

An Automated Method for the Nuclei and Cytoplasm of Acute Myeloid Leukemia Detection in Blood Smear Images

Van-Nhan TRAN^{(1)(*)}, Waidah ISMAIL⁽²⁾, Rosline HASSAN⁽³⁾ and Atsuo YOSHITAKA⁽¹⁾

⁽¹⁾ School of Information Science, Japan Advanced Institute of Science and Technology, Nomi, Ishikawa 923-1292, Japan

⁽²⁾ Faculty Science and Technology, Universiti Sains Islam Malaysia, Bandar Baru Nilai, 71800, Nilai, Negeri Sembilan, Malaysia

⁽³⁾ Department of Hematology, Universiti Sains Malaysia, 16150, Kota Bahru, Kelantan, Malaysia

^(*)nhanv@jaist.ac.jp

Abstract—Leukemia is a cancer of white blood cells that affect the blood forming cells in the body. Acute Myeloid Leukemia (AML) is a form of leukemia and are caused by replacement of normal bone marrows with leukemic cells, which cause a drop in red blood cells, platelets, and normal white blood cells. Early classification of the subtype of AML cells is necessary for proper treatment management. We classify the subtype based on the features of AML cells, which include the nuclei and cytoplasm. In this paper, we developed an automate method for the nuclei and cytoplasm detection from the blood cells images that are captured as microscope images. In contrast to other methods that focus on identifying the nuclei, we proposed a method based on the color conversion, intensity threshold and gradient magnitude. Our method detected both the nuclei and the cytoplasm at the same time. We test our method on 301 images, which contain 643 AML cells. The accuracy of both nuclei and cytoplasm detection is over 82.9% (increase 17% when was compared with the existent method).

Keywords—Acute Myeloid Leukemia (AML), cytoplasm, nuclei, gradient magnitude.

I. INTRODUCTION

In recent decades, image processing is applied to many applications of life. In particular, imaging applications are emerging as a new opportunity for innovation at the meeting point between medicine and computer science. By the help of image processing, we can extract many useful information from medical images in order to assist and improve patient diagnosis especially in the cancer area. Currently, with a blood cell image of leukemia patient, the process of detection and classification depends on human look and it takes up to a few days. It is a vital step in providing the correct form of treatment. If we develop an automate method for the detection and classification of leukemia cells from the blood cells images, the diagnostic process will be reduced in terms of its time span from a few days to a matter of a few hours and the cost of all processes.

Leukemia is a cancer of white blood cells, where the disease basically develops in the bone marrow, which is the spongy tissue that fills the inside region of the bones. There are four major different forms or types of leukemia, which develop in cancer patients according to the growth speed and the improper overproduction of leukemia cells: acute lymphocytic leukemia (ALL), acute myeloid leukemia (AML), chronic lymphocytic leukemia (CLL), chronic myeloid leukemia (CML)

[1]. This study will focus on AML type. In AML the bone marrow makes large numbers of abnormal immature white blood cells. The immature cells are called as blast cells. The recognition of the blast cells in the bone marrow of the patients suffering from myeloid leukemia is a very important step. It is followed by categorizing into subtypes which will allow the proper treatment of the patients. In 1971, the diagnosis of leukemia cells was based on the morphology [2]. The whole process is currently manual in nature and thus is time consuming and exhausting.

Nowadays, there are several research groups focusing on the development of image processing application for medical images that collaborates with the clinicians. From the images of blood smear microscope slides, before classify the subtype of AML cells, we had to detect and locate the AML cells. There are two parts of the AML cell, called nuclei and cytoplasm. We can easily segment the nuclei because the contrast between the nuclei and other components of the images are high. However, the color difference between background stain and cytoplasm is minimal, thus it is difficult to extract cytoplasm from the background. Therefore, in this paper, we will analysis the feature of AML cells and propose a method to solve this problem.

