Eleyon SOUTH ASIAN JOURNAL OF ENGINEERING AND TECHNOLOGY

Full Length Article

CLASSIFICATION OF CELL TYPES IN ACUTE MYELOID LEUKEMIA (AML) OF M4,M5 AND M7 SUBTYPES WITH SUPPORT VECTOR MACHINE CLASSIFIER

S.S ARAAL Dr.T.SRIHARI

[a] *PG Student, Department of Electrical and Electronics Engineering , KSR Institute For Engineering and Technology,Tiruchangode-637215 , Tamilnadu, India*

[a] *Professor, Department of Electrical and Electronics Engineering , KSR Institute For Engineering and Technology,Tiruchangode-637215 , Tamilnadu, India*

ABSTRACT Acute Myeloid Leukemia (AML) is one of cancer type that attack white blood cells in myeloid descendants. On the clinical examination of leukemia, the number of each blast cell in the laboratory is calculated. However, in some subtype of AML like M4, M5 and M7 are affected by the same type of precursor cells. The precursor cell of them are myeloblast, monoblast and megakaryoblast, which needs more detailed analysis to distinguish. Classification is performed on cell types of precursors cells derived from bone marrow preparations. The stages that have been completed are preprocessing, segmentation, extraction and feature selection, and classification. Features used as input of classification stage are area, nucleus ratio, circularity, perimeter, mean, and standard deviation. The support vector machine classification results in the best performance test data are achieved by Linear kernel. The performance was obtained by combining six features for eight cell types from the maturation of the three precursor cells. In this work the method isolates and determines the type of blood cancer in the smear which could be either normal or abnormal and determines the presence of blood white cell. Taking into consideration that abnormal white blood cells indicate to the associated blood leukemia.This research aims to classify cells based on cell type on AML subtypes M4, M5 and M7. The techniques used are the K -Means and some additional methods of segmentation and also multiclass Support Vector Machine using one-vs-rest comparison model for cell of the development of precursor cells. The results of this study are expected to reduce errors and inconsistencies in classification of cell types in AML subtypes M4, M5 and M7.The results are verifying using MATLAB/SIMULINK environment

Keywords: ***acute myeloid leukemia, bone marrow, segmentation, classification, support vector machine***

1 Introduction

Leukemia is one of the blood cancers that attack white blood cells that form in the bone marrow. In patients with leukemia, the bone marrow produces white blood cells (blast) excessively. Excessive white blood cells will cause the accumulation of young white blood cells in the bone marrow. It can affect the inhibition and decrease of healthy blood cells.The French-American-British (FAB) hematologic classification system divides leukemia into four types based on its forming cells; they are Acute Myeloid Leukemia (AML), Acute Lymphocytic Leukemia (ALL), Chronic Myeloid Leukemia (CML) and Chronic Lymphocytic Leukemia (CLL) [2]. Acute leukemia is characterized by the rapid development of blast cells in the blood. If not treated immediately, it can lead to death in a matter of weeks or even days. Furthermore, acute leukemia is divided into eight subtypes; they are M0, M1, M2, M3, M4, M5, M6, M7 [3] [4]. In some subtype of AML like M4, M5 dan M7 are affected by the same type of precursor cells. The precursor cell of them are myeloblast, monoblast and megakaryoblast [5], which needs more detailed analysis to distinguish. Generally, leukemia is identified by counting the number of white blood cells and red blood cells through a microscope based on cell morphology by hematologists [6]. However, the procedure of calculating blood cells under a microscope is still relatively time-consuming, highly dependent on the operator's ability and fatigue factor [7]. This research aims to classify cells based on cell type on AML subtypes M4, M5 and M7. The techniques used are the K -Means and some additional methods of segmentation and also multiclass Support Vector Machine using one-vs-rest comparison model for cell of the development of precursor cells. The results of this study are expected to reduce errors and inconsistencies in classification of cell types in AML subtypes M4, M5 and M7.

