ISSN:0975-9646

An Approach to classify Acute Myelogenous Leukemia Using LBP Based Features

Neena George[1], Lisha Mathai[2]

[1]*Dept of CSE, SNGCE, Kadayiruppu*
[2]*Dept of CSE, SNGCE, Kadayiruppu*

***Abstract*- Acute Myelogenous leukemia (AML) is a specific kind of cancer where the blood cells, bone marrow get affected. AML is commonly seen in adults with an average age of 65 years. The current method for the detection of Acute Myelogenous leukemia needs the manual examination of blood smear, so it is not efficient. Its accuracy depends on the operator's ability and it is time consuming. This paper presents a simple approach that classifies AML into its subtypes using DRLBP and DRLTP features. The proposed approach mainly comprises of preprocessing which involves the conversion of RGB images into CIELAB color space. This step is followed by segmentation which extracts the blue nuclei from an AML image. Discriminative Robust Local Binary Pattern (DRLBP) and Discriminative Robust Local Ternary Pattern (DRLTP), which is based on LBP, are extracted for efficient classification.**

***Keywords*- Acute Myelogenous Leukemia (AML), Discriminative Robust Local Binary Pattern (DRLBP), Discriminative Robust Local Ternary Pattern (DRLTP)**

I. INTRODUCTION

Acute myelogenous leukemia can be defined as a rapidly developing cancer consisting of blood and bone marrow, in which hematopoietic harbingers are seized in the early stages of development [1]. It can also be said as the outcome of a maturational seize of the cells of the bone marrow in the premature evolution stage. In case of AML the bone marrow present in the human body generates numerous cancerous cells, which are also known as leukemic blasts. The development of these leukemic blasts comes to a halt and is not properly ameliorated, thus failing in fighting infections. These leukemic blasts results in seizing the production of normal WBC's, RBC's and platelets. This type of leukemia is also known as "myelogenous" because myeloid cells (a group of white blood cells) are affected due to it. This type of leukemia is caused due to antecedent hematologic malfunctions, environmental vulnerabilities, hereditary syndromes and prolonged usage of drugs. It is also known as acute myelocytic leukemia, acute granulocytic leukemia, or sometimes simply AML. It has been discovered that approximately 15000 people of the total countries population gets affected by this type of leukemia. The quotidian symptoms of acute myelogenous leukemia include anemia, decreased count of WBC's, decreased count of platelets, increased number of leukemic cells [2]. AML is initially treated by the implementation of bone marrow or blood transplant.

Most signs and symptoms of AML are caused by the replacement of normal blood cells with leukemic cells. A lack of normal white blood cell production makes people more susceptible to infections; while the leukemic cells themselves are derived from white blood cell precursors, they have no infection-fighting capacity [3]. Figure 1 represents AML and non-AML images.

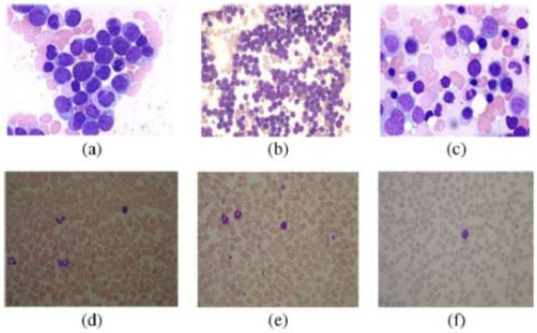

Fig. 1. (a)-(c) AML cells. (d)-(f) Healthy cells from non AML patients

II. RELATED WORKS

Many attempts have been made for classifying AML cells from the blood smear [1-10]

Gupta et al. [11] developed a system using relevant vector for identifying lymphoblast. The system identifies three types of lymphoblasts. This system works well for the ALL of children but not for ALL of adults. The problem with this system is that the Otsu's algorithm has been used for segmenting the lymphoblast which is not strong method.

Piuri [12] performed white blood cell segmentation. The preprocessing step here used is to enhance the input image and it selects the white cells by separating from other blood's component. Low pass and band pass filters are used for removing the noise contents presents in the image. He used edge detection technique for the each leucocyte. He used a neural network for the classification. He trained neural network by morphological features to recognize lymphoblast. The disadvantage of the system is that the processing time required is higher and the computational complexity is also higher.

Halim [13] proposed an automated system which counts number of blasts in the microscopic blood image. He applied some threshold operation on S component of HSV color space to detect white blood cells from the image. The results achieved were quite amazing but the problem in the system is that no method is mentioned for selecting the optimum threshold for the better segmentation. There is no features extracted and no classifier has been used.

III. PROCESS OVERVIEW

The overall system is depicted in Figure 2. The overview of the system shows the sequence of steps that are to be followed for the efficient classification of acute myelogenous leukemia. The first step is the preprocessing step which includes the conversion of RGB images into the L*a*b color space image. This step is followed by the segmentation technique, which uses k-means clustering for extracting blue nuclei. DRLBP and DRLTP features, which are based on LBP, are extracted for classification.

