

# Automatic Blasts Counting for Acute Leukemia Based on Blood Samples

Nurul Hazwani Abd Halim<sup>1</sup>, Mohd Yusoff Mashor<sup>1</sup>, and Rosline Hassan<sup>2</sup>

<sup>1</sup>Electronic & Biomedical Intelligent Systems (EBItS) Research Group, School of Mechatronics Engineering, University Malaysia Perlis, Perlis, Malaysia

<sup>2</sup>Department of Hematology, School of Medical Science, University Science Malaysia, Kubang Kerian, Kelantan, Malaysia

Email: nurul.hazwani43@yahoo.com, mymashor@yahoo.com.my, roslin@kb.usm.my

**Abstract** – Currently, the conventional method of manual counting under a microscope affords a time-consuming, susceptible to error procedure and put an intolerable amount of stress to medical laboratory technicians. As a solution to this problem, this paper proposed an image processing technique for automatic counting the number of blasts present in the slide of leukemia. In automatic process, the segmentation technique for white blood cell (WBC) is one of most crucial stages. Segmentation based on HSV (Hue, Saturation and Value) color space will be used in order to eliminate the white blood cells (WBC) from the background. A simple morphological operator such as erosion plays importance role especially for the overlapping cells. The experimental results show that the proposed system has provided the highest average accuracy of 97.8% for counting both ALL and AML cases.

**Keywords** – Acute Leukemia, Blasts, Segmentation, HSV Color Space, Erosion, Counting

## 1. Introduction

Leukemia is a type of blood cancer involving white blood cells. It is a bone marrow disorder that arises when abnormal white blood cell begins to continuously replicate itself [1]. These cells do not function normally which is to fight off infections. As they accumulate, they inhibit the production of other normal blood cells in the marrow, leading to anemia, bleeding, and recurrent infections. Over time, the leukemic cells spread through the bloodstream where they continue to divide, sometimes forming tumors and damaging organs such as the kidney and liver. Acute leukemia is classified according to the French-American-British (FAB) classification into two types: Acute Lymphoblastic Leukemia (ALL) and Acute Myelogenous Leukemia (AML). Acute leukemia is a rapidly progressing disease that affects mostly cells that are unformed (not yet fully developed or differentiated) [2]. ALL is most common in children while AML mainly affects adults but can occur in children and adolescents.

In 2009, it is estimated that approximately 31,490 individuals will be diagnosed with leukemia and 44,510 individuals will die of the disease in the United States [3]. In Malaysia, a total of 529 cancer cases of Myeloid Leukemia and 433 cases of Lymphatic Leukemia were reported comprising 4.5% of the total number of cancers in Peninsular Malaysia in 2003 and registered under National Cancer Registry. Out of this value, males predominated at a ratio of 1.7: 1 for Lymphatic Leukemia & 1.1: 1 for Myeloid Leukemia. It was the 4th commonest cancer with 7.1% in males after Lung Cancer, 13.8%, Nasopharynx Cancer, 8.8%, and Colon Cancer, 7.6%. In females, it was the 7th commonest cancer with 4.0% after Breast Cancer, 31.0%, Cervix Uteri, 12.9%, Colon Cancer, 6.0%, Corpus Uteri, 4.3%, Rectum Cancer, 4.1%, and Ovary Cancer, 4.1% [4]. Based on the high number of leukemia cases, the requirement for fast and cost-effective production of blood cell count reports and analyzing are of paramount importance in the healthcare industry.

The early and fast identification of the leukemia type, greatly

aids in providing the appropriate treatment for the particular type. Its detection starts with a complete blood count (CBC). The patient should perform bone marrow biopsy if there are abnormalities in this count. Therefore, to confirm the presence of leukemic cells, a study of morphological bone marrow and peripheral blood slide analysis is done. In order to classify the abnormal cells in their particular types and subtype of leukemia, a hematologist will observed some cells under a light microscopy looking for the abnormalities presented in the nucleus or cytoplasm of the cells. This classification is very important in predicting the clinical behavior of the disease and the prognosis in order to determine which treatment should be given to the patient [5].

