REPORT

Nuclei segmentation of leukocytes in blood smear digital images

Naveed Abbas*[1], Dzulkifli Mohamad[1], Abdul Hanan Abdullah[1], Tanzila Saba[2], Mznah Al-Rodhaan[3] and Abdullah Al-Dhelaan[3]

[1]ViCube Research Lab, Department of Software Engineering, Faculty of Computing University Technology Malaysia
[2]College of Computer and Information Sciences (CCIS) Prince Sultan University Riyadh, KSA
[3]Computer Science Dept., College of Computer & Information Sciences, King Saud University, Riyadh, KSA

Abstract: The Leukocytes are differentiated from each other on the basis of their nuclei, demanded in many Medical studies, especially in all types of Leukemia by the Hematologists to note the disorder caused by specific type of Leukocyte. Leukemia is a life threatening disease. The work for diagnosing is manually carried out by the Hematologists involving much labor, time and human errors. The problems mentioned are easily addressed through computer vision techniques, but still accuracy and efficiency are demanded in terms of the basic and challenging step segmentation of Leukocyte's nuclei. The underlying study proposed better method in terms of accuracy and efficiency by designing a dynamic convolution filter for boosting low intensity values in the separated green channel of an RGB image and suppressing the high values in the same channel. The high values in the green channel become 255 (background) while the nuclei always have low values in the green channel and thus clearly appear as foreground. The proposed technique is tested on 365 images achieving an overall accuracy of 95.89%, while improving the efficiency by 10%. The proposed technique achieved its targets in a realistic way by improving the accuracy as well as the efficiency and both are highly required in the area.

Keywords: Leukocytes, nuclei segmentation, medical applications, convolution, leukemia.

INTRODUCTION

Blood is composed of Leukocytes, Erythrocytes and Platelets (Ramoser *at el.,* 2006). The different components of blood have different responsibilities like Leukocytes are responsible for the immune system; Erythrocytes are responsible for oxidation and platelets are responsible for clotting. The blood tests in case of many diseases like Leukemia, Malaria, Multiple Myeloma, dengue fever and in many other diseases are frequently suggested by the physicians to know the disorder in the blood cells or to note their number. Any malfunction due to attack of parasites in the blood cells and the number of blood cells is a crucial indicator about many diseases. The number of blood cells in both cases, i.e. Reduction and abundance give important insights about fatal diseases like in Leukemia blood Cancer, Bone Marrow Cancer etc. the leukocytes are called leukemia cells and are made in abundance in the bone marrow (Nordquist, 2009). This abundance of leukemia cells is due to the disorder in a specific type of leukocytes. The identification of that specific type is a challenging job because their ambiguous appearance. This study is intended to automate the basic and challenging step of segmentation demanded in various medical studies for automating various processes, but here specifically we applied the developed technique in segmentation of

Leukocytes' nuclei. The Leukocytes are mainly divided into five types, i.e. Basophil, Eosinophil, Monocyte, Lymphocyte and Neutorphil as presented in (fig. 1) and also mentioned in the study of (Tabrazi *at el.*, 2010). These Leukocytes are identified by the medical experts through their nuclei to know about the disorder in the specific type of Leukocytes and prescribe suitable medicine according to the type and severance of the situation. The Hematologists perform the tests on thin blood smears, prepared for the purpose of detailed study of the blood components for any disorder under light microscope, which is Gold standard, but time consuming, laborious and has jeopardy of too much human errors also highly dependent on the expert's knowledge. This study contributes the area in its first and most challenging process, i.e. segmentation of Nuclei of Leukocytes by automating the process through the use of computer vision concepts. The mentioned process is already addressed by many previous studies, but the underlying paper is intended to address it on the basis of improving the accuracy and efficiency.

