIALS AND METHODS

The techniques for cell segmentation are mainly divided into two type's i.e, (thresholding, Edge based, intensity based) and texture based as the proposed method employ the first one for fast performance. There are two plus points of any type of leukocyte's nuclei, the first one that they appeared with low values in the three channels of the input RGB image and the second one that they encompassed the largest area in the image. Both of these points will help in the correct segmentation of the nuclei.

Steps of the proposed method: The proposed method can process the input images according to the given steps:

- Read the input image in RGB color-space as A(x,y)
- Convolution with the kernel h(x,y) of size 2*2/6 is applied to A(x,y) gives convoluted output C(x,y)
- Original input image is added with convoluted image to suppress the high values in the three channels of the RGB image as R(x,y) = A(x,y)+C(x,y)
- The resultant R(x,y) is binarized using OTSU global thresholding
- To get the binary nuclei mask the small areas are identified as Ar(x,y) using regional property Area of the white objects in the binary image R(x,y) while areas of the nuclei are find out using top hat mean.
- Small areas identified are removed from R(x,y) by first taking complement of Ar(x,y) and then subtracting it from R(x,y) as R(x,y) = R(x,y)-Ar(x,y)e.g:1-1 = 0
- After removing the small areas only the nuclei of the leukocytes are left as largest areas as R(x,y) e.g:1-1 = 0 called binary nuclei mask.
- Morphological operation dilation is applied to enlarge the segmented nuclei of the leukocytes with fixed structuring element of size 4 with disk shape is applied to the binary nuclei mask R(x,y) to get R'(x,y).
- Further the binary nuclei mask R'(x,y) is converted to RGB R''(x,y) shows clearly visual nuclei of the leukocytes

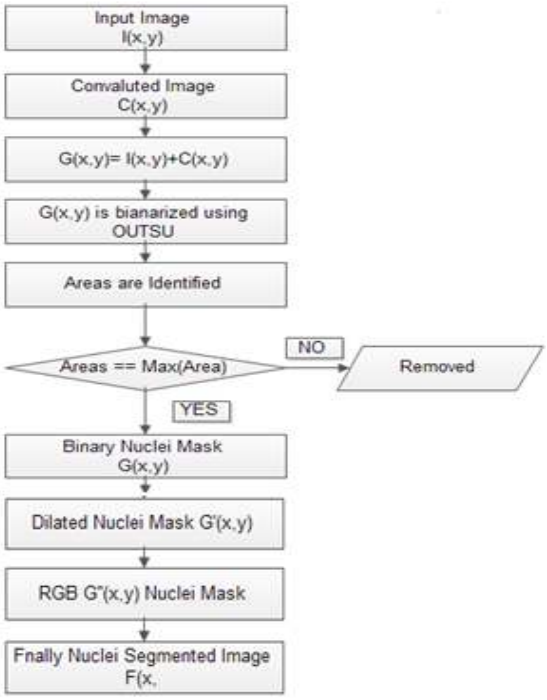

Fig. 1: Steps of proposed algorithm to segment nuclei of leukocytes

R/C	R/C	R/C
R/C	R/C	R/C
R/C	R/C	R/C

Fig. 2: Convolution kernel of size 3*3

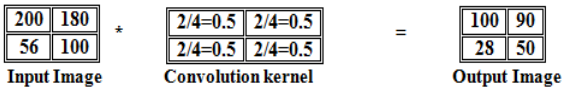

Fig. 3: The convolution and its effect on the input image

Fig. 4: The addition of original Input image with the convoluted image

- Finally the R''(x,y) is added with the original input image A(x,y) to show the segmented nuclei in the original image

Figure 1 shows the steps of the proposed algorithm along with output images at each step.

Details of the proposed algorithm: As mentioned earlier, taking advantage of the first plus point of the nuclei of the leukocytes, the input RGB image in step1 is convoluted in step 2 with the window of fixed size of R*C divided by a constant H as shown in the Fig. 2.