Conventionally, manual counting done under the microscope was performed in order to count the white blood cells in leukemia slides. This way is troublesome and time consuming if the counting process is interrupted; it has to be started all over again. Thus, the traditional method of manual counting under the microscope affords susceptible to error procedure and put an intolerable amount of stress on the medical laboratory technicians. Although there are hardware solutions such as the Automated Hematology Counter to perform counting, certain developing countries are not capable to deploy such an expensive machine in every hospital laboratory in the country [6]. Furthermore, hematologists still carry out the manual counting based on slide of blood and bone marrow samples to confirm the case. Therefore, this paper aimed to provide an alternative solution to the problem where white blood cell counting can be carried out automatically at low cost by designing an automated system to count white blood cells in leukemia slides.

## 2. Materials and Methods

This section discusses the related theoretical and methods to automated counting Acute Leukemia based on blood samples. The proposed method consists of the following four steps, as illustrated in Figure 1.

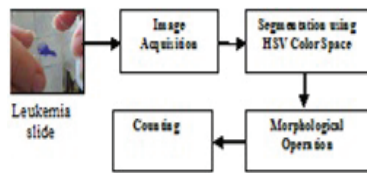


Figure 1. Block diagram of the proposed automated counting Acute Leukemia based on blood samples.

### 2.1. Image Acquisition

The acute leukemia slides were provided by the Hematology Department, Hospital Universiti Sains Malaysia (HUSM), Kelantan. Images of leukemia slides were acquired using the Luminera Infinity 2 digital camera mounted on the Leica light microscope. The slides were analyzed under 40x magnifications. The system captured 24-bit RGB images at a resolution of 806×608 pixels and they were saved in the bitmap (.bmp) file format. 50 samples of leukemia images were taken for the testing of the system. Figure 2 (a) and (b) shows examples image for acute leukemia cells. The digital images captured by the camera were verified by the hematologist to determine the appropriateness and type of the blood cell based on the microscopic details of features and contrast of individual cell.

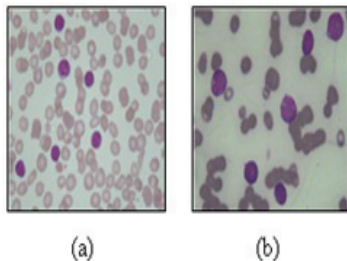


Figure 2. Sample of peripheral blood smear of acute leukemia consists of (a) ALL and (b) AML blast cells.

In order to classify whether the leukemia is ALL or AML, the morphological features will be discovered based on these characteristics [7]:

- The size and shape of white blood cells nucleus. For example, ALL rather small, uniform blast cells with scanty cytoplasm and rounded with usually a single nucleolus. In contrast, AML have blasts with large, often irregular and nuclei with one or more nucleoli.
- The presence of Auer rod and multiple nucleoli inside the nucleus, which are prominent in AML.

### 2.2. Segmentation using HSV Color Space

Image segmentation is a process that partitions an image into its constituents regions or objects [8]. Effective segmentation of complex images such as blood cell is the most challenging tasks in image processing due to complex nature and overlapping these cells. Various segmentation techniques have been proposed to achieve efficient and accurate results such as thresholding techniques, boundary based, region based segmentation [9]-[10]. Moreover, many algorithms for segmentation have been developed for gray scale images [11]. Kumar et al. [12] discussed about an approach for color image segmentation. Kumar stated that color images provide a better description of a scene as compared to gray scale images as it is very rich source of information. In fact, the actual screening process by hematologists is carried out on stained slide where the leukemia is detected based on color and size of blast. Thus, color segmentation becomes a very important issue.