The majority of the previous approaches, achieved high segmentation accuracy and efficiency, but still there is a wide space for improvement. The study made by (Dorini *et al.*, 2007) employs watershed transform for the segmentation of nuclei of leukocytes in gray scale images and then extract the cytoplasm on the Granulometry measure of the size of leukocytes. They also used some

*Corresponding author: e-mail: naveed23a@yahoo.com

morphological operators, but the results are validated by visual inspection on image set of only two images as the problem we faced in our study previously that there are a lot of variations in the size and shape of nuclei. The work by the same authors with the only change in the number of images in the results was carried out in (Dorini *et al.*, 2012). However, both of the studies of (Dorini *et al.*, 2007) and (Dorini *et al.*, 2012) are highly dependent on the size and shape of nuclei of the leukocytes. In the study of (Chan *et al.,* 2010), the authors used edge-based segmentation through a weighted Sobel operator in the grayscale images. The results were presented with two images while also some calculation had been made for finding the number of lobs in the nucleus. The results on the experimented image were good, but there was a lack of universality in the method. The Global threshold OTSU and morphological operator opening is used in the work of (Mohammad *at el.,* 2012a) and create a nucleus mask, but this approach will highly depend on the size of the nuclei of the leukocytes, having much variation but still the work achieved an overall segmentation accuracy of 90.1%. The work carried out by (Mohammad and Far, 2012b), the authors used additional step of Gram Schmidt Orthogonalzation for improving the processing time but we experience the same time as in previous studies (Mohammad *at al.,* 2012a) of the same authors. The same authors carried out the solutions in (Mohammad and Far, 2012b) on the same basis with addition of minimum filters of kernel size 3*3 and achieved an overall segmentation accuracy of 80%. The authors mentioned the problem of overlapped nuclei in (Mohammad *at el.,* 2012a), (Mohammad and Far, 2012b) and (Mohammad and Far, 2012c) but the actual problem is the variation in the size and shape of nuclei of leukocytes while the problem of overlapping can be easily solved through the method of (Abbas *et al.*, 2014a) or (Abbas *et al.*, 2015). The more suitable for overlapping is (Abbas *et al.*, 2014b), because based on boundaries, independent of morphology of the nuclei. The authors on the work carried out by them in (Ramesh *et al.,* 2012) used two-step classifications of leukocytes. First, they did segmentation of the nuclei in the converted image from RGB to JCV color-space through cropping by manual method. In the second step they did classify the leukocytes. The work carried out by (Rezatofighi *et al.*, 2009), the authors stated that their work was new and in the same way used by (Mohammad and Far, 2012c), they used the Gram Schmidt Orthogonalzation with difference that they further did classification of the leukocytes on the basis of segmented nuclei with the mentioned method and achieved overall segmentation accuracy of 93.02%. However, the method adopted will be affected by the variation in shape and size of nuclei of leukocytes. The initial work performed by the authors of (Abbas and Mohamad, 2014) proposed a unique and robust approach of color nuclei segmentation of leukocytes through a simple method of suppressing the high color values in the

three channels because the nuclei always in the low colors due to the staining process. However, achieved an overall segmentation accuracy of 96.04%, but affected when the erythrocytes were densely overlapped while also processing the three channels will be time consuming in massive testing. The study of color processing of leukocytes microscopic images in the YCbCr color-space gives a unique identity to the nuclei of the leukocytes in terms of colors, mentioned in the work (Abbas and Mohamad, 2013) and is very easy for segmentation especially in the case of binning algorithms.
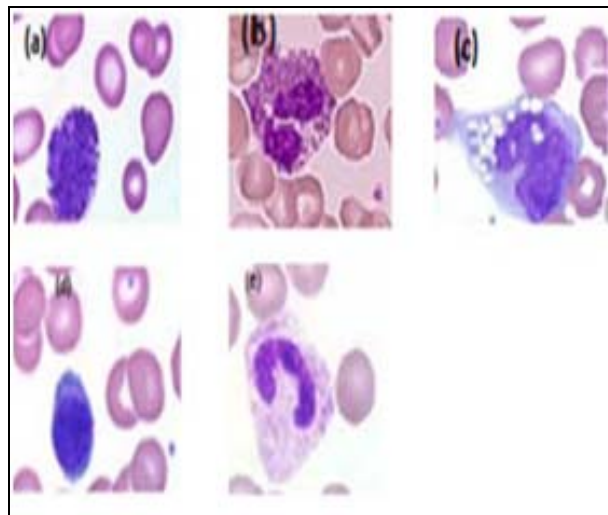

