Image Segmentation for Acute Leukemia Cells Using Color Thresholding and Median Filter

Leow Bin Toh[1, 2], M.Y. Mashor[1], P. Ehkan[2], H. Rosline[4], A.K. Junoh[3], N.H. Harun[5]
[1]Electronic & Biomedical Intelligent Systems (EBItS) Research Group, School of Mechatronics Engineering,
[2]School of Computer and Communication Engineering,
[3]Institute of Engineering Mathematics,
Universiti Malaysia Perlis, Perlis, Malaysia.
[4]Department of Hematology, School of Medical Sciences, Universiti Sains Malaysia, Kelantan, Malaysia.
[5]Data Science Research Lab, School of Computing, College of Arts and Sciences,
Universiti Utara Malaysia, Kedah, Malaysia.
leowbin@gmail.com

***Abstract*—Acute leukemia is a kind of the malignant disease which may lead to death due to its characteristic of rapid development of immature blood cells. Recently, several image processing techniques have been implemented to assist the task of acute leukemia diagnosis. The segmentation of acute leukemia cells is an important key to determine the accuracy of its classification task. This paper proposed a combined technique of color thresholding based on the RGB color information from acute leukemia slide images and median filter to segment the leukemia cells from the unwanted regions such as background and red blood cells. The presented results proved that the proposed technique was successfully segmented the acute leukemia cells from the Acute Myeloid Leukemia and Acute Lymphocytic Leukemia slide images, with the average accuracy rate of 97.63% and 97.64% respectively. Therefore, the proposed image segmentation technique could benefits the classification process of acute leukemia.**

***Index Terms*—Acute Leukemia; Color Thresholding; Image Segmentation; Median Filter.**

I. INTRODUCTION

Leukemia is a type of blood disease or cancer that begin in the bone marrow and spread into the human blood [1]. According to World cancer Report from World Health Organization (WHO), there are 352,000 people around the world developed a form of leukemia and caused 265,000 deaths in the year 2012 [2]. Acute leukemia is characterized with the rapid progression of immature blood cells, which was critical and may cause death. There are two major types of acute leukemia which known as Acute Lymphocytic Leukemia (ALL) and Acute Myeloid Leukemia (AML) [3].

Conventional diagnosis of acute leukemia required hematologists to identify the maturity and types of blood cells manually under the light microscope. This process could be time intensive and highly dependent on expertise. Recently, several image processing techniques have been developed and applied to the acute leukemia slide images to assist the diagnosis task of acute leukemia. Image segmentation is the key process in image processing which may impact directly on the classification task of acute leukemia. The purpose of implementing image segmentation techniques on acute leukemia slide images is to segment out the blast cells from the unwanted regions in order to ease the features extraction process and classification process later on. Several image segmentation techniques had been explored on acute leukemia slide images, such as fuzzy c-means clustering [4], k-means clustering [5], watershed transform [6], thresholding [7][8], and active contour [9].

This paper proposed an image segmentation technique using color thresholding method based on the RGB color information that was extracted from the slide images of acute leukemia in order to segment the blast cells from the unwanted regions, such as background and red blood cells. After that, a median filter technique was applied to the resultant images to remove the noises that presented in the segmented images. In order to evaluate the performance of the proposed technique in quantitative manner, the accuracy of the segmented image is evaluated by comparing the total number of pixels of blast cells in the segmented image with the total number of pixels in the manually segmented image.

II. METHODOLOGY

A. Implementation Steps

Image segmentation is one of the main sectors in image processing. It has been widely used in computer vision, pattern recognition, medical imaging and so on. The goal of image segmentation is to segment the region of interest in the image from the unwanted regions for further analyzing. Generally, an acute leukemia slide image can be divided into three main regions, there are: background, red blood cells (RBC) and abnormal white blood cells (blasts). A blast cell contains a nucleus and cytoplasm, which was the region of interest for the acute leukemia slide image segmentation because it contain important features that could be used in the classification process, such as the shape of nucleus and cytoplasm, the size of blast cells, and the ratio of size between nucleus and cytoplasm. Whereas, the RBC and background regions are the non-region of interest that will be removed from the acute leukemia slide image. The features that could be obtained from the blast cells can be divided into two different types, which are statistical features and morphological features. The statistical features including the size of blast cells and the shape of nucleus and cytoplasm, whereas the morphological features including the presence of multiple nucleoli and Auer rod inside the nucleus [10]. Generally, the proposed segmentation technique can be divided into five steps of implementation:

Step 1: Acquire the AML and ALL slide images from light microscope.
Step 2: Acquire the value of threshold from the RGB color information of the blast cells.
Step 3: Implement color thresholding technique to the images by using the extracted RGB information.
Step 4: Apply median filter to the resultant images for noise removing.
Step 5: Retrieve the original RGB color space to the segmented blast cells.