II. RELATED WORK

Due to the accuracy of the subsequent feature extraction and classification depends on the correct detection of cells, hence the detection step is very important. There are many methods were applied to segment the cells. In [3], Byoung et al. proposed two different schemes for segmenting the nuclei and cytoplasm of white blood cells (WBC). They used stepwise merging rules for nuclei segmentation and boundary removal rules and gradient vector flow snake for cytoplasm segmentation. Some results were slightly over-segmented because of a weak difference between the cytoplasm and the background or connected red blood cells (RBC). In [4], Putzu et al. present a complete and fully automated method for WBC identification and classification. This approach isolates the whole leucocyte and then separates the nuclei and cytoplasm by using the color conversion and threshold method. A group of clustering methods are also applied to solve the problem [5]-[6]. These methods only detect the nuclei of cell after that using some region growing method to detect the cytoplasm.

However, they cannot detect the complete cytoplasm. All of above method are used for segmentation WBC cells and ALL cells, which belong to Acute Leukemia type. For the AML cells, follow by W. Ismail, [7], the combination of Cellular Automata and Heuristic Search was applied. Although there are a few methods was proposed, however, there are also some problems that need to solve such as the overlapping of blood cells, the noise of images, etc.

III. THE PROPOSED METHOD

Other methods aim first to identify the nuclei of AML cells, which are more prominent than other components. We proposed a method for both the nuclei and cytoplasm detection based on the color conversion and the improvement of threshold technique with the gradient magnitude of images. The flow diagram of the proposed method is show in Fig.1 which consist of 6 small steps.

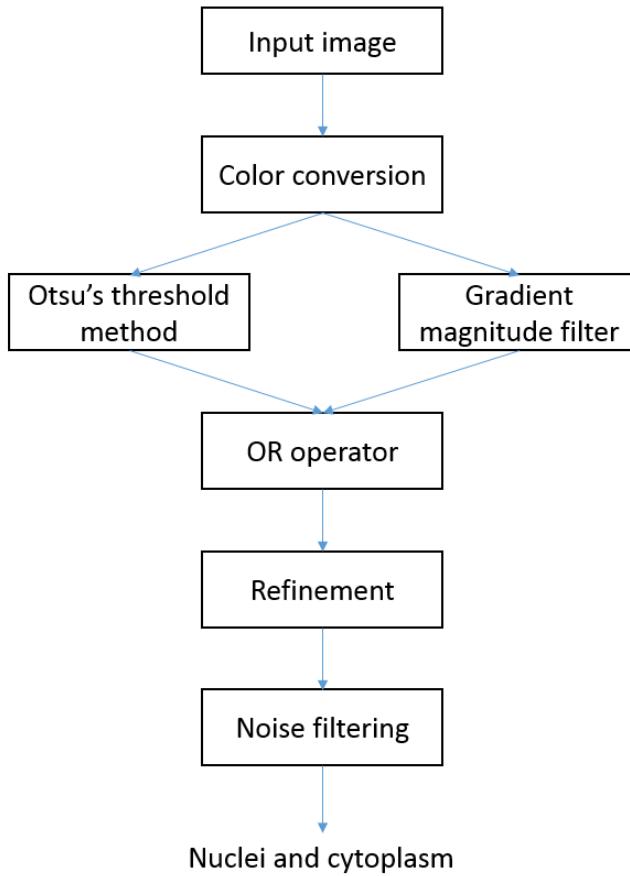


Fig. 1: Flow diagram of proposed system

The AML images, which are captured by microscopes, are usually in RGB color space. It is given as input for segmentation. Firstly, we increase the contrasted between the AML cells and another component such as background, white blood cell. We convert the RGB color space to CYMK space. In fact, AML cells are more contrasted in the Y component of CMYK color model because the yellow color is present in all elements of the image, except AML cells. Redistribution

of image grey levels is necessary to make subsequent segmentation process more easier (Fig. 2 show an example). As

