Performance Comparison of Segmentation Techniques for Nucleus in Chronics Leukemia

R.Adollah[1], M.Y Mashor[1], R. Hassan[2] and N.H Harun[3]
[1]*School of Mechatronic Engineering, Universiti Malaysia Perlis*
[2]*Hospital Univ. Science Malaysia, Kubang Kerian.*
[3]*Data Science Research Lab, School of Computing, Universiti Utara Malaysia*
robiyanti@unimap.edu.my

***Abstract*—Morphological criteria have been used by haematologists to identify malignant cells in the blood smear sample under a light microscope. Experienced hematologist must perform this screening operation. However, manual screening using microscope is time-consuming and tedious. Thus, an automated or semi-automated image screening and diagnosis system are very helpful. An ideal automated screening system will acts as a human expert during the procedure. To formulate this idea, there are few steps involves in this process which is the acquisition of image, image segmentation, features extraction and recognition of image data for further analysis in computer-based. However, segmentation of a region of interest is the most crucial task to extract features for further learning and diagnose. This paper represents two segmentation techniques and their performance comparison based on clustering approach which are k-means and moving k-means clustering algorithms. The segmentation process is performed on ten chronics leukaemia images. The performance of segmentation based on the proposed techniques was evaluated. The proposed segmentation techniques offer high accuracies of segmentation which is more than 97% for both techniques.**

***Index Terms*— Blood Images; Chronicleukemia; Component; Image Segmentation; Nucleus Segmentation; Segmentation Performance.**

I. Introduction

Leukaemia is a cancer of the bone marrow and blood. It is characterised by the uncontrolled accumulation of blood cells [1]. The majority of patients are found at a late stage of the disease. In people with leukaemia, the bone marrow produces abnormal white blood cells, called leukaemia cells. At first, leukaemia cells function almost normally. However, over time, as more leukaemia cells are produced, they may crowd out the healthy white blood cells, red blood cells, and platelets. This makes it difficult for the blood to carry out its normal functions. Leukaemia is divided into four categories: myelogenous or lymphocytic, each of which can be acute or chronic [2-4].

Generally, diagnosing leukaemia begins with a medical history, physical examination, complete blood count and bone marrow biopsy. The microscope screening of peripheral blood smear provides important qualitative and quantitative information concerning the presence of cancer. This operation has to be performed by experienced haematologists, which basically perform two main analyses. The first is the qualitative study of the morphology of the cells and second task involved quantitative approach by performing differential counting of white blood cells types [3-4]. Morphological criteria have been used by haematologists to distinguish malignant cells in the blood smear sample. Furthermore, laboratory diagnosis of chronic leukaemia in modern haematology practice is increasingly relying on guidelines that require the availability of relatively expensive machines with a consistent need for continuous quality control, kits supply and maintenance [4-5].

In order hand, manual screening of peripheral blood smear is time-consuming and tedious. Repetitive works by medical expert lead to exhaustion and tiredness. Thus, an automated or semi-automated image screening and diagnostic system very helpful. However, the success of an automated image screening and analysis mostly depends on the proper segmentation of images before morphological features of chronic leukaemia images can be extracted for computer based training and learning for further classification tools [6-8]. Various segmentation for image segmentation has been proposed. This study will utilise two algorithms for segmentation technique to be performed on chronic leukaemia slide images for segmenting the nucleus of white blood cell (WBC) as the interest of region (ROI).

II. Methodology

This study involves a few steps as illustrated in Figure 1. Images are captured under a light microscope, the captured image will be split into S-component based on HSI (hue, saturation, intensity) and this component will go through segmentation techniques. The accuracy of this segmentation will be evaluated and will be discussed in the next section.

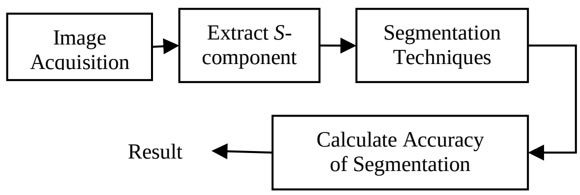
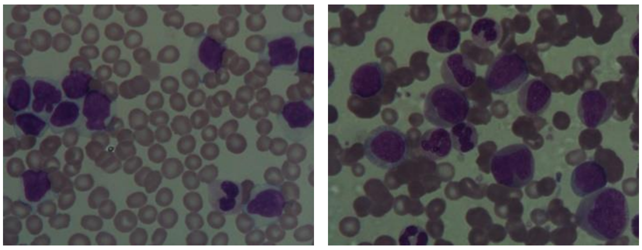