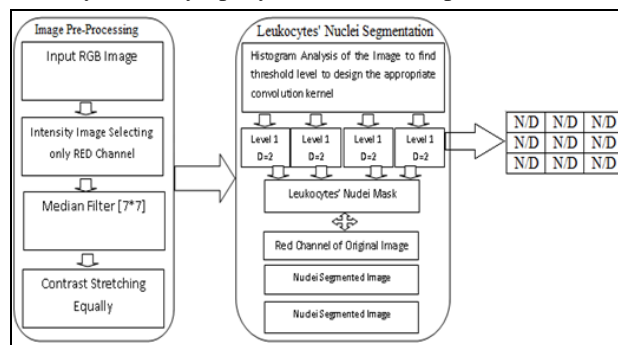**Fig. 1**: Leukocytes, a) Basophil, b) Eosinophil, c) Monocytes, d) Lympocyte and e) Neutrophil

Fig. 2: Block diagram depicted overall methodology of the process

MATERIAL AND METHODS

The proposed methodology for the segmentation of Leukocytes' nuclei starts from image pre-processing as the blood smear images are highly subjective to noise and contrast problems, thus we reduce the noise with median filter of kernel size of [7*7] while stretch the contrast equally. In the proposed method we designed a dynamic convolution filter of size [3*3] window based on the threshold level effectiveness metric (EM) for its dynamicity. The convolution filter will emboss lower intensity values in the RED channel (we selected the RED

channel for the purpose that the intensity values of the Erythrocytes are high in this channel due to their Red color and this study is intended deeply in this property of any other object then Leukocytes' Nuclei) of the input RGB image while suppress the high intensity values to 255. After filtering the mask of Leukocytes nuclei is resulted that is converted to binary mask. The original intensity image (Red Channel of the original Input RGB image) is also converted according to the thresholding level resulted from the analysis of the histogram in the designing of convolution filter. The original binary image and the binary mask are added to each other, resulted in the nuclei of the Leukocytes as foreground and the rest of the objects as background. The overall methodology and the window of dynamic convolution are depicted in (fig. 2) while the conceptual design of the whole process is depicted as (fig. 3).
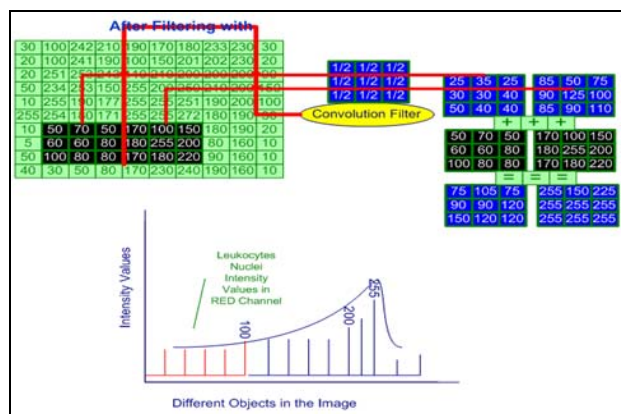

Fig. 3: Conceptual design of the idea behind the leukocytes' nuclei segmentation

Algorithm
Input: Pre-processed grayscale (Red channel) of the RGB Microscopic thin blood smear Image I
Output: Binary images ← Nuclei segmented
Begin:
\\Design the Convolution kernel
1.C ← *Ones* ← *3*3*
\\To find the denominator of the convolution kernel
2.D ← *Consult the Histogram*
\\Decide the denominator K of the kernel C for each cell on the basis of D in the mentioned four levels
3.F ← *Filter I with the convolution kernel C*
\\Add the Original Image I with the filtered F
4.P ←*I+F*
\\Resultant grayscale Image has Leukocyte's Nuclei
\\Conversion to binary through level D
5.BWP ← *Binary P with level D*
\\Show BW as parasites segmented image
\\Convert to binary the original image with the same threshold D
6. BWI ← *Binary I with level D*
END