2 Related Research

Some studies have done a lot of enhancement technique of image quality in recent years. One of them have been done by using contrast stretching technique [8]. Contrast stretching is used to improve the process of diagnosis of acute leukemia images, thus providing additional information on the cytoplasm and nucleus cell [8]. Research on the improvement of image quality among others has been done by using median filter [9]. Median filter could provide a smoother appearance by retaining the edge detail of the object [9].At the stage of identification of medical object, one of the important stage is segmentation. Segmentation aims to separate the part of the cell body consisting of nucleus and cytoplasm. The thresholding method also can be used for segmentation [10] [11]. On the other hand, another technique use color as a medium for object separation by using K-Means[7][13].Besides these two methods, watershed distance transform also can be used to perform separation of stacked cells [10]. Morphology reconstruction technique is also applied with good result for segmentation

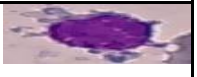
Feature extraction has an important role in the classification technique as the stage performed after the segmentation process. Features are used as input values that represent the characteristics of the observed object. The shape feature like circularity is very influential in determining the characteristics of the object. It is proved by a maximum accuracy value of 95.70% [10] [12]. Other features such as area, ratio of nucleus, and perimeter are also the characteristics that can be used to recognize objects with an average accuracy of 89.68% [12] [14]. In addition to these four features, color features like mean and standard deviation of RGB colors can also be applied [12].The classification stage of objects based on features has been done by using supervised classification techniques with support vector machine. Classification with SVM techniques can be applied linearly and non-linearly [15] and are very powerful for data splitting based on hyperplane classifier [16]. The SVM technique has been used to separate objects by applying linear separation with an accuracy of 93.5% [7] and used to separate six cell types with an accuracy of 97% [13].

3. Results and discussion

3.1 Cell Type Identification

The obtained images of AML M4, M5 and M7 consist of eight cell types of different amounts. These cell types are obtained by expert identification support.The eight cells are derived from 3 precursor cells; they are myeloblast, monoblasts and megakaryoblasts. These three precursors undergo a process of maturation with their respective cell.Myeloblast develop into 3 types, namely myeloblast, promyelosit and granulocytes. Monoblast develop into monoblast, promonosit, monocytes. Whereas in megakaryoblast the derived cell is megakaryoblast. In addition to the seven cells, there is a support / lymphocyte cell derived from lymphoblast stem cells encountered in all identified AML subtypes. The sample of each cell is shown in Table I.

TABLE I. SAMPLE OF EACH CELL

No.	*Cell Type*	*Image of Object*
1.	Myeloblast	
2.	Promyelocyte	
3.	Granulocyte	
4.	Monoblast	
5.	Promonocyte	
6.	Monocyte	
7.	Megakaryoblast	
8	Support (Lymphocyte)	

3.2 Cell Segmentation Results

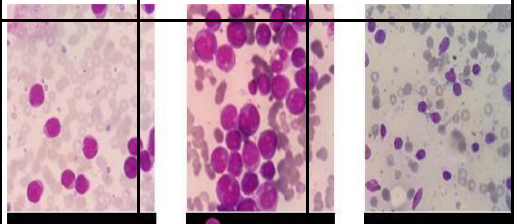
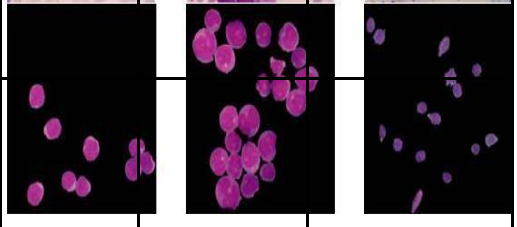
The segmentation results show that from the total of 105 images consisting of 35 images of each subtype, 1500 cells are correctly segmented, and 210 cells are incorrectly segmented. The detailed comparison of cell numbers of each subtype is shown in Table II. Moreover, sample of cell segmentation results in each subtype are shown in Table III.

TABLE II.