Figure 1: Block Diagram of Proposed Segmentation Procedure

A. Image Aquisition

In this study chronics, leukaemia slides were provided by the Hematology Department, *Hospital Universiti Sains Malaysia* (HUSM), Kelantan. Images of leukaemia slides were acquired using the *Luminera Infinity 2* digital camera mounted on the *Leica* light microscope. The slides were analysed under 40X magnifications. The images were captured at the resolution of 800×600 pixels, and they were saved in the bitmap (.bmp) file format. The digital images captured by the camera were verified by the haematologist to determine the appropriateness and type of the blood cell. Ten samples of leukaemia images were used in this clustering segmentation. Figure 2(a) and Figure 2(b) shows an example of an original image captured for chronic lymphoid leukaemia (CLL) and chronic myeloid Leukemia (CML).



(a) CLL (b) CML

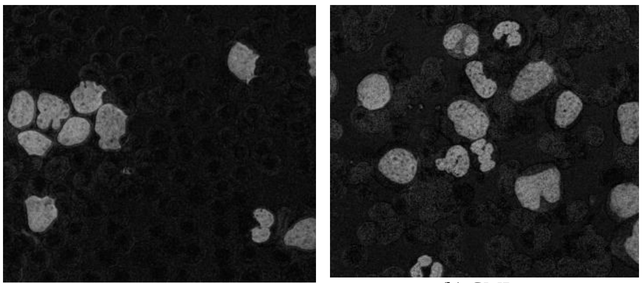
Figure 2: Chronic Leukemia Images

B. Saturation Components

The saturation component is a measurement of the degree of white light added to the pure colour. The mathematical equation of saturation formula that applied to the original chronics leukaemia images as given below [9]:

$$Saturation = 1 - \frac{3}{R+G+B}\min(R,G,B) \tag{1}$$

where, R, G, B and min are a red component, green component, blue component and a minimum of RGB component in colour space. In this study, clustering segmentation techniques were performed on saturation component. There is 3 cluster involves which is immature white blood cell or blast, red blood cell and background. Therefore, combination segmentation steps on the saturation component were chosen and proved by other researchers to reduce the computational effort and to ease the clustering process [9-10]. Figure 3 shows saturation images for (a) CLL and (b) CML.



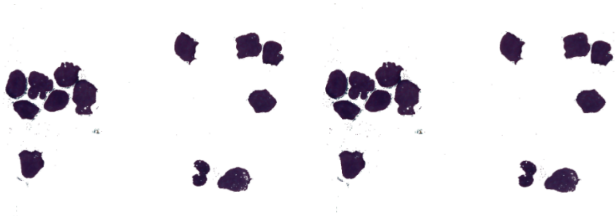
(a) CLL (b) CML

Figure 3: Saturation Component

C. Segmentations

Segmentation refers to the process of the partition where objects and background are separated into non-overlapping sets and only remains the object of interest at the end. Two types segmentation techniques are performed in this study which is based on k-means and moving k-means clustering algorithms. This segmentation algorithm was performed on saturation component of the images in HSI colour space conversion from RGB original image. k-means clustering (KM) segmentation algorithm step can be referred to [10]. The conventional K-mean clustering widely used for automatic image segmentation and do not always produce a good result due to dead centre, centre redundancy and trapped in local minima.

Another segmentation techniques applied in this study is moving k-means (MKM). A modified k-means algorithm version proposed by Mashor in the year 2000 [11]. This algorithm can minimise dead centres, centres redundancy problems and the effect of trapped in local minima problems. The method was originally used for positioning the centre of radial basis function (RBF) network, but later found suitable for image segmentation [12]



(a) KM (b) MKM

Figure 4: Segmentation CLL using KM and MKM

(a) KM (b) MKM

Figure 5: Segmentation CML using KM and MKM

III. SEGMENTATION PERFORMANCE

To evaluate the performance of segmentation techniques, an assessment of the accuracy of automatic segmentation techniques can be done by comparing the proposed clustering segmented image to manually segmented image as a reference by performing pixels subtraction. The manual segmented image was prepared by manually editing to segment out the object of interest which is nucleus using Adobe Photoshop 7.0. The segmentation performance of KM and MKM algorithm is evaluated by determining the percentage of accuracy based on Equation (2).

$$accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \tag{2}$$

where TP, TN, FP and FN are true positive, true negative, false positive and false negative respectively. The percentage of accuracy as in Equation (2) was calculated based on a comparison between pixels in a segmented image using the segmentation techniques and the pixels of the manually segmented image (expected segmented region of interest (ROI) which is nucleus).