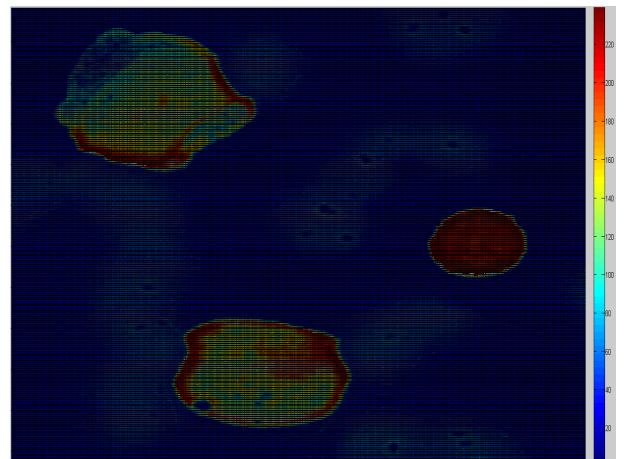
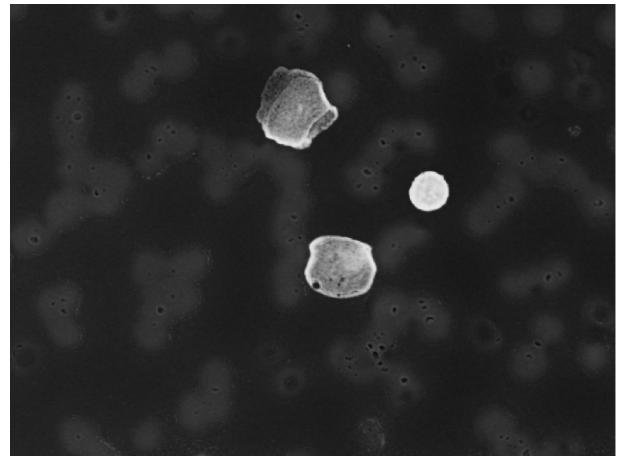
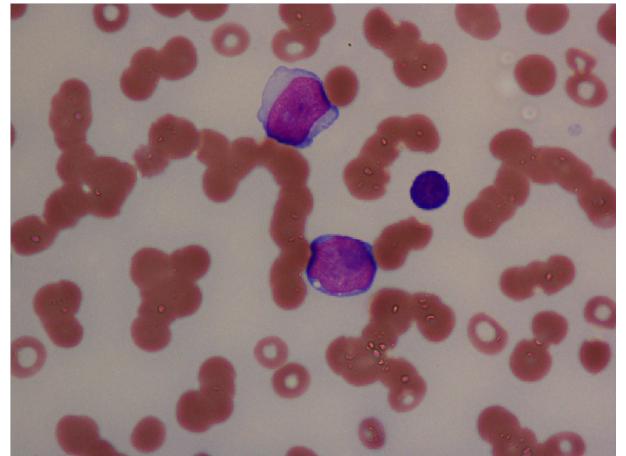


Fig. 2: Top to bottom: original sample image, Y component image, redistribution of image grey level

we can see in the Fig. 2, we can easily recognize the region of AML cells in the Y component images. Segmentation is achieved using an automatically calculated threshold. Here, we use the threshold valued based on the Otsu's method. This is the basic threshold method based on a simple idea: find the

threshold that minimizes the weighted within-class variance. However, in some case, this threshold method cannot segment the cytoplasm because of the similar intensity between the cytoplasm and background. So, we have to find a method to solve the problem in this case. Based on the idea that the gradient magnitude between AML cells region and the background is different, we design a mask filter to calculate the gradient of image. After that, we choose a threshold and filter the region of cytoplasm. The final result after applying it to threshold method is encouraging. We segment exactly the cytoplasm and the nuclei of AML cell. Figure 4 show the change of gradient magnitude of each point in the image. We can imagine the AML cells like a mountain on the ground land.