RESULTS

The results are obtained from the experimentation on the digital image dataset provided by (Muhammad, 2012) that has also a dataset of the ground truth images marked by two medical experts Dr. Yaser Hasan and Dr. Mohamed Albasher from school of medicine, Cario University were taken through an Hp Laser jet color CP6015 dn (Q3932A) printer. The image dataset was further verified by Dr. AmreekLal, Dr. Ikram-ul-Mabood, Dr. Imran and Dr. Mukamil Shah for proof marking of nuclei again and they found the previously marked image dataset accurate. The proof observer in the same way marked 365 microscopic digital images as marked by the previous Medical experts. Moreover, in evaluation the marked image dataset was converted to binary with the automatic thresholding mentioned by (Ridler and Calvard, 1978) and matched with the automatic segmented nuclei of leukocytes by the proposed method for calculation of the segmentation accuracy, sensitivity or recall, precision and F-measure of the recall and precision. The results are presented through visual inspection along with ground truth images and also through statistical metrics to note the segmentation accuracy for comparison with other methods in the area while finally on the basis of average processing time to note the efficiency improvement and compared with other methods in the area.

Table 1: Evaluation of the proposed method on the basis of comparing with ground truth data

Types of Leukocytes	Recall	Precision	F-Measure
Basophile	0.8922	0.9411	0.91599784
Eosinophil	0.9222	0.9222	0.9222
Lymphocyte	0.9111	0.9623	0.936000352
Monocyte	0.8911	0.9284	0.909367672
Neutrophil	0.9211	0.9333	0.927159868
Overall	0.8955	0.9216	0.908362556

Visual inspection based comparisons with the ground Truth
Visual inspection based comparison involved the results in the form of images in that the nuclei are segmented through the proposed technique with the ground truth data marked and verified by the medical experts. The visual inspection based results are presented in (fig. 4).

According to (fig. 4), each row representsa different single image, (1) Input original RGB image, (2) Red Channel of the RGB (enhanced image), (3) is the ground truth image marked and verified by the mentioned medical experts and finally the (4) is the Leukocytes' Nuclei segmented image through the proposed method. Some small areas are also segmented due to the same intensity values as that of the Leukocytes' nuclei that are eliminated from the final image to overcome the problem