IV. PREPROCESSING

A Image Acquisition

The collection of AML images are provided by American Society of Hematology *(ASH)*, which is a web based image library. It provides high quality images with different resolutions. The resolution used for the images are 200 X 200 pixels.

B. CIELAB Color Correlation

The digital microscope will generate RGB images which is difficult to segment. According to color and intensity the blood cells and image background varies greatly. The reasons for this can be camera settings, aging stain and varying illumination. So for making the segmentation robust, we are converting the RGB image into CIELAB color space [14], [15] or more correctly L*a*b color space. A Lab color space is a color-opponent space with the dimension L for lightness and a and b for the color-opponent dimensions, based on nonlinearly compressed (e.g. CIE XYZ) coordinates. However, Lab is now more often used as an informal abbreviation for the L-a-b representation of the CIE 1976 color space . The lightness, L*, represents the darkest black at L* = 0, and the brightest white at L* = 100. The color channels, a* and b*, will represent true neutral gray values at a* = 0 and b* = 0. The red/green opponent colors are represented along the a* axis, with green at negative a* values and red at positive a* values. The yellow/blue opponent colors are represented along the b* axis, with blue at negative b* values and yellow at positive b* values.

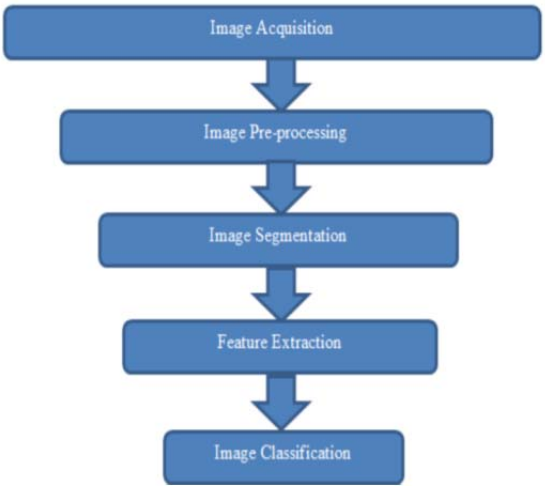

Fig. 2. Overall System

V. SEGMENTATION

Image segmentation is the process of partitioning a digital image into multiple segments (sets of pixels, also known as super-pixels) [16]. The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze. Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. More precisely, image segmentation is the process of assigning a label to every pixel in an image such that pixels with the same label share certain characteristics.

The result of image segmentation is a set of segments that collectively cover the entire image, or a set of contours extracted from the image. Each of the pixels in a region is similar with respect to some characteristic or computed property, such as color, intensity, or texture. Adjacent regions are significantly different with respect to the same characteristic(s). When applied to a stack of images, typical in medical imaging, the resulting contours after image segmentation can be used to create 3D reconstructions with the help of interpolation algorithms like marching cubes.

The K-means clustering algorithm is an iterative technique that is used to partition an image into *K* clusters. In this paper clusters corresponds to nucleus (high saturation), background (high luminance and low saturation), and other cells (e.g. erythrocytes and leukocyte cytoplasm). Here, every pixel is assigned to one of these classes using the properties of the cluster center [17]. After performing k-means clustering, only the edges of the nuclei are obtained as opposed to the whole image of the nuclei. So inorder to preserve the edges of the nuclei, some morphological operations like dilation, erosion, imclose, imopen are performed.

VI. FEATURE EXTRACTION

Feature extraction is a process of redefining a large set of data into a smaller set of features. The selection of features is an important step, as the Classifier performance is influenced by the feature selection. The features are extracted for the blue nuclei. The features extracted were DRLBP and DRLTP, which is based on the LBP feature.

A. DRLBP

The LBP code at location (x,y) is computed as follows:

$$LBP_{x,y} = \sum_{b=0}^{B-1} s(p_b - p_c)2^b,$$

$$s(z) = \begin{cases} 1, & z \geq 0 \\ 0, & z < 0 \end{cases}$$

Where pc is the pixel value at (x,y), pb is the pixel value estimated using bilinear interpolation from neighbouring pixels in the b-th location on the circle of radius R around pc and B is the total number of neighbouring pixels.