B. Color thresholding based on ratio of RGB color space

A study of the color information of the Acute Leukemia slide images had been carried out in order to get the usable information that can be used for the threshold value. This study had carried out on both AML and ALL slide images, which comprise of underexposed images, properly exposed images, and overexposed image. From the study, a color thresholding technique based on the ratio of RGB color space had been carried out to segment the blast cells from the non-region of interests such as RBCs and backgrounds. The proposed technique implements the ratio of RGB color space of the blast cells for thresholding. Based on the RGB information, Equation 1 was formulated to obtain the thresholding algorithm:

$$g(x,y)=\begin{cases} g(x,y) & \text{for } \left[red(x,y)\le blue(x,y)\right]\cap\left[green(x,y)/blue(x,y)\le 0.85\right] \\ g(x,y) & \text{for } \left[red(x,y)> blue(x,y)\right]\cap\left[green(x,y)/blue(x,y)\le 0.75\right] \\ 255 & \text{otherwise} \end{cases} \quad (1)$$

where $g(x,y)$ is the color pixel values for the image itself, $red(x,y)$, $green(x,y)$ and $blue(x,y)$ are the pixel values for the red, green and blue color space respectively. In order to optimize the segmentation performance, the value of green to blue pixel ratio had been selected as 0.85 and 0.75 respectively from the study of RGB color information over the nucleus and cytoplasm from the blast cells. As in (1), the first condition can be used to segment normal blast cells with no color degradation. However, some acute leukemia slide images tend to degrade after some times, the color of blast cells tends to be more pinkish rather than purplish, and the color of RBC will became paler. Thus, the second condition had been introduced to overcome this problem due to the color of blast cells were more pinkish and the value of $red(x,y)$ will be more than the value of $blue(x,y)$ in the degraded slight images. Apart from the first and second conditions, all the remaining pixels will be changed into a value of 255 for all RGB channels, which indicates that they were the non-region of interest.

C. Median Filter

In order to remove the noise that presented in the segmented image after the thresholding process, median filter technique had been proposed in this study due to its effectiveness in reducing the noises [11]. The neighborhood of the median filter can be described as a square mask which had $N \times N$ pixels, and the number of N must be odd numbered either 3, 5, 7 and so on. Figure 1 shows an example of square mask with $N \times N$ pixels, where $N = 3$.

114	116	117
110	130	115
121	111	122

Figure 1: A square mask with $N \times N$ pixels, where $N = 3$.

As shown in Figure 1, the neighborhood values for the central pixel value of 130 are 110, 111, 114, 115, 116, 117, 121 and 122. Thus, the central pixel value of 130 will be replaced with the median value, which is 116. Besides that, when the median filter straddles an edge, it will not create new unrealistic pixel values since the value of median must actually be the value of one of the pixels in the neighborhood. In addition, it also characterized by a smoothing effect which is capable of closing the small holes in the segmented blast cells. In this study, a median filter with $N = 7$ had been found to be suitable for removing the pepper noise, preserves the edge of blast cells, and fills up the small hole in the segmented blast cells.

III. RESULT AND DISCUSSION

The proposed segmentation technique began with the captured of acute leukemia slide images by using a light microscope with 40x magnification. In this studied, three images with different exposure which were underexposed image, properly exposed image and overexposed image had been chosen for the implementation of proposed image segmentation technique. Figures 2 shows the original AML images with different exposure, whereas Figures 3 shows the degraded ALL images with different exposure.

After the slide images with different exposure had been acquired, the proposed color thresholding technique was implemented on the images according to the Equation 1. Figure 4 and Figure 5 show the threshold images for AML and ALL respectively. Next, the median filter with $N = 7$ had been applied to the segmented images in order to remove the unnecessary pepper noises. Figure 6 shows the resultant images for Figure 2, whereas Figure 7 shows the resultant images for Figure 3 respectively.