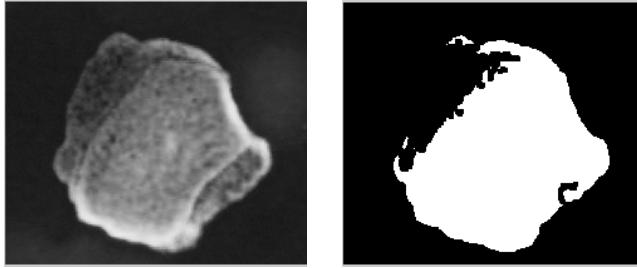


Fig. 3: Left to right: Y component image of an AML cell, result of threshold method

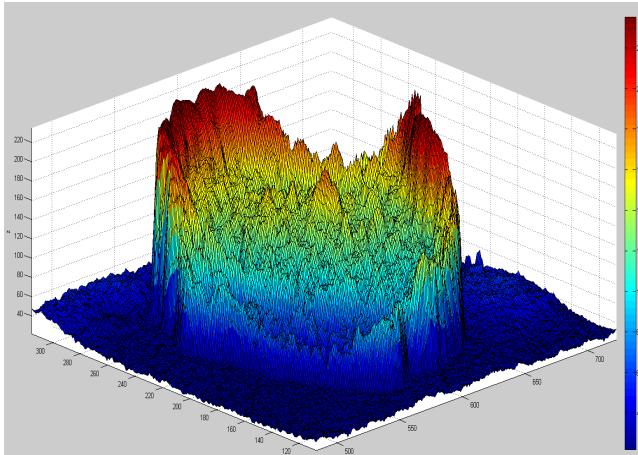


Fig. 4: The visual of the gradient magnitude and the intensity of AML cell

The gradient of a two-variable function at each image point is a 2D vector with the components given by the derivatives in the horizontal and vertical directions. We uses two 3×3 kernels which are convolved with the image to calculate approximations of the derivatives - Gy for horizontal changes, and Gx for vertical. The value of two kernels are:

$$Gx = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}, Gy = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \quad (1)$$

At each point in the image, the resulting gradient can be

combined to give the gradient magnitude, using:

$$G = \sqrt{Gx^2 + Gy^2} \quad (2)$$

We set the value of threshold T to filter by computing the mean of all values in G . The result is showed in the left image of Fig.5. The value of each point in the result image Ir is calculated by:

$$\forall(x, y) \in Ir, \quad Ir(x, y) = \begin{cases} 1 & G(x, y) \geq T \\ 0 & G(x, y) < T \end{cases} \quad (3)$$

In fact, the region of cytoplasm that could not be segmented by threshold method is covered by the result of gradient magnitude filter. After that, we have a binary images, which are result of segmentation step. Because every point inside the AML cells are connected as a block, so we connect them together and get the result like the right image of Fig. 5. There are many hole inside the object and we have to fill these hole.



Fig. 5: Left to right: the result of gradient magnitude filter, the AML cell segmentation

When we apply the gradient magnitude filter to Otsu threshold method, there are some regions in the images satisfy the condition and it becomes the noise of images (Fig 6). Therefore, we set a condition to eliminate them. In the noiseless region, the value of intensity is low and the size is small. Based on this, we have the constraints for eliminating a region:

$$\begin{cases} \text{area} \leq \max(\text{all region})/4 \\ \text{mean(intensity of region)} < \text{mean(intensity of image)} \end{cases} \quad (4)$$

This condition eliminated almost the noise in the binary images and kept all AML cells. In fact, the noises which connected to the AML cells could not be eliminated. This is a problem that we solve in the future. After all above steps, we receive a binary image which contain the region of AML cells.

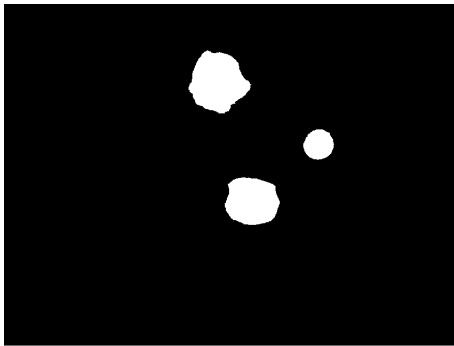
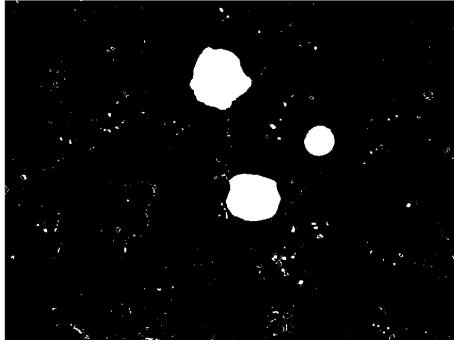