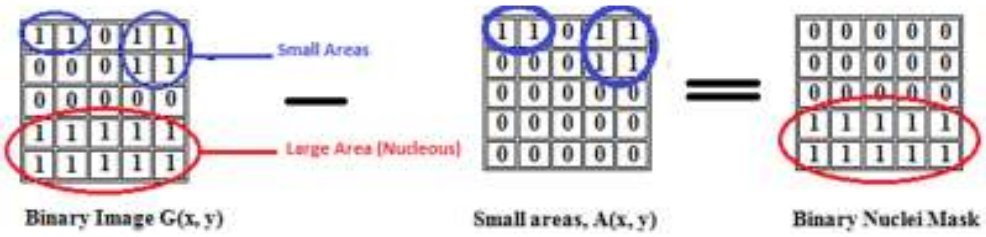

Fig. 5: Example of small areas removal

As the convolution is a local operation and is done in the way that a window of some fixed size is scanned across the image, the output pixels are the weighted sum of the input pixels in the window where the weights are the values of the filter assigned to each and every pixel of the window itself. Mathematically, the convolution for a discrete signal is given with equation as shown:

$$g[m,n] = I[m,n] \oplus h[m,n] = \sum_{i=-i_o}^{i_o} \sum_{j=-j_o}^{j_o} h[i,j]a[m-i,n-j]$$

The main idea behind the convolution in the proposed method is that it made the values lower down in all the three channels of the RGB image and when the original input image is added with convoluted image the high values become 255 means white while the low intensity values in all the three channels of the RGB while become brighter but not disappears as in this way we are able to suppress the high intensity values in the three channels of the RGB image as mentioned in the Fig. 3:

After applying convolution the resulted image is added with the original input image in step 3 which results in the suppressing of high values in all the three channels of the RGB image as shown in Fig. 4:

Further, we have the image with white background and a nuclei appears as more clear than in the original image. Moreover, if non-*In-vitro* slides are used also the output image will contains platelets as they are also low in intensity values in the three channels of the RGB image but due to the second benefit of the nuclei of the leukocytes we will select the largest area as our area of interest and platelets will be removed. Further the output Image is binarized with the OTSU thresholding to apply morphological operation on the output image of step 3 to make the nuclei clear. OTSU performs clustering based thresholding which proves that it minimizes the interclass variance and in the same way it is possible to maximize the interclass variance as expressed in the equation as terms of class probabilities and class means and can be updated iteratively mentioned in the equation:

$$\sigma_a^2 = \sigma^2 - \sigma_\omega(i) = \omega_1(i)\omega_2(i)\,[\mu_1(i) \quad \mu_2(i)]^2$$

After getting the binary image in step 4, small and large areas are identified using the regional properties of the bright object in the image in step 5. As taking advantage of the second plus point of the nuclei of leukocytes largest areas are considered as nuclei of the leukocytes, while small areas are removed by subtracting the small areas complemented image of step 5 from the binary image of step 4 in step 6 and the resultant image is called as binary nuclei mask obtained from step 7, the subtraction process is shown in Fig. 5:

Table 1: Number of leukocytes images used in testing

Type of Leukocyte	Number
Basophile	06
Eosinophil	40
Lymphocyte	37
Monocyte	21
Neutrophil	276
Total Images used	380

The binary nuclei mask is enlarged using the morphological operator, dilation in step 8 and we get the dilated Nuclei Mask.

Morphological dilation is applied with a structuring element D on a binary image G as presented in the equation:

$$G \oplus D = \{ z \epsilon I | (D^e) \epsilon Z \cap G \neq \emptyset \}$$

where, D^e is the structuring element and is defined by:

$$D^e = \{x \in I | -x \in D\}$$

where, I is an integer grid.

Finally the binary nuclei dilated mask of step 8 is converted to RGB color-space in step 9 while in step 10 the RGB nuclei mask is added with the original image of step 1 to segment the nuclei in the original input image.

Microscopic blood image database: Peripheral blood smears stained by Gimesa staining technique were taken for the experimentation. The regular light microscope with 100x objective lens is used for blood slides to obtained the digital images with VGA resolution of (640*360) through a CCD color camera attached to microscope in Ihsan Medical Complex, Mardan, Pakistan in the presence of expert Dr. Bakht Biland but only 15 images were obtained. The rest 365 images were obtained from the dataset published by Mohamed and Far (2012b) in MATLAB file exchange with the permission of comparison and reproduction. Table 1 shows the different types of leukocytes and quantity in database.