According to the resultant images in Figure 6 and Figure 7, the blast cells in both AML and ALL images with different exposure had been segmented from the non-region of interest such as background and RBCs. In order to further evaluate the performance of the proposed technique in quantitative manner, the accuracy of the segmented image is evaluated by comparing the number of pixels of blast cells of the segmented image with the number of pixels of the manually segmented image. Table 1 shows the number of pixels of segmented images by proposed technique, number of pixels of manually segmented images, and the result of segmentation accuracy for Figure 6 and Figure 7 respectively. Besides that, a color-based segmentation using *k*-means clustering had been applied to the AML and ALL images for the comparison purpose [12][13]. Table 2 shows the number of pixels of segmented images by *k*-means clustering, number of pixels of manually segmented images, and the result of segmentation accuracy for Figure 8 and Figure 9 respectively.

Table 1
No. of pixels and accuracy of the segmented image by proposed technique

Image no.	No. of pixels	No. of pixels for manually segmented image	Accuracy	Average Accuracy
6(a)	14716	14841	99.16%	
6(b)	24992	25915	96.44%	97.63%
6(c)	17443	17927	97.30%	
7(a)	82661	85966	96.16%	
7(b)	66152	67196	98.45%	97.64%
7(c)	42683	43417	98.31%	

Table 2
No. of pixels and accuracy of the segmented image by *k*-means clustering method [12][13]

Image no.	No. of pixels	No. of pixels for manually segmented image	Accuracy	Average Accuracy
8(a)	14034	14841	94.56%	
8(b)	22102	25915	85.29%	88.40%
8(c)	15301	17927	85.35%	
9(a)	83848	85966	97.54%	
9(b)	70955	67196	94.41%	96.25%
9(c)	44800	43417	96.81%	

According to Table 1, the average accuracy for both AML and ALL images in different exposure are 97.63% and 97.64% respectively. Hence, the small noises or artifacts that cannot be fully eliminated by the proposed technique presented in Figure 7, which probably caused by the staining process does not bring a big impact to the overall segmentation results as the number of pixels from the noise were not significant. In comparison, the average accuracy for both AML and ALL images in different exposure by k-means clustering are 88.40% and 96.25% respectively according to the Table 2. Therefore, the proposed color image segmentation technique had provided higher segmentation accuracy as compare with the *k*-means clustering method [12][13]. Overall, the proposed technique had been proven that capable to segment the blast cells from the AML and ALL images with different exposure.

IV. CONCLUSION

This paper proposed a combined technique of color thresholding based on the ratio of the RGB color space and the median filter. The experiment had carried out on both AML and ALL slide images, which comprise of underexposed images, properly exposed images, and overexposed images respectively. The presented results showed that the proposed technique was capable to segment the blast cells for the AML and ALL images regardless of the exposure of the images. Hence, the proposed technique could be used to segment the underexposed and overexposed acute leukemia images that probably caused by the capability and performance of camera or inappropriate operation by a human operator during the capturing process.

ACKNOWLEDGMENT

This research was supported by Fundamental Research Grant Scheme (FRGS) under The Ministry of Higher Education (MOHE) Malaysia. The authors gratefully acknowledge the Department of Hematology, Universiti Science Malaysia (USM) and the group members of Electronic & Biomedical Intelligent Systems (EBItS) Research Group who provided insight and expertise that greatly assisted the research.

REFERENCES

[1] National Cancer Institute of U.S., "What Is Cancer?," 2015. [Online]. Available: https://www.cancer.gov/about-cancer/understanding/what-is-cancer. [Accessed: 20-Dec-2016].

[2] World Health Organization, "World Cancer Report," 2014.

[3] G. C. C. Lim, "Overview of cancer in Malaysia," *Jpn. J. Clin. Oncol.*, vol. 32, no. suppl 1, pp. S37–S42, 2002.

[4] N. Theera-Umpon, "White Blood Cell Segmentation and Classification in Microscopic Bone Marrow Images," in *Fuzzy Systems and Knowledge Discovery SE - 98*, vol. 3614, L. Wang and Y. Jin, Eds. Springer Berlin Heidelberg, 2005, pp. 787–796.