A color can be represented by the red, green and blue (RGB) components. It is a conventional method. Unfortunately, the RGB

are not suitable for describing colors in terms that are practical for human interpretation. In this research, the segmentation technique based on HSV color space images will be used. HSV color model is a kind of method to define color according to the three basic features of the color: hue, saturation and value. The transformation from the source RGB color space to HSV color space is performed based on the following equations [8]:

$$H = \begin{cases} \theta & \text{if } B \leq G \\ 360 - \theta & \text{if } B > G \end{cases} \quad (1)$$

Where

$$\theta = \cos^{-1} \left\{ \frac{1/2 [R - G] + (R - B)}{[R - G]^2 + (R - B)(G - B)]^{1/2}} \right\}$$

$$S = 1 - \frac{3}{(R + G + B)} [\min(R, G, B)] \quad (2)$$

$$V = \frac{\max(R, G, B)}{255} \quad (3)$$

The illustration in Figure 3(a) has shown that the original of leukemia images obtained by the digital image are all in RGB color space. The image will be converted to HSV color space Figure 3(b). Behind this conversion, the individual Hue, Saturation and Value Band can be extracting from original HSV respectively as shown in Figure (c), Figure (d) and Figure (e). Hue is a color attribute that describes a pure color (pure yellow, orange or red), whereas saturation gives a measure of degree to which a pure color is diluted by white color, and value refers to the brightness [8].

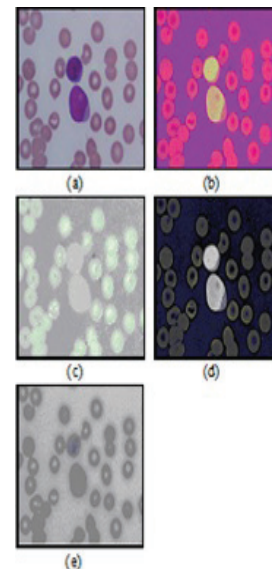


Figure 3. (a) Original RGB image (b) Original HSV image (c) Hue Band (d) Saturation Band (e) Value Band.

According to the fact, the H component in HSI color space contains most of the white blood cells information while the S component contains the structure information of the white blood cells nucleus [13]. Based on this information, color segmentation based on S component image for both blast and nucleus is proposed. Thus, out of these three color spaces, only the S component is used for transforming the RGB image.

The utmost goal of acute leukemia blood cell segmentation is to extract morphological component such as blast from its complicated blood cells background by using thresholding based on S-component. There are 5 steps involved in this processing method.

- Step 1: Extract the S-component information from original HSV image.
- Step 2: Plot the histogram to determine the threshold value.
- Step 3: Segment the WBC (blast) by applying the same threshold value (TH=110) on S component image of ALL and AML.
- Step 4: Convert to binary image.
- Step 5: Implement N x N (N=3) the median filter to the resulted images.

### 2.3. Mathematical Morphological Operation

Mathematical morphology which was introduced by Matheron [14] plays an important role in computer vision since it deals directly with the geometrical properties of objects. This quality makes it an efficient tool for extracting shape information from an images developed from geometry. Based on set theory, mathematical morphology is established by introducing fundamental operators applied to two sets. The basic mathematical morphological operators are dilation and erosion. Only erosion will be implemented in this project.

The erosion of  $f$  by a flat structuring element  $b$  at any location  $(x, y)$  is defined as the minimum value of the image in the region coincident with  $b$  when the origin of  $b$  is at  $(x, y)$ . The equation for the erosion process at  $(x, y)$  of an image  $f$  by a structuring element  $b$  is given by: [15]

$$[f \ominus b](x, y) = \min_{(s, t) \in b} \{f(x + s, y + t)\} \quad (4)$$

Where  $x$  and  $y$  are incremented through all values required so that the origin of  $b$  visits every pixel in  $f$ . The erosion of  $f$  by  $b$  is found by placing the origin of the structuring element at every pixel location in the image. The erosion at any location is determined by selecting the minimum value of  $f$  from all the values of  $f$  contained in the region coincident with  $b$ . Similarly, the dilation of  $f$  by a flat structuring element  $b$  at any location  $(x, y)$  is defined as the maximum value of the image in the window outlined by  $b$  when the origin of  $b$  is at  $(x, y)$ . The equation for dilation is:

$$[f \oplus b](x, y) = \max_{(s, t) \in b} \{f(x - s, y - t)\} \quad (5)$$

Where  $b = (-x, -y)$ . Erosion is a fundamental morphological operator that subtracts pixels from the boundary of any image and causes it to shrink; this operator has the ability to eliminate the irrelevant detail (in terms of size) from a binary image. The effects for dilation are the opposite of those obtained with erosion.