Fig. 6: Top to bottom: The noise in the image, The result after eliminating the noise

We located the AML cells by using the Connected Component Analysis (CCA) [8] technique in binary image. It provides the information about the position of all point inside AML cells and the boundary of AML cells. In Fig. 7, the AML cells were circled by the green line.

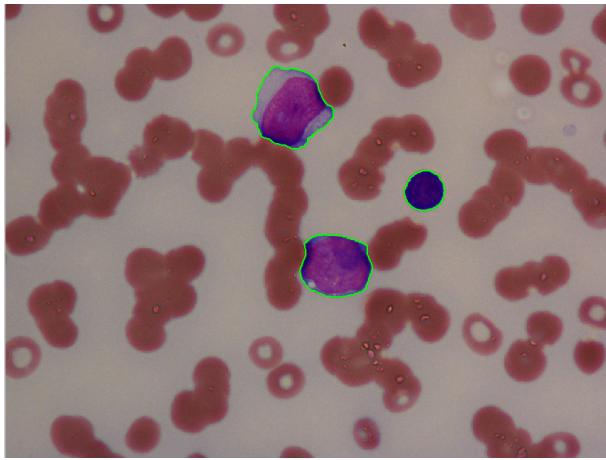


Fig. 7: The final result

IV. EXPERIMENTAL RESULTS

A. Sample Dataset

The dataset consists of 301 real images, 1280 by 960 pixels in size, all from patients suffering from AML. They were provided by Department of Haematology in the Universiti

Sains Malaysia (USM) in Kota Bahru, Kelantan, Malaysia. There are 4 subtypes of AML cells in this dataset, which are M1, M2, M3, M5. There are totally 643 AML cells inside all images. Table I shows the number of available images for each individual subtype. The images were taken with the same lighting and contrast but different microscope zoom magnifications of 40x and 60x. The examples of images from the dataset are shown in the Fig. 8.

TABLE I: DATASET SAMPLE

AML subtype	Images	Cells
M1	44	82
M2	125	249
M3	90	154
M5	42	158
Total	301	643

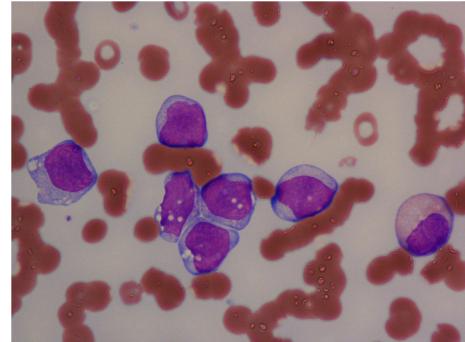
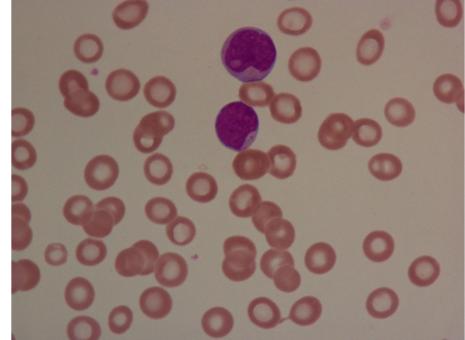


Fig. 8: Example of a bone marrow image of an AML patient

B. Experimental results

We evaluate the performance of the proposed method by the state of AML cells which were segmented. We defined three case of the result: nuclei and cytoplasm, partial cytoplasm and noise. The successful case detected the complete nuclei and cytoplasm (Fig. 10). The failed case detected only the partial cytoplasm (the right image of Fig. 11) or also detect the noise (the left image of Fig. 11). The noiseless cases are caused by the other components, which not belong to the AML cell and were segmented as a part of AML cells. Note that, in noiseless case, the complete cytoplasm are also segmented.