[5] H. N. Lim, M. Y. Mashor, N. Z. Supardi, and R. Hassan, "Color and morphological based techniques on white blood cells segmentation," in *Biomedical Engineering (ICoBE), 2015 2nd International Conference on*, 2015, pp. 1–5.

[6] H. N. Lim, M. Y. Mashor, and R. Hassan, "White blood cell segmentation for acute leukemia bone marrow images," *Biomedical Engineering (ICoBE), 2012 International Conference on*. pp. 357–361, 2012.

[7] T. Zerfaß, E. Haslmeyer, T. Schlarb, and M. Elter, "Segmentation of leukocyte cells in bone marrow smears," in *Computer-Based Medical Systems (CBMS), 2010 IEEE 23rd International Symposium on*, 2010, pp. 267–272.

[8] R. Adollah, E. U. Francis, M. Y. Mashor, and N. H. Harun, "Bone marrow image segmentation based on multilevel thresholding," in *Biomedical Engineering (ICoBE), 2012 International Conference on*, 2012, pp. 457–461.

[9] N. I. C. Marzuki, N. H. Mahmood, and M. A. A. Razak, "Segmentation of White Blood Cell Nucleus Using Active Contour," *J. Teknol.*, vol. 74, no. 6, 2015.

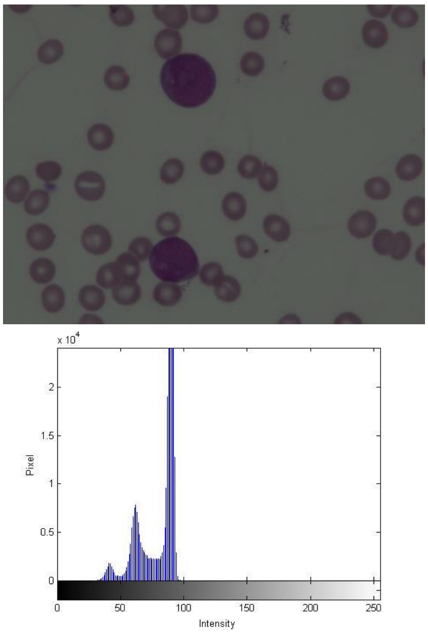
[10] S. Konoplev and C. E. Bueso-Ramos, "Advances in the pathologic diagnosis and biology of acute myeloid leukemia," *Ann. Diagn. Pathol.*, vol. 10, no. 1, pp. 39–65, Feb. 2006

[11] X. Deng, Y. Xiong, and H. Peng, "A new kind of weighted median filtering algorithm used for image Processing," in *2008 International Symposium on Information Science and Engineering*, 2008, vol. 2, pp. 738–743.

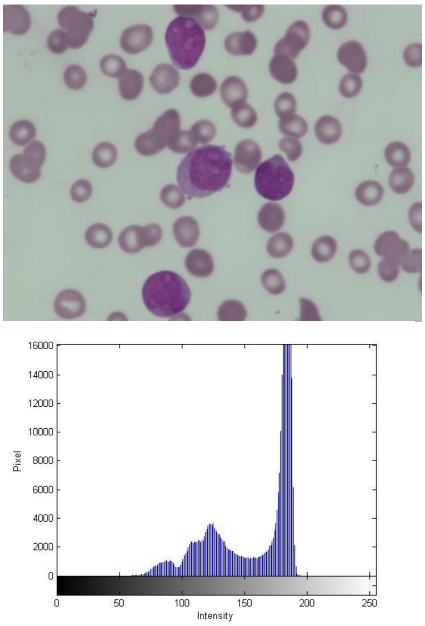
[12] A. Z. Chitade and S. K. Katiyar, "Colour based image segmentation using k-means clustering," *Int. J. Eng. Sci. Technol.*, vol. 2, no. 10, pp. 5319–5325, 2010.

[13] M.-N. Wu, C.-C. Lin, and C.-C. Chang, "Brain tumor detection using color-based k-means clustering segmentation," in *Intelligent Information Hiding and Multimedia Signal Processing, 2007. IIHMSP 2007. Third International Conference on*, 2007, vol. 2, pp. 245–250.