After interpolating and bridging the gaps at the membranes boundary, the inner region is then filled using a combination of the morphological operators [15]. The process of region filling is defined as:

$$X_K = (X_{K-1} \oplus B) \cap A^C, K = 1, 2, 3 \dots \quad (6)$$

Where  $X_K$  is the starting point inside the boundary,  $B$  is structuring element and  $A^C$  is the complement of  $A$ . This process is repeated until  $X_K$  is equal to  $X_{K-1}$ .

Finally, the result is augmented with the original boundary.

### 2.4. Counting

At this stage, counting process will be done by using the function BWLABEL of Matlab 7.0 to count the number of cells automatically:

$$[L, NUM] = BWLABEL(SEGMCELLS, 8) \quad (7)$$

The BWLABEL function labels blood cell in binary image as shown in Figure 3(h) and returns the number of components found from the image in the output value, numObjects. Due to the irregular shape of cell,  $n$  was selected as 8 which mean that any 8-connected object (such as blood cell) in the image SEGMCELLS will be counted. The shape of the structuring element used is based on 8-connectivity between pixels [16]. The objects in the image were labeled by extracting the connected components in the image. 8-connected pixels are neighbors to every pixel that touches one of their edges or corners. These pixels are connected horizontally, vertically, and diagonally. The pixels were labeled 0 for the background. The pixels labeled 1 made up of one object, while the pixels labeled 2 made up of a second object, and so on. The number of white blood cells present in the image can be determined by the number of connected objects found in labeled connected components.

The analysis is done in order to determine the capability of the system for counting the blood samples accurately. Hence, the accuracy between manual counting and automatic counting has been evaluated by using this formula.

$$Accuracy = \frac{\text{Number of WBC in Manual Count}}{\text{Number of WBC in Auto Count}} \times 100 \quad (7)$$

In contrast, based on Demir & Yener [17], the accuracy, sensitivity and specificity can be defined as below:

$$Accuracy = \frac{\text{Number of TP} + \text{Number of TN}}{\text{Total Blood Sample}} \times 100 \quad (8)$$

$$Sensitivity = \frac{\text{Number of TP}}{\text{Number of TP} + \text{Number of FN}} \times 100 \quad (9)$$

$$Specificity = \frac{\text{Number of TN}}{\text{Number of TN} + \text{Number of FP}} \times 100 \quad (10)$$

Where:

True positive (TP): The system yields positive test result for the sample and the sample really has leukemia.

False positive (FP): The system yields positive test result for the sample but the sample does not actually have leukemia.

True negative (TN): The system yields negative test result for the sample but the sample does not actually have leukemia.

False negative (FN): The system yields negative test result for the sample but the sample actually has leukemia.

In briefly, the accuracy of counting system for leukemia is determined based on correct classification of the blood samples. Sensitivity is the probability of a positive screening test among persons that have the disease. Specificity is the probability of a negative screening test among persons that don't have the disease.

## 3. Results and Discussion

This section is dedicated to the discussion of the performance of the proposed automatic counting based on blood samples. Matlab 7.0 software has been used in this study to develop a system for blood cell counting. The development of this system



involved several stages. These include image pre-processing techniques such as image acquisition, image segmentation, morphological operation and last counting. 50 blood images from acute leukemia cases were obtained using the Luminera Infinity 2 digital camera mounted on the Leica light microscope.

In this study, we successfully designed a system that can segment the blast (abnormal white blood cell) and evaluated the accuracy of the proposed methods.

The results of the blood cells image pre-processing applied to the Acute Myelogenous Leukemia (AML) and Acute Lymphoblastic Leukemia (ALL) images for blood cell counting are shown in Figure 3, Figure 4 and Figure 5 respectively. Figure 3(a) and Figure 5(a) shows the captured AML image at a resolution of 806 x 608, whereas Figure 4 (a) shows the original image of ALL.

In order to obtain the fully segmented nucleus, the images had been segmented by using the color image segmentation based on HSV color space.