Pak. J. Pharm. Sci., Vol.28 No.5, September 2015, pp.1801-1806

1803

Table 2: Comparison of the proposed technique with other techniques in the area

Techniques	Types of Leukocytes					
	Basophile	Eosinophil	Lymphocyte	Monocyte	Neutrophil	Overall
(Rezatofighi *at el.*, 2009)	0.757	0.889	0.796	0.896	0.823	0.832
(Mohamed and Far, 2012c)	0.786	0.901	0.783	0.830	0.585	0.584
(Mohamed and Far, 2012b)	0.81	0.71	0.845	0.859	0.812	0.806
(Abbas and Mohamad, 2014)	0.8333	0.8755	0.8108	0.9047	0.9782	0.9613
Proposed	0.8431	0.8823	0.8241	0.8915	0.9688	0.9589
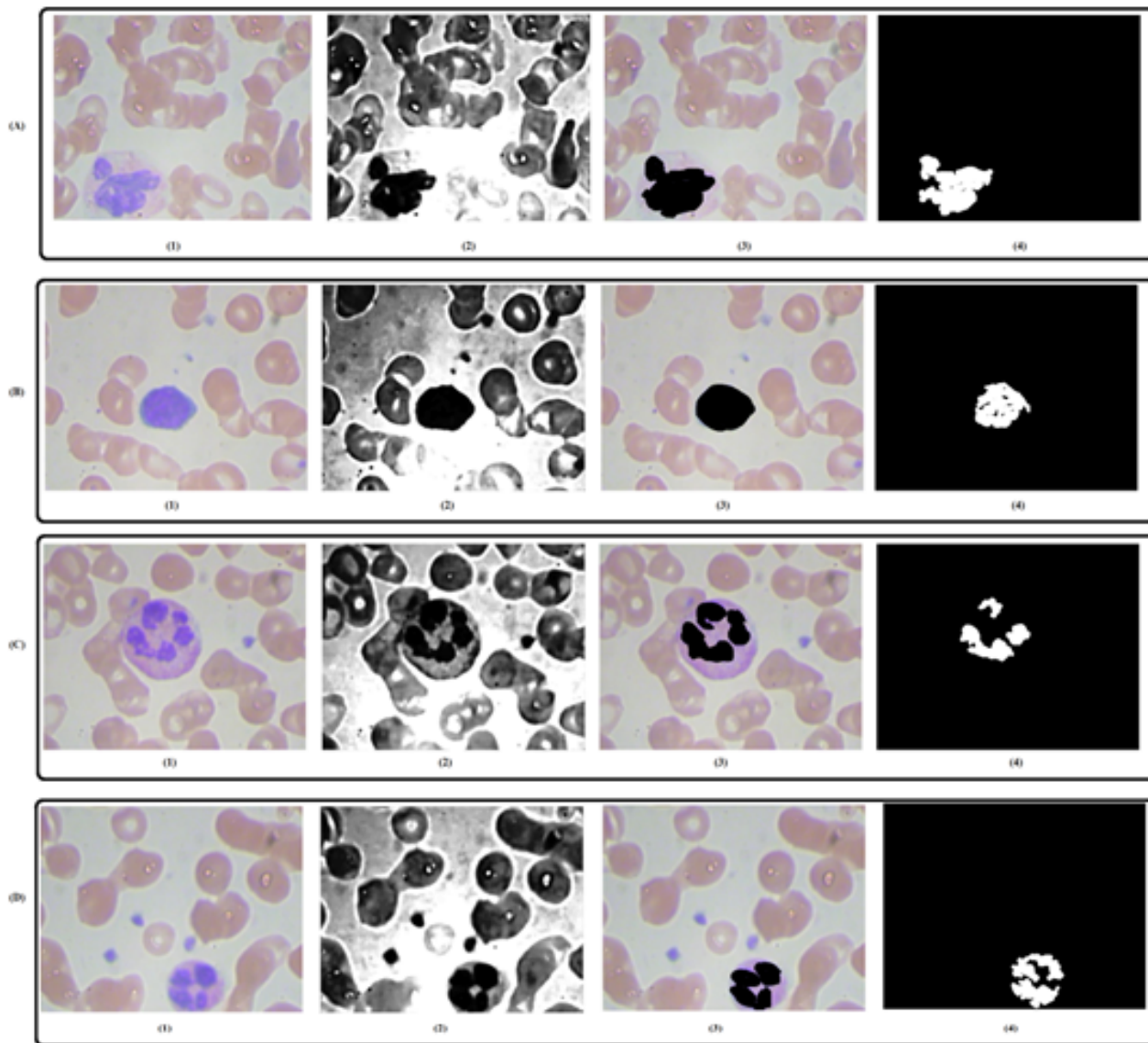
Fig. 4: Results of different input images along with ground truth data.

of over segmentation. More, results are given at the end as (fig.6), (fig.7), (fig.8) and (fig.9).

Statistical metrics based analysis of the segmentation Accuracy

The statistical metrics involved in measuring the segmentation accuracy are Precision, Recall and F-measure, given as equations (1), (2) and (3) and are calculated for each type separately as well as overall and presented in (table 1).

$$Sensitivity \, or \, Recall = \frac{|BIS_{auto} \cap BIM_{manual}|}{|BIS_{auto}|} \quad (1)$$

$$Precision = \frac{|BIS_{auto} \cap BIM_{manual}|}{|BIM_{Manual}|} \quad (2)$$

$$F - Measure = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall} \qquad (3)$$

Comparison based on efficiency

In this comparison, the proposed technique is compared with the same number of images used by the mentioned state-of-the-art techniques based on average processing time to measure the efficiency, which is also one of the demanding hypotheses of the underlying study. The results are depicted in graphical form mentioned in fig. 5. According to table 1, the segmentation accuracy is not improved too much, but still considered as a contribution in the area on such accuracy with high efficiency. The improvement in efficiency is due to reducing the processing to a single channel.
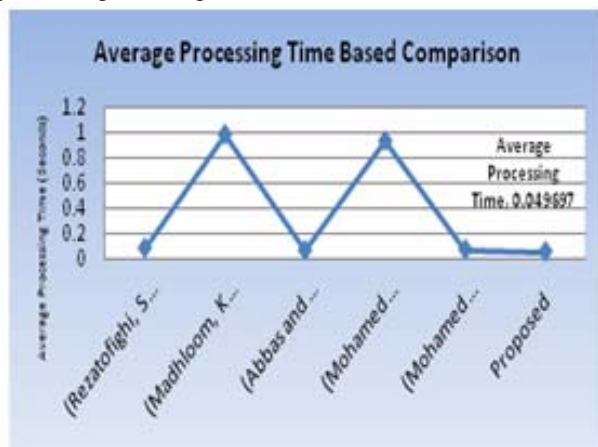

Fig. 5. Average processing time based comparison of the proposed technique

Proposed method comparison with other methods in the area

The proposed technique is compared with state-of-the-art on the same grounds and number of images and presented in (table 2).