DETAILS OF SEGMENTATION RESULT OF EACH SUBTYPE

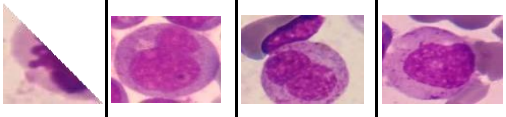
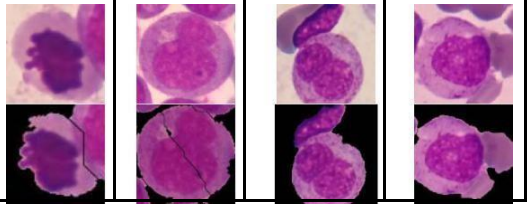
Type	**Detail**			
	Correct Segmentation	***Incorrect Segmentation***	***Total Cell***	***Percentage***
AML M4 Preparation	300	41	341	87.98%
AML M5 Preparation	448	134	582	76.97%
AML M7 Preparation	752	35	787	95.55%
Total	**1500**	**210**	**1710**	**87.72%**

TABLE II SAMPLE OF WHOLE CELL SEGMENTATION RESULT

Image Type	Subtype		
	AML M4	***AML M5***	***AML M7***
Original Image			
Segmentation Results			

There are some cells that are categorized as incorrect in the segmentation process. Cells that fail or incorrect to be segmented are not included as input data in the classification stage. The example of incorrect segmented cells along with the reason is shown in Table IV.

TABLE IV. INCORRECT CELL TERMS AND REASONS

Image Type	Reason			
	Non-WBC Cell	***Cell body is cut off***	***Cell area failed to separate***	***Cell area exceeds the Body***
Original Image				
Segmentation Result				

3.3 Feature Extraction Results

All the features of each cell data are stored and act as input data at the classification stage. The average values of cells have been calculated and shown in Table V.

TABLE V. AVERAGE FEATURE EXTRACTION RESULT OF EACH CELL TYPE

Feature	*T1*	*T2*	*T3*	*T4*	*T5*	*T6*	*T7*	*T8*
Area	9097.808	13167.818	10404.577	11941.691	12997.013	16690.750	2713.324	2895.563
Nucleus Ratio	0.731	0.562	0.521	0.676	0.685	0.533	0.760	0.768
Perimeter	347.466	465.752	403.381	406.249	423.651	600.205	202.372	192.585
Circular-Ity	0.936	0.876	0.787	0.899	0.897	0.589	0.840	0.920
Mean	126.048	137.439	133.256	120.318	117.407	144.822	103.126	101.208
St. Deviation	16.294	19.780	24.854	17.847	17.384	22.765	21.247	21.698

The columns T1 to T8 in Table V stand for the eight successive cell types, they are myeloblast, promyelocyte, granulocyte, monoblast, promonocyte, monocyte, megakaryoblast and support cell.

From the average data value of features can be seen characteristics of cell types. For example, monocyte cells have the largest area, whereas myeloblasts have the highest circularity value compared to other cell types.

3.4 Classification Result

The classification stage begins with the training of data with a model as the output and be used in the testing phase. Based on the test results of 20% of data from each cell type, experiments were conducted on linear and non - linear SVM kernels. Several experiments also performed by combination of features obtained from Minimal Overlap Probability calculations.

The parameters used in the training and data testing phase are the best C (cost) and Û (gamma) values. From the experimental results on feature combinations, RBF kernel is getting better results than the linear kernel, but in the final results with the highest accuracy average values obtained by the linear kernel that includes all of the features. Details of each accuracy on each feature combination with the linear kernel and RBF is shown in Fig.3

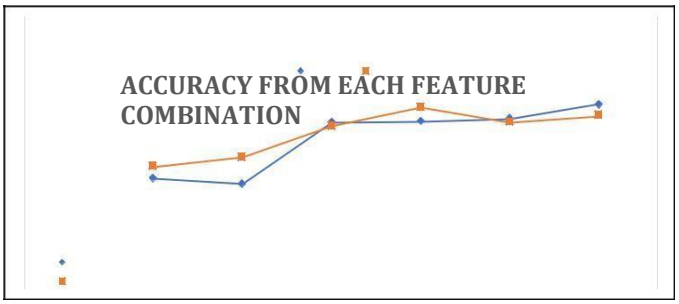


Fig 3.Accuracy on each feature combination ith linear kernel

Furthermore, the accuracy of each type of cell is selected based on kernel type with the highest accuracy. The kernel that obtains the highest accuracy is linear kernel that includes all combinations of the six extracted features. The detailed of the highest accuracy of each cell obtained by the linear kernel is shown in Fig 4.