(a) CLL (b) CML

Figure 6: Manual Segmentation CML and CLL using Adobe Photoshop

True positive (*TP*) occurs when the technique correctly indicates the pixels that represent the blast as the blast (object interest). Then, if the technique could correctly indicate the pixels that represent the background as a background, then it will be label as true negative (*TN*). However, if the test could falsely indicate the nucleus when for the true, then it labels as false negative (*FN*). Finally, if the test could falsely indicate the pixels that represent the background, it will be label as false positive (*FP*).

IV. RESULT AND DISCUSSION

Result for this study is shown in Figure 2 to Figure 8 for CLL and CML type with two segmentation techniques respectively. Then, for quantitative assessment, the accuracy of automatic segmentation techniques was performed to evaluate the best segmentation method give the best performance by comparing proposed clustering segmented image to manual segmented image as a reference using pixel subtraction. Percentage accuracy of proposed segmentation techniques was calculated based on Equation (2). Accuracy segmentation performance was calculated based on this remain pixel after performing pixel subtraction using Equation (2). Table 1 shows an example of the mapping pixels subtraction (in percentage) between manual segmentation and automated segmentation clustering.

Table 1
Example of Mapping Pixel Subtraction Between Manual Segmentation and Proposed Clustering Techniques for One Image

Method	Acc. (%)	TP (%)	TN (%)	FP (%)	FN (%)
KM	99.773	5.52	94.25	0.08	0.44
MKM	99.775	5.54	94.23	0.09	0.41

Furthermore, percentage accuracy in different types of chronic leukaemia was shown in Table 2 and Table 3. There were five images of each type of chronic leukaemia respectively. Mean accuracy of segmentation performance in CLL was 99.45% and 99.43% for KM and MKM respectively. Meanwhile, mean accuracy of segmentation performance in CML was 98.15% and 97.80% respectively. Segmentation performance CLL type has higher accuracy compared to segmentation performance in CML type. This was caused by the pixels redundancy of cytoplasm intensity of WBC with the intensity of red blood cell. Frequently, almost all cytoplasm of the blast were eliminated during segmentation process because of the condition pixels redundancy. In fact, the ratio of nucleus to cytoplasm in CML much higher compared to CLL. As a result, more eliminated pixels occurs in CML images compare to CLL images. Consequently, when pixels subtraction procedure was applied, the object of interest pixels was considered as background where this condition is called false negative condition that will lead to less accuracy when segmentation performance was calculated.

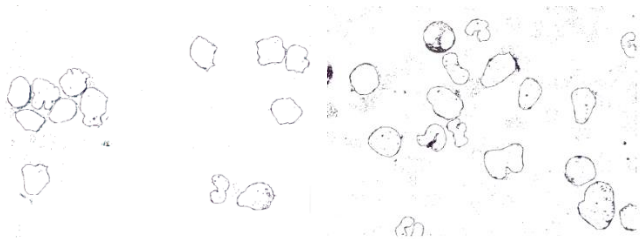
Table 2
Segmentation of Accuracy in 5 of CLL Images

Image	Accuracy Segmentation (%) KM	MKM
1	98.75	98.7
2	99.77	99.7
3	99.22	99.23
4	99.86	99.86
5	99.61	99.54
Mean	99.45	99.43

Table 3
Segmentation of Accuracy in 5 of CML Images

Image	Accuracy Segmentation (%) KM	MKM
1	98.06	97.88
2	98.45	98.38
3	97.37	97.34
4	97.82	97.80
5	98.15	98.56
Mean	98.15	97.80

Figure 7 and Figure 8 were ghost images produced from pixels subtraction between proposed segmentation techniques and manual segmented image of CLL and CML. Ghost image represents the unsuccessful segmented object appeared in original colour. The less pixel mapped on the ghost image picture indicates, the better of the segmentation technique.