According to the graph presented as fig. 4, the proposed technique shows significant improvement in average processing time due to the use of a single channel of the RGB input image. However, the accuracy is not improved due to the selection of single channel of the RGB input image because the detailed information existed in RGB images as compared to a grayscale image.

DISCUSSION

The reduction of three channels in one has a drastic effect on the computational performance. The color processing has its own benefits that having more and more details, but here the objects of interest are large enough that conversion to single channel will have no effect while the processing time is reduced.

Due to the involvement of health we were too vigilant during the experimentation and experienced some cases

of over-segmentation. These cases of over-segmentation are due to the diver occlusion of erythrocytes. This problem can be easily solved if we follow the method of elimination of erythrocytes mentioned by (Naseemullah *et al.*, 2014) during the slide preparation process.

CONCLUSION

The proposed technique achieved the desired goals by improving the accuracy to 0.8955%, while trimming the processing by at least 25%. As both of these goals are highly demanded in the area because the accuracy is needed due to life involvement and efficiency due to a recommendation of the World Health Organization that to check or test such conditions in at least 70 to 100 windows (WHO, 1991). Moreover, still they're more gaps existed due to noise and illumination correction, as future work if the mentioned problems are properly addressed with more robust techniques than the accuracy will rise further.

ACKNOWLEDGEMENTS

REFERENCES

Abbas N and Mohamad D (2013). Microscopic rgb color images enhancement for blood cells segmentation in ycbcr color space for K-means clustering. *JATIT*, **54**(1): 117-125.

Abbas N and Mohamad D (2014a), Automatic color nuclei segmentation of leukocytes for acute leukemia. *Res. J. Appl. Sci. Eng. Technol,* **7**(14): 2987-2993.

Abbas N and Mohamad D (2014 b), Accurate Red Blood Cells Automatic Counting In Microscopic Thin Blood Smear Digital Images *Sci.Int.(Lahore)* **26**(3): 1119-24.

Abbas N, Mohamad D and Abdul HA (2015). Occluded red blood cells splitting via boundaries analysis and lines drawing in microscopic thin blood smear digital images. *VFAST Trans. Soft. Eng.*, **5**(1): 10-17.

Abbas N, Muhamad D and Abdul HA (2014). Semi-automatic clustered red blood cells splitting and counting in thin blood smear images. *1ˢᵗ IRICT Conference, University Technology Malaysia*, 342-351

Chan YK, Tsai MH, Duang DC, Zheng ZH and Hung KD (2010). Leukocyte nucleus segmentation and nucleus lobe counting. *BMC bioinformatics*, **11**(1): 558.

Dorini LB, Minetto R and Leite NJ (2007). White blood cell segmentation using morphological operators and scale-space analysis. IEEE XX brazilian symposium on

computer graphics and image processing. *SIBGRAPI.* pp.294-304.

Dorini LB, Minetto R and Leite NJ (2013). Semiautomatic white blood cell segmentation based on multiscale analysis. *IEEE J. Biomed. Health Inform.* **17**(1): 250-256.

Madhloom HT, Kareem SA, Ariffin H, Zaidan AA, Alanazi HO and Zaidan BB (2010). An automated white blood cell nucleus localization and segmentation using image arithmetic and automatic threshold. *J. Applied Sci.,* **10**(1): 959-966.

Mohamed M (2012). Image dataset with ground truth images and code. Retrieved on14-May-2013; Matlab File exchange from: http://www.mathworks.com/matlabcentral/fileexchange /36634-an-efficient-technique-for-white-blood-cells-nuclei-automatic-segmentation/content/WBC_SegProposed.m.