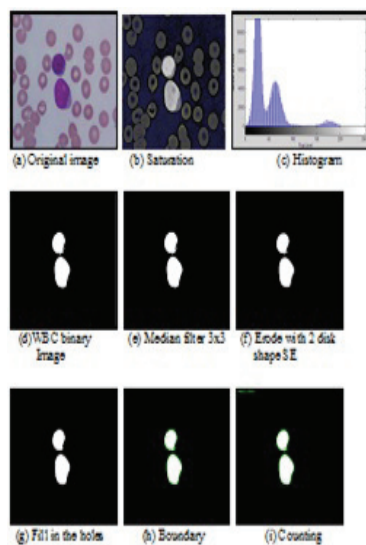


Figure 3. Application of image processing system to Acute Myelogenous Leukemia (AML) image.

Figure 3(b), Figure 4 (b) and Figure 5(b) represent the S component images for both ALL and AML images, respectively. In an image processing context, the histogram of an image normally refers to a histogram of the pixel intensity values.

This histogram is a graph showing the number of pixels in an image at each different intensity value found in that image. Based on the resultant of intensity histogram in Figure 3(c), Figure 4(c) and Figure 5(c), the narrow range of data in original image will be stretched linearly to a wider part of the intensity histogram so that its dynamic range is fulfilled.

The threshold value was selected at 110 (TH=110) and used to segment the WBC from the image background. Even though the threshold value used is fixed to 110, the value is found to be suitable for both ALL and AML images.

After applying the thresholding method, the binary images of WBC for each blood cell images were obtained as shown in Figure 3(d), Figure 4(c) and Figure 5(c).

Object pixels (WBC) will have the value of 1 and the background pixels will have the value of 0. Based from these binary images, there are small spots of object appearing in the WBC background images.

Figure 3(e), Figure 4(e) and Figure 5(e) referred to the result of the image after filtering stage by using median filter. This type of filter produced better result compared to other commonly used type of filters.

The filtering process removed noise that appeared in the

image so that it would not affects the counting process later. After this filtering stage, the blood plasma and dust particles in WBC images were cleaned than before in order to obtain the accurate number of WBC in blood cell counting.

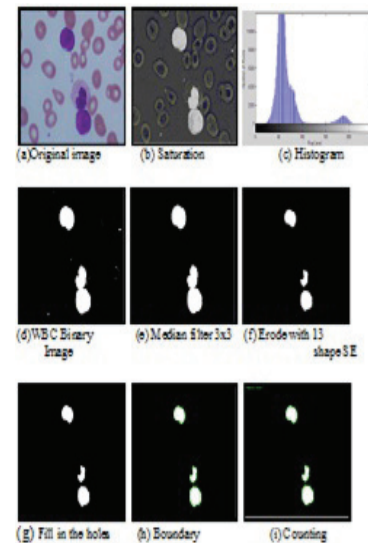


Figure 4. Application of image processing system to Acute Lymphoblastic Leukemia (ALL) image.

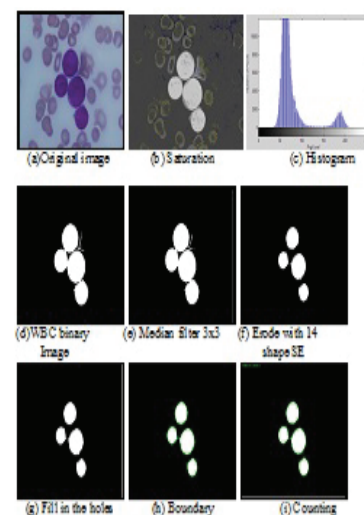


Figure 5. Application of image processing system to Acute Myelogenous Leukemia (AML) image.

In some images, there were overlapping region between one white blood cell and the neighboring cells as shown in Figure 4(e) and Figure 5(e). In this stage, the overlapping regions were separated so that it would not be treated as one cell during counting.

The overlapped white blood cells in the result were separated where the merged region of the cells were split up into different cells. The number of white blood cells present in the image can be determined by the number of connected objects found in labeled connected components.

Not all overlapped regions were able to be separated as can be seen in the result. Additional task was performed by the system to ensure precise counting.

Thus, morphological operation such as erosion will be implemented to the resulted images as shown in Figure 3(f), Figure 4(f) and Figure 5(f).