Table 2: Percentage accuracy of segmentation

	Leukocytes					
Technique	Basophil	Eosinophil	Monocyte	Lymphocyte	Neutrophil	Total
Proposed	83.33	87.5	81.08	90.47	97.82	96.13

RESULTS AND EVALUATION

The evaluation has been made in three ways with the corresponding results for ground truth:

Evaluation of proposed method for different types of leukocytes: The different types of leukocytes have different morphology, color and size as shown in Fig. 6, due to these variations the evaluation of the proposed method has been done by comparing the automatic segmented leukocytes with the manually identified leukocytes by an expert. The metric used for the evaluation is:

For a single image:

$$S_{automatic} = 1, if\ S_{automatic} = S_{maual}$$

$$S_{manual} = 1, if\ S_{automatic} \neq S_{mauual}$$

For a class of images:

$$S_{M=} \frac{Total(S_{automatic})}{Total\ number\ of\ Images\ in\ class} * 100$$

where,

$$S_{automatic} = Automatic\ Segmented\ Image$$

$$S_{manual} = Manually\ marked\ Image$$

$$S_{Matric} = \%\ age\ Segmentation\ Metric$$

The S_{Matric} technique is best qualitative technique for finding the accuracy in percentage. The complete evaluation results of the different types are given in Table 2 and graphically plotted in Fig. 7.

Results assessment for all leukocyte's types: It is superiority of the presented technique that an overall segmentation accuracy of 96.13% was achieved in which the highest was recorded as 97.82% for the Netruophils. The tested sample of the Neutrophils segmentation is shown in Fig. 8:

The lowest segmentation accuracy was recorded as 81.08% for the Monocytes because the monocytes have spots which deviate in some slides from the criteria made for the proposed method. The Monocyte segmentation sample is presented in Fig. 9.

For the rest of the types the segmentation accuracy ranges from 83.33 to 90.47% and the results are presented in Fig. 10 to 13.

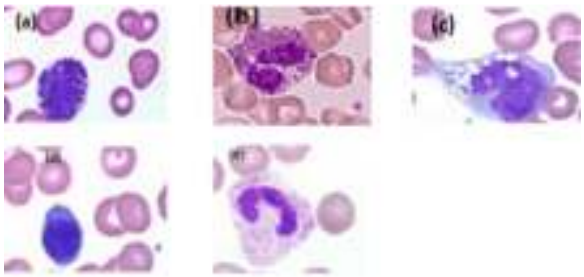

Fig. 6: Leukocytes, a) Basophil, b) Eosinophil, c) Monocyte, d) Lymphocyte, e) Neutrophil

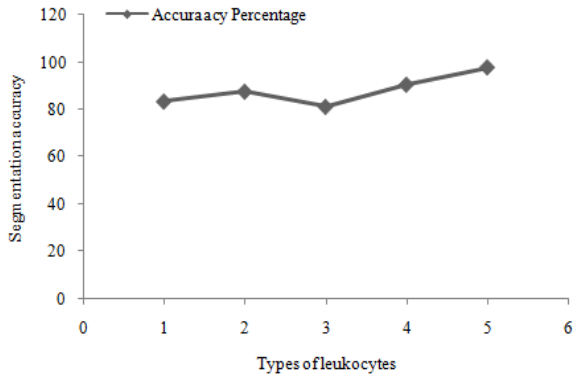

Fig. 7: Plotted segmentation accuracy of the proposed method

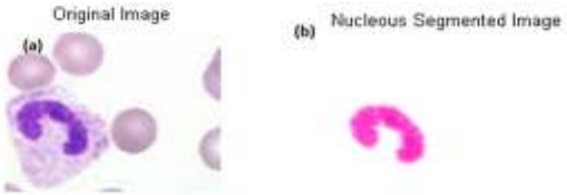

Fig. 8: Neutrophil sample, a) Original input image, b) Nucleus segmented image

Fig. 9: Monocyte sample, a) Original input image, b) Nucleus segmented image

Performance evaluation: Performance of the proposed method was evaluated in two ways i.e., with respect to percentage segmentation accuracy with five other