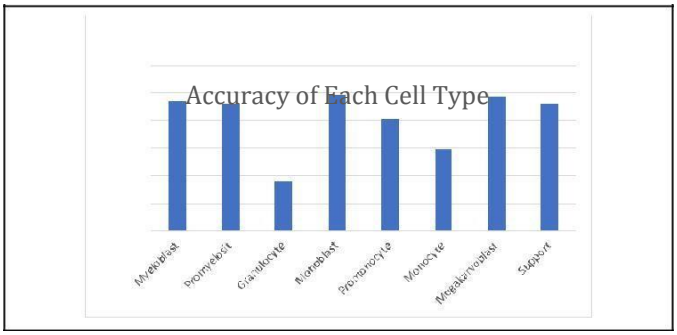
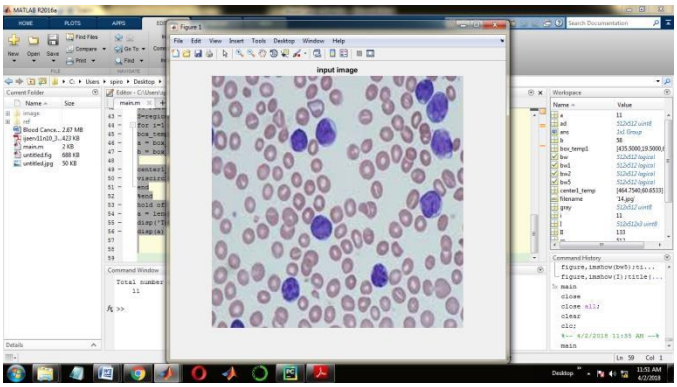


Fig. 4. Results of classification accuracy ofeach cell type

OUTPUT SNAPSHOTS

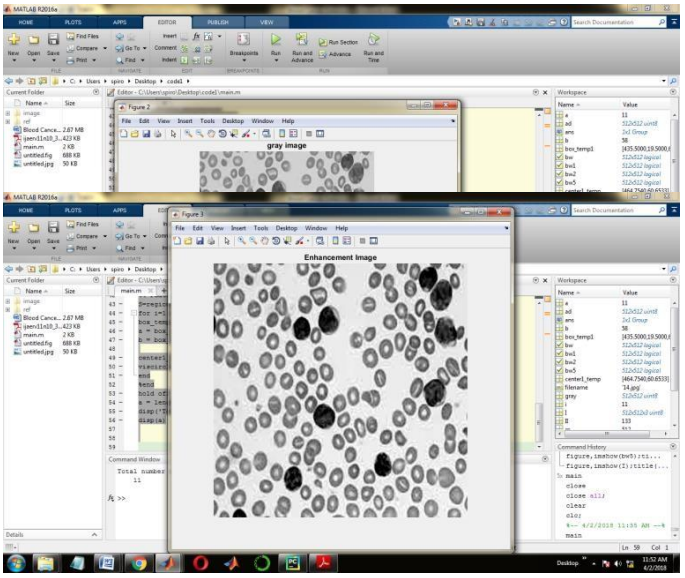
Input image



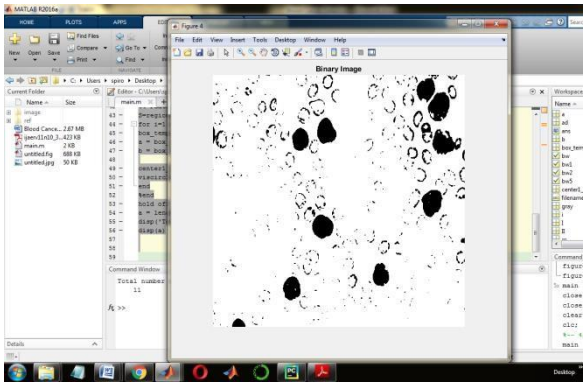
2) Gray Converted Image

3) Enhancement Image

OUTPUT SNAPSHOTS



4) Binary Converted Image



4 Conclusions

Segmentation technique using K-Means with WDT and morphological reconstruction operation can be well applied to segment the bone marrow blood cell object. It proved by the result of an average segmentation of 87.72% from 1710 total real cells. The presentation of correct segmented object data from each AML M4, M5 and M7 preparations were 87.98%, 76.97% and 95.55%, respectively.