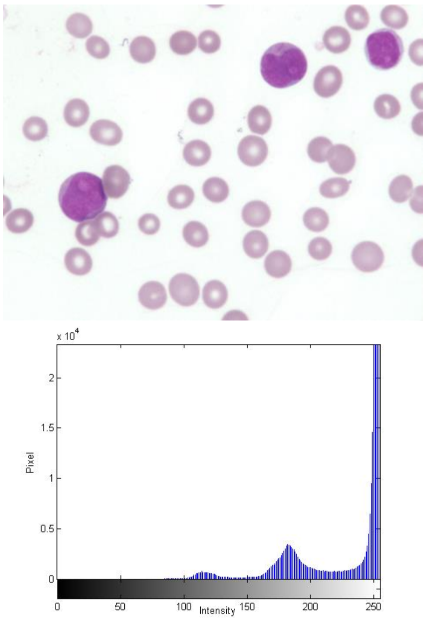
APPENDIX



(a) AML Image in Underexposure

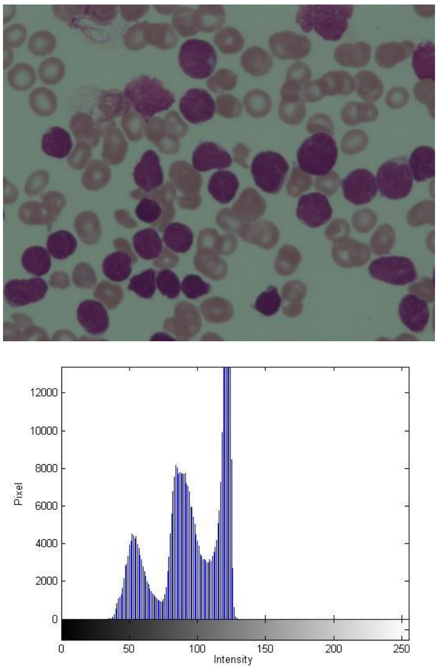


(b)AML Image in Proper Exposure

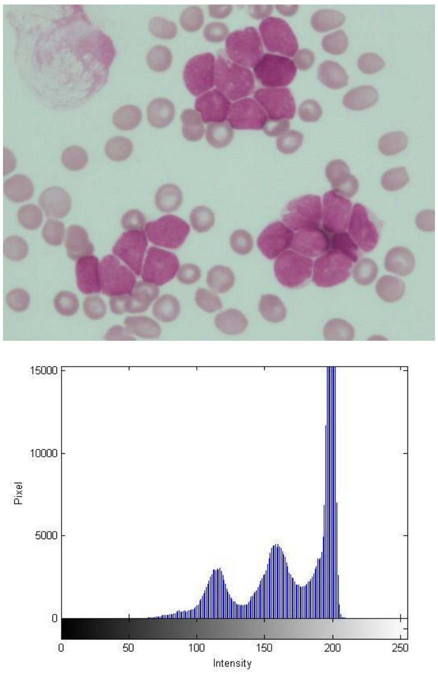


(c) AML Image in Overexposure

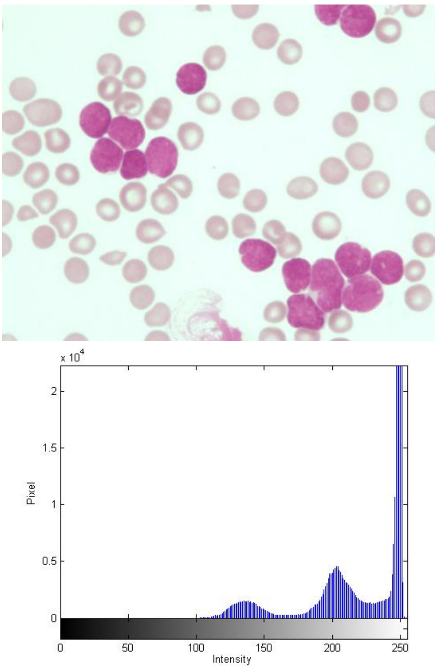
Figure 2: AML images and its intensity histogram in different exposure.