The reasons of noiseless case are the value of threshold for gradient magnitude filter. It is very difficult to determine the optimal value of threshold. In this paper, we choose the simplest way to set the value of threshold.

The performance of proposed method is showed in table II. We properly individuated 533 of 643 AML cells, for an average accuracy of 82.9%.

TABLE II: PROPOSED METHOD PERFORMANCE

Cells	Cytoplasm and Nuclei (%)	Partial Cytoplasm (%)	Noise (%)
M1	82	65 (79.3%)	1 (1.2%)
M2	249	194 (77.9%)	27 (10.8%)
M3	154	150 (97.4%)	0 (0.0%)
M5	158	124 (78.5%)	18 (11.4%)
Total	643	533 (82.9%)	46 (7.2%)
			64 (9.9%)

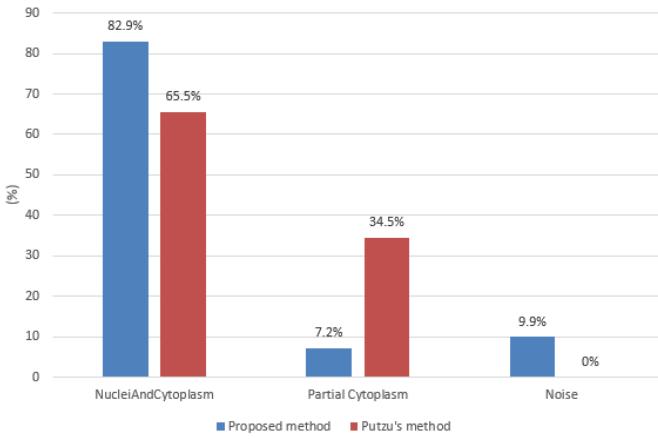


Fig. 9: The comparison of two methods

Currently, there are no research that focus on segmenting both the nuclei and cytoplasm of AML cells. Therefore, we compared our method with one in [4]. This method also used the color conversion and the threshold value based on the triangle method [9]. Figure 9 shows the comparison of two methods.

V. CONCLUSION

In this study, we proposed a method to detect the nuclei and cytoplasm of AML cells. We use the change of gradient magnitude to filter the region of cytoplasm. It is applied to the thresholding. We tested the proposed method with 301 images which total 643 AML cells. The proposed method was demonstrated to improve the detection performance when compared to another method. Experimental results confirmed that our method can efficiently segment the nuclei and cytoplasm of AML cells. We can also use this proposed method for another type of cell such as white blood cell, another type of leukemia cell.

The proposed method is advantageous especially for the images which the difference between cytoplasm and background is low. However, in some case, we also segment some regions