Fig. 10: Basophil sample, a) Original input image, b) Nucleus segmented image

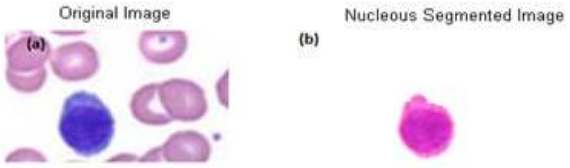

Fig. 11: Lymphocyte sample, a) Original input image, b) Nucleus segmented image

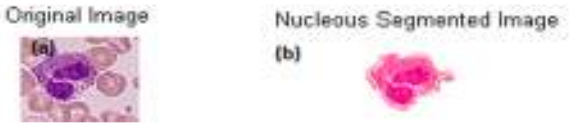

Fig. 12: Eosinophil sample, a) Original input image, b) Nucleus segmented image

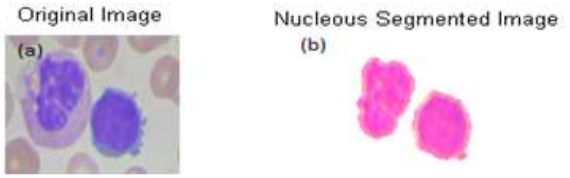

Fig. 13: Segmentation of two cells neutrophil and lympocyte, a) Original image, b) Nuclei segmented image

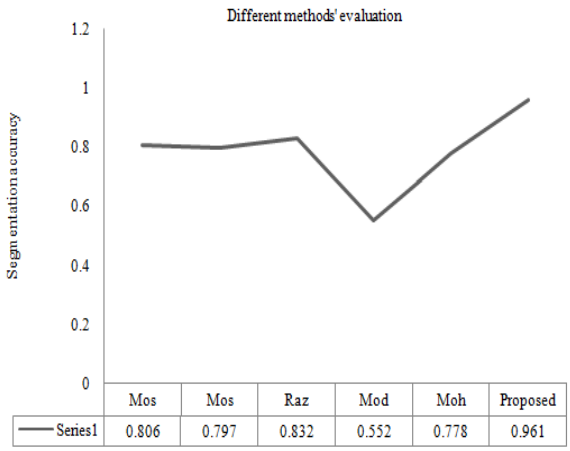

Fig. 14: Total segmentation accuracy of each method and its comparison

proposed methods and in the same way with respect to average segmentation processing time with the same number of other proposed methods. The five

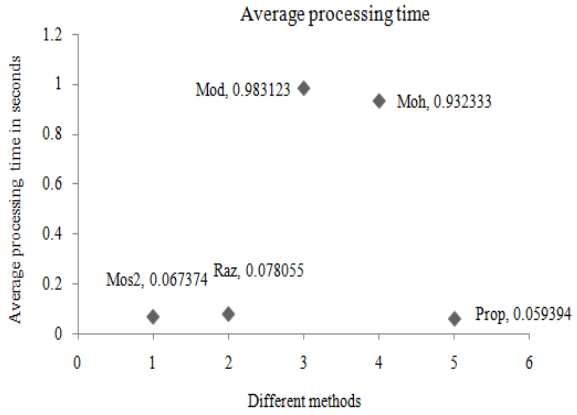

Fig. 15: Average processing time of different compared methods with the proposed method plotting

other methods were proposed by Mohamed and Far (2012a, b), Ramoser *et al*. (2006), Madhloom *et al*. (2010) and Mohamed *et al*. (2012). The metric used for evaluation among different method is:

$$S_{ammm} = \frac{S_{automatic} \cap S_{manual}}{\max(S_{automatic}, S_{manual})}$$

where,

S_{ammm} = Segmentation accuracy multi-methods metric
$S_{automatic}$ = Automatically segmented Nuceli images
S_{manual} = Manually marked images by expert

The segmentation accuracy is recorded in Table 3 and plotted on the basis of overall accuracy of each method in Fig. 14.

The processing time of different algorithms is calculated on the basis of processing time of each method for each time then calculating its average as an overall processing time. The developed method is superior in the performance as well as in efficiency as recoded overall average processing time for 380 images of different types 0.049391 sec. The proposed methods is compared with the other methods on the same set of images found much more faster as compared to other methods as their average processing times are recorded and plotted in the graph as Fig. 15.

Limitations: As the method achieved 96.13% overall accuracy and much faster as compare to other methods but the results are slightly affected as over segmentation due to very large clumps of cells having high degree of overlap and also due to improper slide formation as thin smear. However, the accuracy is not disturbed up to a clump of three cells but higher clumps with high degree of overlap will results in the low intensity and large area of the cells and in the same way when the blood is not properly spread on the slide results in low intensity of red blood cells as shown in Fig. 16.