The erosion operation is useful for removing the small objects. However, it has the disadvantage that the entire remaining objects shrink in size depends on the value of structure element (SE), but in terms of counting process it was acceptable.

Table 1. Evaluation of the proposed method

SAMPLE	M A N U A L COUNT	AUTO COUNT	Accuracy (100%)
Sample 1	2	2	100%
Sample 2	2	3	67%
Sample 3	4	4	100%
Sample 4	1	1	100%
Sample 5	4	4	100%
Sample 6	1	1	100%
Sample 7	6	6	100%
Sample 8	1	1	100%
Sample 9	3	3	100%
Sample 10	3	3	100%
Sample 11	3	3	100%
Sample 12	1	1	100%
Sample 13	1	2	50%
Sample 14	2	2	100%
Sample 15	5	5	100%
Sample 16	3	4	75%
Sample 17	3	3	100%
Sample 18	1	1	100%
Sample 19	1	1	100%
Sample 20	1	1	100%
Sample 21	3	3	100%
Sample 22	1	1	100%
Sample 23	1	1	100%
Sample 24	1	1	100%
Sample 25	1	1	100%
Sample 26	2	2	100%
Sample 27	1	1	100%
Sample 28	1	1	100%
Sample 29	1	1	100%
Sample 30	1	1	100%
Sample 31	1	1	100%
Sample 32	1	1	100%
Sample 33	2	2	100%
Sample 34	2	2	100%
Sample 35	1	1	100%
Sample 36	1	1	100%
Sample 37	3	3	100%
Sample 38	1	1	100%
Sample 39	3	3	100%
Sample 40	3	3	100%
Sample 41	1	1	100%
Sample 42	1	1	100%
Sample 43	1	1	100%
Sample 44	1	1	100%
Sample 45	1	1	100%
Sample 46	1	1	100%
Sample 47	1	1	100%
Sample 48	1	1	100%
Sample 49	1	1	100%
Sample 50	1	1	100%

$$\text{Average Accuracy} = \frac{(\text{Total Accuracy})}{(\text{Total Images Tested})}$$

$$= \frac{4892}{50}$$

$$= 97.8\%$$

Sometimes there are the holes inside the interest region which is referring to the WBC. Thus, the imfill function in Matlab will be used in order to fill image regions and holes in binary images as indicated in Figure 3(g), Figure 4(g) and Figure 5(g) respectively. Hence, the further step of boundary detection will be easy since the holes of blast have been imfill with the previous technique. Since the boundaries of the blast have been detected as presented in Figure 3(g), Figure 4(g) and Figure 5(g) respectively, automatically the counting process will be done.

Fifty images were tested using the automated system to determine the accuracy and reliability of the system. The testing began with manual counting of the abnormal white blood cells which is referring to blast in the original image. This was performed so that the number of blasts gained through manual counting can be compared with the number of blasts produced by the system later as demonstrate in Table 1. If the number of blasts acquired by the system is similar with the manual count, then the system had performed an accurate counting or else the system is not that accurate.

Overall the average accuracy of the proposed system is a very high as 97.8 %. Thus, the system is reliable to count white blood cells that do not overlap with the adjacent cells as indicated in Figure 3(j) and Figure 5(j). The counting will be not that accurate when the white blood cells have very large difference in sizes. In other words, counting is not so accurate in case of AML and ALL type images that contained abnormal white blood (blasts) cells that coexist with the normal white blood cells such as neutrophil as shown in Figure 4(j).

This is the limitation of the system where counting of blasts cells was not totally accurate. The normal white blood cell might be treated as one cell because the system can't distinguish between the normal and abnormal cells in a binary image. Although the accuracy is not totally accurate for this cases, the inaccuracy was acceptable because it does not create much problem as the counting results only differ 1 or 2 cells in most images tested. Thus, the error is quite low and is acceptable. Therefore, the system is successful to achieve its objective which is to count the blasts that present in the leukemia slides.