(a) CLL (b) CML

Figure 7: Ghost Images for CLL and CML for KM segmentation

(a) CLL (b) CML

Figure 8: Ghost Images for CLL and CML for MKM segmentation

Overall, the mean average of segmentation accuracy and performance comparison between two segmentation techniques shows segmented nucleus have high accuracies to segment out the objects of interest in the chronics images.

V. CONCLUSION

This paper presents two techniques of clustering segmentation and compare performance between image segmentation techniques on chronics leukaemia via k-means and moving k-means.

In this comparison, the accuracy is very similar, and both are equally good. However, KM requires proper initialisation of centres to produce good performance while this is not a requirement for MKM.

In diagnosing leukaemia type, when screening chronic leukaemia by a medical expert, they will also consider the cytoplasm as one of the components for determining the type of leukaemia. However, this study only focuses on the nucleus and the accuracy of nucleus segmentation itself. Furthermore, the purposes of this study are to extract features morphology of nucleus such as nucleus size and shape for automated classification of leukaemia subtype.

ACKNOWLEDGEMENT

The authors would like to express our thanks to University Hospital, University of Science Malaysia Kubang Kerian (HUSM), Kelantan for providing images slide sample and University Malaysia Perlis (UniMAP) for full support and encouragement. This research is funded under Fundamental Research Grant Scheme (Grant No. 9003 00129).

REFERENCES

[1] S. Miwa, "Atlas of Blood Cells", *Okinaka Memorial Institute*, Japan, 1998.

[2] C. B. Emmanuel, "Chronic Myelogenous Leukemia" articles retrieved on 30 Mac 2017 at http://emedicine.medscape.com/article/199425-overviewJ.

[3] T. J. Hamblin, D. G. Oscier. "Chronic lymphocytic leukaemia: the nature of the leukaemic cell", *In Haematological oncology, Blood Review* , 1997, pp.68–73.

[4] K. Yee, P. Anglin, A. Keating "Molecular approaches to the detection and monitoring of chronic myeloid leukemia: theory and practice", *In Blood Reviews*, 1999.

[5] G.A. Colvin, G.J. Elfenbein, "The latest treatment advances for acute myelogenous leukemia", *Med Health R I*, Vol. 86, No. 8, 2003, pp. 243-246.

[6] V. Piuri, F. Scotti: "Morphology Classification of Blood Leucocytes by Microscope Images", *In IEEE International Conference on Computitional Intelligence for Measurement Systems and Applications*, Boston, MA, USA, 14-16 July, 2004.

[7] T. Markiewicz, S. Osowski, B. Marianska, L. Moszezynski: "Automatic Recognition of the Blood Cells of Myelogenous Leukemia Using SVM", *In Proceeding of International Joint Conference on Neural Network*, Montreal, Canada, July 31- August 4, 2005.

[8] R. Adollah, M.Y. Mashor, N.F. Mohd Nasir, H. Rosline, H. Mahsin, H. Adilah, "Blood Cell Image Segmentation: A Review", *In Proceeding on Biomed2008*, Vol 21,2008, pp. 141–144.

[9] N.H. Harun, M.Y. Mashor, N.R. Mokhtar, A. S.Abdul-Nasir, R.Hassan, R.A.A.Raof, M.K.Osman. "Comparison of Acute Leukemia using HSI and RGB color space", *In 10th International Conference on Information Science, Signal Processing and their Applications* (ISSPA 2010), 2010, pp. 740-752

[10] N.H. Harun, M.Y. Mashor, "Automated Blasts Segmentation Techniques Based on Clustering Algorithm for Acute Leukemia Blood Samples", *In Journal of Advanced Computer Science and Technology Research 1* , 2011, pp. 96-109.

[11] M.Y. Mashor, "Hybrid Multilayered Perceptron Networks, *International Journal of System Science*, 31(6), 2000,pp771-785.

[12] N.H. Harun, M.Y. Mashor, "Segmentation Technique for Acute Leukemia Blood Cells Images Using Saturation Component and Moving K-Mean Clustering Procedures", *In Journal of Electrical, Electronic Engineering and Technology, Vol 1*, 2011, pp.23-35 .

[13] N.A. Mat-Isa,"Automated Edge Detection Technique for Pap Smear Images Using Moving KMeans Clustering and Modified Seed Based Region Growing Algorithm", *International Journal of The Computer, the Internet and Management* 45–59, 2005.