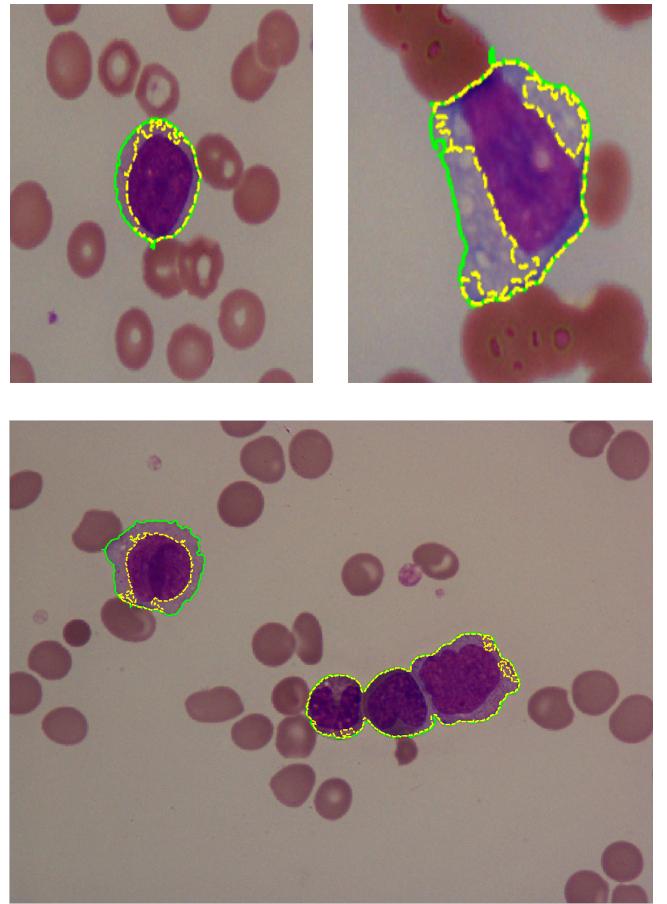


Fig. 10: The final results of two methods. The green lines are our result and the yellow lines are Putzu's result

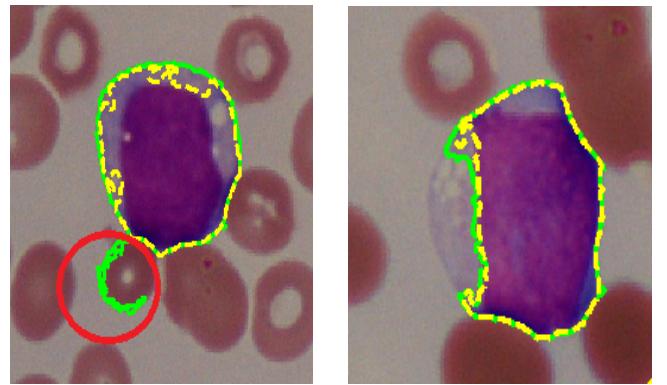


Fig. 11: Left to right: The noiseless case, the partial cytoplasm segmentation. The noiseless region inside the red circle. The green lines are our result and Putzu's result

which are not the cytoplasm. In future work, we propose a method that eliminating the noiseless case. Another side, we will separate the group of AML cells to the single cell. Finally, we extract the features from the single AML cells and classify it into 4 subtypes.

REFERENCES

- [1] C. Haworth, A. Heppleston, M. Jones, et al., *Routine bone marrow examination in the management of acute lymphoblastic leukaemia of childhood*, J Clin Pathol, 1981.
- [2] R., Hassan, *Diagnosis and outcome of patients with Acute Leukemia*, In: Haematology department, University Sains Malaysia, Malaysia, 1996.
- [3] C.K. Byoung, J. Gim, J. Nam, *Automatic white blood cell segmentation using stepwise merging rules and gradient vector flow snake*, Micron 42, 2011.
- [4] L. Putzu, G. Caocci and C. Ruberto, *Leucocyte classification for leukaemia detection using image processing techniques*, Artificial Intelligence in Medicine 62, 2014, pp 179-191.
- [5] O. Sarrafzadeh and A.M. Dehnavi, *Nucleus and cytoplasm segmentation in microscopic images using K-means clustering and region growing*, Adv Biomed Res 2015;4:174.
- [6] F. Jabar, W. Ismail, R.A. Salam, R. Hassan, *Image segmentation using a hybrid clustering technique and mean shift for automated detection acute leukemia blood cells images*, Journal of Theoretical and Applied Information Technology, 2015, vol. 76, pp 88-96.
- [7] W. Ismail, R. Hassan, et al., *Detecting Leukaemia (AML) Blood Cells Using Cellular Automata and Heuristic Search*, Advances in Intelligent Data Analysis IX, 2010, Vol 6065, pp 54-66.
- [8] R.C. Gonzalez, R.E. Woods, *Digital Image Processing*, second edition, ch. 2, sec. 2.5, pp. 66-69.
- [9] G. Zack, W. Rogers, S. Latt, *Automatic measurement of sister chromatid exchange frequency*, J Histochem Cytochem 1977; 25:741-53.