The supervised classification with multiclass support vector machine with one-vs -rest model has also been applied with good results. The best kernel that can separate the eight data types based on the experiment is linear kernel. The best result of it includes all feature data with the accuracy value of each cell types of 98.67%, 98.01%, 84.05%, 99.67%, 95.35%, 89.70%, 99.34% and 98.01% respectively.

References

[1] N. R. Mokhtar, N. H. Harun, M. Y. Mashor, H. Roseline, N. Mustafa, and R. Adollah, "Image Enhancement Techniques Using Local, Global, Bright, Dark, and Partial Contrast Stretching For Acute Leukemia Images," *Engineering*, vol. I, pp. 3–8, 2009.

[2] Halim, N.H.A. et al., "Nucleus segmentation technique for acute leukemia", *Proceedings - 2011 IEEE 7th International Colloquium on Signal Processing and Its Applications*, CSPA 2011, pp.192–197, 2011

[3] L. Putzu and C. Di Ruberto, "White Blood Cells Identification and Counting from Microscopic Blood Image," *International Journal of Medical, Health, Biomedical, Bioengineering and Pharmaceutical Engineering,* vol. 7, no. 1, pp. 20–27, 2013.

[4] L. H. Nee, M. Y. Mashor, and R. Hassan, "White Blood Cell Segmentation for Acute Leukemia Bone Marrow Images," *Journal of Medical Imaging and Health Informatics*, vol. 2, no. 3, pp. 278–284, 2012.

[5] A. A. Nasir, M. Y. Mashor, and R. Hassan, "Classification of acute leukaemia cells using multilayer perceptron and simplified fuzzy ARTMAP neural networks," *International Arab Journal of Information Technology,* vol. 10, no. 4, pp. 356–364, 2013.

[6] M. M. Amin, S. Kermani, A. Talebi, and M. G. Oghli, "Recognition of Acute Lymphoblastic Leukemia Cells in Microscopic Images Using K Means Clustering and Support Vector Machine Classifier," *Journal of Medical Signal and Sensors* , vol. 5, no. 1, pp. 49–58, 2015.

[7] E. Suryani, U. Salamah, and A. A. Wijaya, "Identifikasi Penyakit Acute Myeloid Leukemia (AML) Menggunakan 'Rule Based System' Berdasarkan Morfologi Sel Darah Putih Studi Kasus: AML2 dan AML4," *Seminar Nasional Teknologi Informasi, dan Komunikasi Terapan 2014,* pp. 193–199, 2014.

[8] D. Goutam and S. Sailaja, "Classification of Acute Myelogenous Leukemia in Blood Microscopic Images using Supervised Classifier," *International Conference on Engineering and Technology (ICETECH),* pp. 3–7, 2015.

[9] K. Laosai, Jakkrich and Chamnongthai, "Acute Leukemia Classification by Using SVM and K-Means Clustering," *Electr. Eng. Congr. (iEECON), 2014 Int*., pp. 1–4, 2014

[10] Leukimia and Lymphoma Society, "The AML Guide Information for Patients and Caregivers", 1311 Mamaroneck Avenue, Suite 310, Whute Plains, New York, 2012.

[11] American Cancer Society, "Colorectal Cancer Facts & Figures 2014-2016". Atlanta, Ga: American Cancer Society, 2014.

[12]G. Abdul-hamid, "Classification of Acute Leukemia," *Univ. Aden/Hematology Unit Yemen,* 1999.

[13]A. Bell and S. Sallah, "The Morphology of Human Blood Cells – Seventh Edition", Abott, Memphis, 2005.

[14]Dacie and Lewis, "Practical Haematology, Eleventh Edition", Elsevier Churcill Livingstone, London, ISBN-13: 9780702034084, 2011.

[15]K. Anwar, A. Harjoko, and S. Suharto, "Feature Selection Based on Minimum Overlap Probability (MOP) in Identifying Beef and Pork," *International Electrical Engineering Congress (iEECON) 2014,* vol. 7, no. 3, pp. 316–322, 2016.

[16]N. Pradana T.P, E. Suryani, Wiharto, "Pemanfaatan Seed Region Growing Segmentation dan Momentum Backpropagation Neural Network Untuk Klasifikasi Jenis Sel Darah Putih", *Seminar Nasional Teknoin 2013 FTI UII*, Yogyakarta, 2013.