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

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

Mohamed MMA and Far B (2012). A fast technique for white blood cells nuclei automatic segmentation based on gram-schmidtorthogonalization. IEEE 24th International Conference on Tools with Artificial Intelligence (ICTAI), **1**: 947-952.

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

Ramesh N, Dangott B, Salama ME and Tasdizen T (2012). Isolation and two-step classification of normal white blood cells in peripheral blood smears. *J. pathol. Inform.*, **3**(13): 1-10.

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

Rezatofighi SH, Soltanian-Zadeh H, Sharifian R and Zoroofi RA (2009). A new approach to white blood cell nucleus segmentation based on gram-schmidtorthogonalization. IEEE International Conference on Digital Image Processing. pp.107-111.

Ridler T and S Calvard (1978), Picture thresholding using an iterative selection method. *IEEE Trans. Syst., Man, Cybern., Syst.,* **8**(8): 630-632.

Tabrizi PR, Rezatofighi SH and Yazdanpanah MJ (2010). Using PCA and LVQ neural network for automatic recognition of five types of white blood cells. Annual International Conference of the IEEE, Teheran,

Iran on Engineering in Medicine and Biology Society (EMBC) pp.5593-5596.

Ullah N, Khan MF, Jan SU and Mukhtiar M (2015). Study of the effect of inorganic and organic complexes of arsenic metal on the status of GSH in T cells and B. cells of blood. *Pak. J. Pharm. Sci.*, **28**(2): 457-464.

WHO (1991). Basic malaria microscopy: Part I. Learner's guide: Part II. Tutor's guide. (1[st]ed) World Health Organization: Switzerland, pp.1-77.
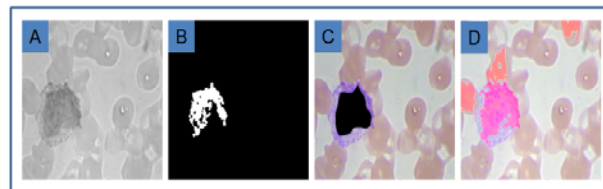
Visual Results

Fig. 6: Matlab results A) presents original image in the Red channel, B) presents segmented nucleus of leukocyte with proposed method, C) presents ground truth image, D) presents segmented nucleus (over segmentation) with (Abbas and Mohamad, 2014).
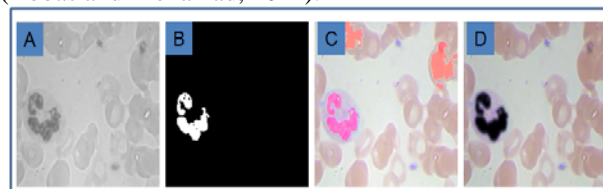

Fig. 7: Matlab results A) presents original image in the Red channel, B) presents segmented nucleus of leukocyte with proposed method, C) presents segmented nucleus (over segmentation) with (Abbas and Mohamad, 2014), D) presents ground truth image.
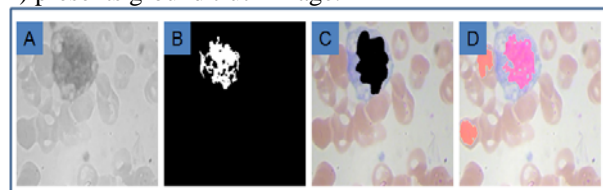

Fig. 8: Matlab results A) presents original image in the Red channel, B) presents segmented nucleus of leukocyte with proposed method, C) presents ground truth image, D) presents segmented nucleus (over segmentation) with (Abbas and Mohamad, 2014).
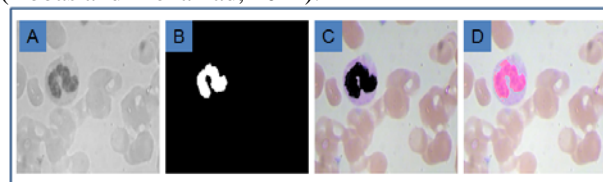

Fig. 9: Matlab results A) presents original image in the Red channel, B) presents segmented nucleus of leukocyte with proposed method, C) presents ground truth image, D) presents segmented nucleus (highly accurate) with (Abbas and Mohamad, 2014).

1806

Pak. J. Pharm. Sci., Vol.28 No.5, September 2015, pp.1801-1806