(a) ALL Image in underexposure

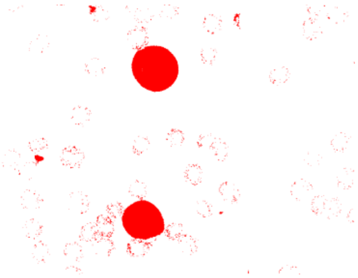


(b) ALL Image in Proper Exposure

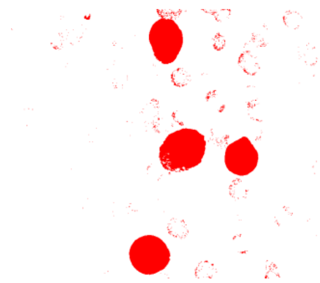


(c) ALL Image in Overexposure

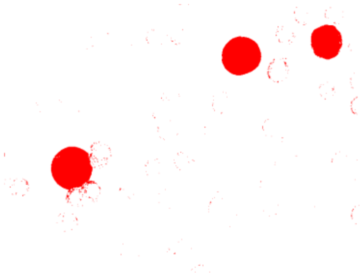
Figure 3: Degraded ALL images and its intensity histogram in different exposure.



(a) Threshold image for AML - Figure 2(a)

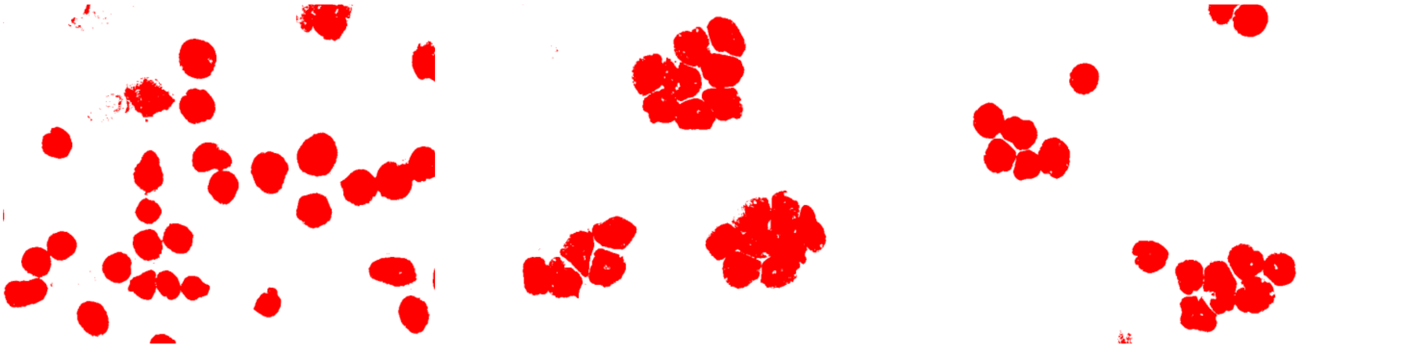


(b) Threshold image for AML - Figure 2(b)



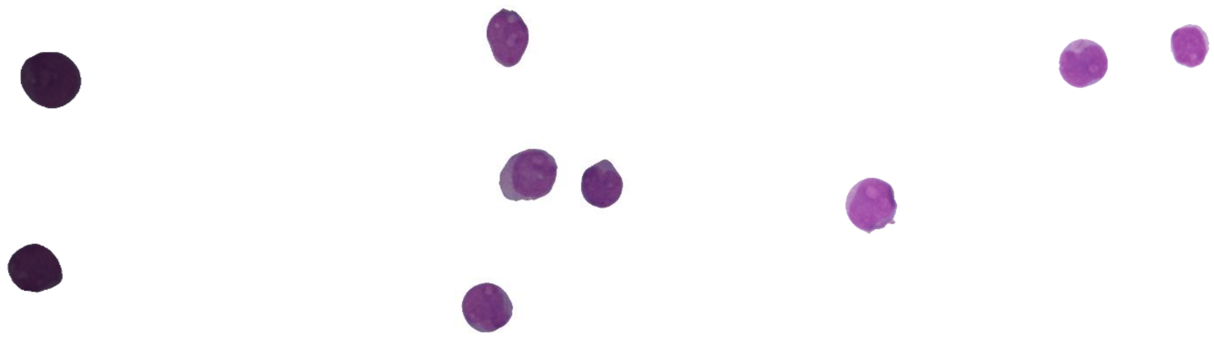
(c) Threshold image for AML - Figure 2(c)

Figure 4: Threshold image for AML in different exposure



(a) Threshold image for ALL - Figure 3(a) (b) Threshold image for ALL - Figure 3(b) (c) Threshold image for ALL - Figure 3(c)

Figure 5: Threshold image for ALL in different exposure.