Table 3: Evaluation of various methods on the basis of different types of Leukocytes

	Types of leukocytes				
Technique	Basophil	Eosinophil	Monocyte	Lymphocyte	Neutrophil
Mos, (Madhloom *et al*., 2010)	0.81	0.71	0.845	0.859	0.812
Mos, (Mohamed and Far, 2012a, b)	0.786	0.901	0.783	0.830	0.585
Raz, (Tabrizi *et al*., 2010)	0.757	0.889	0.796	0.896	0.823
Mod, (Theera-Umpon and Dhompongsa, 2007)	0.623	0.435	0.620	0.576	0.559
Moh, (Mohamed *et al*., 2012)	0.776	0.6619	0.897	0.781	0.787
Proposed	83.335	87.551	81.082	90.473	97.821

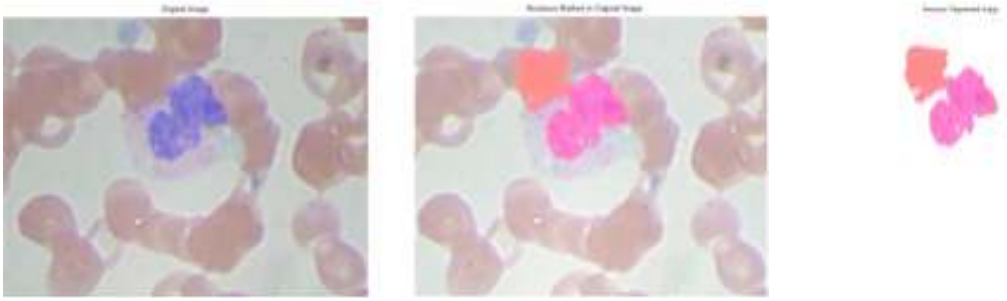

Fig. 16: Over segmentation due to large clumps and improper slide formation

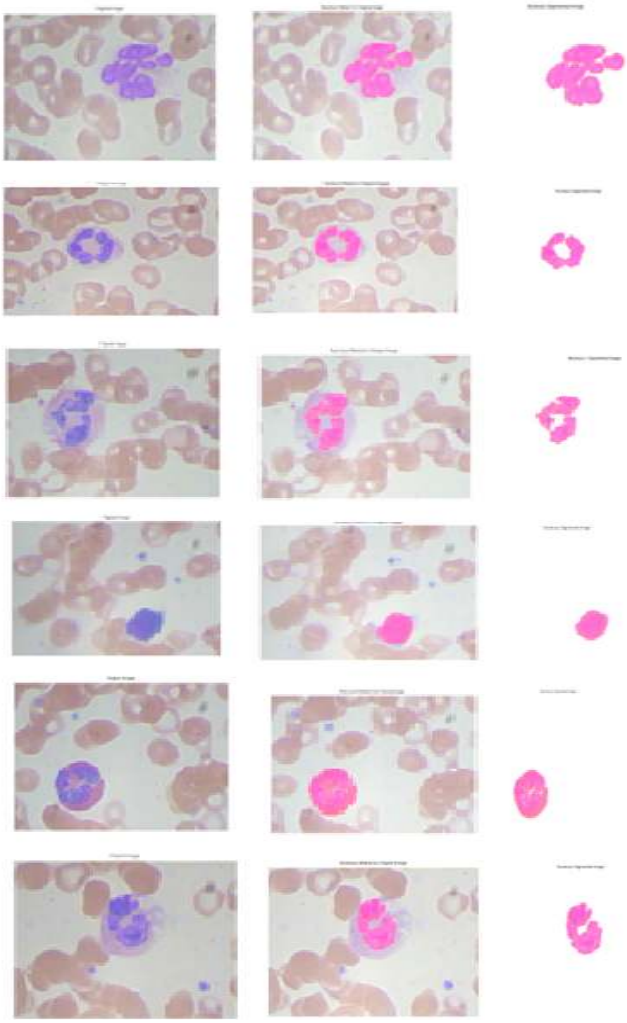

Fig. 17: Nuceli segmentation of different types

CONCLUSION

In this study we introduced a completely new method/algorithm for the segmentation of nuceli of Leykocytes, which achieved an overall segmentation accuracy of 96.13% and also efficient in processing time as compared to other published methods. The areas of Leykocyte's neuclei are the largest and found out through top hat mean to reduce the jeopardy of error. Moreover, the algorithm is carefuly tested through a dataset of 380 images and the results are recorded and in the same way compared on the same set of images with other methods. The higher segmentation accuracy is recorded for the Neutrophil as 97.82% because mostly the slides for Neutrophils have a very low degree of overlapped cells while the lower segmentation accuracy is recorded for the type Monocytes 81.08% because the slides provided for this type have high degree of overlapped cells and also they are overlapped with the Lekocytes. As a futre work the research still has room for improvement of segmentation accuracy in all the types as well as in overall segmentation accuracy. The work in the next phase on the same graound is possible to achieve segmentatiom accuracy of 97% to 99% with more efficiency. Also the futre work will have a strong demand for finding out a universal solution to the problem (Fig. 17).

ACKNOLEDGMENT

I would like to acknowledge my supervisor Prof.Dr. Dzulkifli Mohamad Head of the ViCube Reearch Lab. faculty of Computing, University Technology Malaysia, Johar Bahru Skudai Campus, Malaysia. It is all because his keen interest in my study and gauidence at each and every step of designing the algorithm. I am also thankful to Prof. Dr. Bakht Biland, Director Ihsan Medical Complex Mardan, KPK, Pakistan for his guidence in the medical field and marking the nuclei of Leukocytes in his very busy schedule.

REFERENCES