#### 4. Conclusion

Based on the high average accuracy of the counting, which is 97.8%, our system provides major advantage especially to hematologists as it can provide total count of white blood cells that present in the leukemia slides in an automatic way so that they do not have to count manually as they used to. This will provide a faster solution and effective way for white blood cell counting where it can reduce the burden of the conventional way of counting. Thus, this will speed up the process of early detection of the leukemia disease. Early detection of positive case of leukemia can be arranged to undergo further test to confirm the disease so that early cure for the disease can be administered. This can increase the survival rate of the leukemic patients. For future work, the result of this paper can be used as the basis for extracting the features from the blood slide images.

#### Acknowledgment

The author is gratefully acknowledged and thanks the team members of leukemia research and University Science of

Malaysia (USM). The author also likes to acknowledge Malaysia Government for awarding their National Science Fund (NSF) scholarship and support throughout this project.

## References

- [1] Leukemia Disease, <http://en.wikipedia.org/wiki/Leukemia>, retrieved on 1 June 2011.
- [2] S. Rajendran, "Image Acquisition and Retrieval Systems For Leukaemia Cells" Master's thesis, Department of Computer Science, UM, 2007.
- [3] A. Jemal, R. Siegel, E. Ward, Y. Hao, J. Xu, and M. J. Thun, "Cancer statistics, 2009," CA: A Cancer Journal for Clinicians, 2009.
- [4] National Cancer Registry, Malaysia (2003). Second Report of the National Cancer Registry: Cancer Incidence in Malaysia 2003.
- [5] C.Reta, L.Alamirano, J.A.Gonzales, R. Diaz, and J.S.Guichard, "Segmentation of Bone Marrow for Morphological Classification of Acute Leukemia," in Proceeding of the Twenty-Third International Florida Artificial Intelligence Research Society Conference (FLAIRS 2010).
- [6] G.P.M Priyankara, O.W Seneviratne, R.K.O.H Silva, W.V.D Soysa and C.R. De Silva, "An Extensible Computer Vision Application for Blood Cell Recognition and Analysis", 2006.
- [7] S. Miwa, Atlas of Blood Cells. Tokyo, Japan: Bunkodo Co., Ltd., 1998.
- [8] R.C. Gonzales and R.E.Woods, Digital Image Processing, Prentice-Hall, 2002, pp. 617-620.
- [9] J.Tang "A Color Image Segmentation algorithm Based on Region Growing" in 2010 IEEE 2nd International Conference on Computer Engineering and Technology, Vol. 6, pp. 634 – 637.
- [10] E.Navon, O.Miller, A.Averbuch. "Color Image Segmentation Based on Adaptive Local Thresholds" in International Journal of Image and Vision Computing, 2005, pp 69–85.
- [11] N.Y.Hassan and N.Aakamatsu, "Contrast Enhancement Technique of Dark Blurred Image," in International Journal of Computer Science and Network Security, vol.6 no2A, February 2006.
- [12] R.S.Kumar, A.Verma and J. Singh: " Color Image Segmentation and Multi-Level Thresholding by Maximization of Conditional Entrophy", in International Journal of Signal Processing Vol 3, No. 1, 2006, WASET.ORG.
- [13] J. Wu, P. Zeng, Y. Zhou, and C. Olivier, "A Novel Color Image Segmentation Method and Its Application to White Blood Cell Image Analysis," in 2006 Proceedings of IEEE 8th on ICSP
- [14] J. Serra, Image Analysis and Mathematical Morphology, Academic Press, New York, 1982.
- [15] R.C. Gonzalez, R.E. Woods, Digital Image Processing Using Matlab, Pearson Prentice Hall, Jersey, USA, Inc., 2004.
- [16] A. S. Abdul Nasir, N.Mustafa, N.F. Mohd Nasir, "Application of Thresholding Technique in Determining Ratio of Blood Cells for Leukemia Detection" in 2009 Proceedings of the International Conference on Man-Machine Systems (ICoMMS), 11 – 13 October 2009, Batu Ferringhi, Penang, Malaysia, pp 2A2-1- 2A2-6.
- [17] C.Demir & B.Yener(2005). Automated cancer diagnosis based on histopathological images: a system survey, Tech. Rep., Rensselaer Polytechnic Institute.