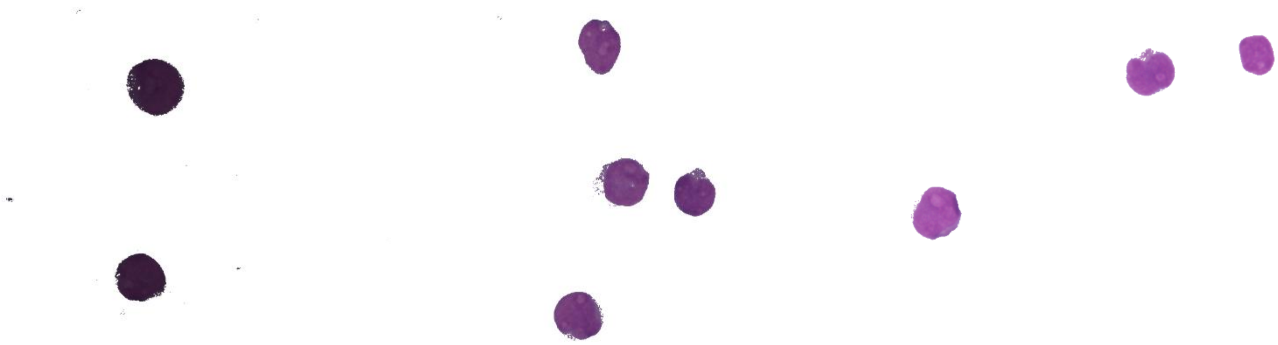


(a) Resultant image for AML - Figure 2(a) (b) Resultant image for AML - Figure 2(b) (c) Resultant image for AML - Figure 2(c)

Figure 6: Segmentation results for AML images with different exposure.

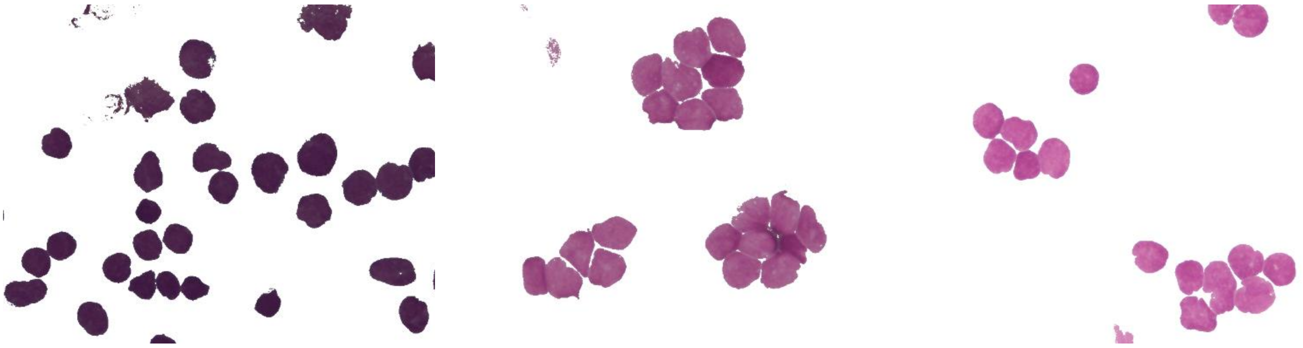
(a) Resultant image for ALL - Figure 3(a) (b) Resultant image for ALL - Figure 3(b) (c) Resultant image for ALL - Figure 3(c)

Figure 7: Segmentation results for degraded ALL images with different exposure.



(a) Resultant image for Figure 2(a) by using *k*-means clustering (b) Resultant image for Figure 2(b) by using *k*-means clustering (c) Resultant image for Figure 2(c) by using *k*-means clustering

Figure 8: Segmentation results for AML images by using *k*-means clustering method.



(a) Resultant image for Figure 3(a) by using *k*-means clustering

(b) Resultant image for Figure 3(b) by using *k*-means clustering

(c) Resultant image for Figure 3(c) by using *k*-means clustering

Figure 9: Segmentation results for degraded ALL images by using *k*-means clustering method.