Abbas, N. and D. Mohamad, 2013. Microscopic RGB color images enhancement for blood cells segmentation in YCBCR color space for k-means clustering. J. Theor. Appl. Inform. Technol., 54(1).

Dorini, L.B., R. Minetto and N.J. Leite, 2007. White blood cell segmentation using morphological operators and scale-space analysis. Proceeding of the 20th Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI, 2007), pp: 294-304.

Madhloom, H., S. Kareem, H. Ariffin, A. Zaidan, H. Alanazi and B. Zaidan, 2010. An automated white blood cell nucleus localization and segmentation using image arithmetic and automatic threshold. J. Appl. Sci., 10: 959-966.

Mohamed, M. and B. Far, 2012a. An enhanced threshold based technique for white blood cells nuclei automatic segmentation. Proceeding of the IEEE 14th International Conference on e-Health Networking, Applications and Services (Healthcom), pp: 202-207.

Mohamed, M.M.A. and B. Far, 2012b. A fast technique for white blood cells nuclei automatic segmentation based on gram-schmidt orthogonalization. Proceeding of the IEEE 24th International Conference on Tools with Artificial Intelligence (ICTAI, 2012), pp: 947-952.

Mohamed, M., B. Far and A. Guaily, 2012. An efficient technique for white blood cells nuclei automatic segmentation. Proceeding of the IEEE International Conference on Systems, Man and Cybernetics (SMC, 2012), pp: 220-225.

Nordqvist, C., 2009. What Is Leukemia? What Causes Leukemia? Retrieved from: http:// www.medicalnewstoday.com/articles/142595. php. (Accessed on: December 12, 2013)

Ongun, G., U. Halici, K. Leblebicioglu, V. Atalay, M. Beksac and S. Beksac, 2001. An automated differential blood count system. Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 3: 2583 - 2586.

Ramoser, H., V. Laurain, H. Bischof and R. Ecker, 2006. Leukocyte segmentation and classification in blood-smear images. Proceedings of the 27th Annual International Conference of the Engineering in Medicine and Biology Society (IEEE-EMBS, 2005), pp: 3371-3374.

Rezatofighi, S., H. Soltanian-Zadeh, R. Sharifian and R. Zoroofi, 2009. A new approach to white blood cell nucleus segmentation based on gram-schmidt orthogonalization. Proceedings of the IEEE International Conference on Digital Image Processing.

Tabrizi, P., S. Rezatofighi and M. Yazdanpanah, 2010. Using PCA and LVQ neural network for automatic recognition of five types of white blood cells. Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC, 2010).

Theera-Umpon, N. and S. Dhompongsa, 2007. Morphological granulometric features of nucleus in automatic bone marrow white blood cell classification. IEEE T. Inform. Technol. Biomed., 11(3): 353-359.