The histogram of LBP code is computed as follows.

$$h_{lbp}(i) = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} \omega_{x,y}\delta(LBP_{x,y}, i),$$

$$\delta(m,n) = \begin{cases} 1, & m = n \\ 0, & \text{otherwise} \end{cases}$$

The RLBP histogram is computed as follows:

$$h_{rlbp}(i) = h_{lbp}(i) + h_{lbp}(2^B - 1 - i), \quad 0 \le i < 2^{B-1}$$

Histogram of discriminative local binary pattern is considered as the absolute difference between the bins of a LBP code and its complement. It is computed as follows:

$$h_{dlbp}(i) = |h_{lbp}(i) - h_{lbp}(2^B - 1 - i)|, \quad 0 \le i < 2^{B-1}$$

The 2 histogram features, RLBP and DLBP, are concatenated to form Discriminative Robust LBP (DRLBP) as follows:

$$h_{drlbp}(j) = \begin{cases} h_{rlbp}(j), & 0 \le j < 2^{B-1} \\ h_{dlbp}(j - 2^{B-1}), & 2^{B-1} \le j < 2^B \end{cases}$$

For B=8, the number of bins is 256 (128 + 128). Using uniform codes, it is reduced to 60 (30 + 30)

B. DRLTP

The LTP code at location (x,y) is computed as follows:

$$LTP_{x,y} = \sum_{b=0}^{B-1} s'(p_b - p_c)3^b,$$

$$s'(z) = \begin{cases} 1, & z \ge T \\ 0, & -T < z < T \\ -1, & z \le -T \end{cases}$$

Where T is a user-defined threshold. For B=8, the histogram has 6561 bins which is very high-dimensional. Hence it is proposed to split the LTP code into its "upper" and "lower" LBP codes [18]. The "upper" code, ULBP, is computed as follows:

$$ULBP = \sum_{b=0}^{B-1} f(p_b - p_c)2^b,$$

$$f(z) = \begin{cases} 1, & z \ge T \\ 0, & \text{otherwise} \end{cases}$$

The "lower" code, LLBP, is computed as follows:

$$LLBP = \sum_{b=0}^{B-1} f'(p_b - p_c)2^b,$$

$$f'(z) = \begin{cases} 1, & z \le -T \\ 0, & \text{otherwise} \end{cases}$$

The histogram of URLBP is computed as follows:

$$h_{urlbp}(s) = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} \omega_{x,y}\delta(\max(ULBP, LLBP), s),$$

The histogram of LRLBP is computed as follows:

$$h_{lrlbp}(t) = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} \omega_{x,y}\delta(\min(ULBP, LLBP), t)$$

The histogram of UDLBP is computed as follows:

$$h_{udlbp}(s) = \left| \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} \omega_{x,y}\delta'(\lambda(ULBP, LLBP), s) \right|$$

$$\lambda(p,q) = \begin{cases} p, & p > q \\ -q, & p < q \end{cases}$$

$$\delta'(m,n) = \begin{cases} 1, & m = n, m > 0 \\ -1, & |m| = n, m < 0 \\ 0, & \text{otherwise} \end{cases}$$

λ(•) determines whether the ULBP and LLBP codes are being swapped. If a swap occurs, the negative maximum code is assigned to the result. δ'(•) checks the value output from λ with s. If the value is positive and matches s, then the bin value is incremented. Otherwise, it is decremented. The histogram of LDLBP is computed as follows:

$$h_{ldlbp}(t) = \left| \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} \omega_{x,y}\delta''(\lambda'(ULBP, LLBP), t) \right|$$

$$\lambda'(p,q) = \begin{cases} q, & p \ge q \\ -p, & p < q \end{cases}$$

$$\delta''(m,n) = \begin{cases} 1, & m = n, m \ge 0 \\ -1, & |m| = n, m < 0 \\ 0, & \text{otherwise} \end{cases}$$

λ'(•) determines whether the ULBP and LLBP codes are being swapped. If a swap occurs, the negative minimum code is assigned to the result. δ''(•) checks the value output from λ with t. If the value is zero or positive and matches t, then the bin value is incremented. Otherwise, it is decremented. The URLBP, LRLBP, UDLBP and LDLBP histograms are then concatenated to form DRLTP.

VII. CLASSIFICATION

Classification is the task of assigning to the unknown test vector, a label from one of the known classes. Since the patterns are very close in the feature space, support vector machines (SVM) are employed for classification. SVM is a powerful tool for data classification based on hyperplane classifier [19]. This classification is achieved by a separating surface (linear or non linear) in the input space of the dataset.

Doctors divide acute leukemia's into myeloid and lymphoblastic leukemia's. But they also divide them into even smaller groups or subtypes. This is called classification.

There are 8 types of AML in the FAB system – M0 to M7. All these types have names as well as numbers. M0, M1 and M2 are all myeloblastic leukemia and together make up just under half of all AML cases (45%). M3 is called acute promyelocytic leukemia (APL) and makes up 1 in 10 cases (10%) of adult AML. Together M4 and M4eos are called acute myelomonocytic leukemia and they make up a quarter of AML cases (25%). M5 is called acute monocytic leukemia and makes up 1 in 10 cases (10%) of AML. M6 is called acute erythroid leukemia – it is very rare. M7 is called acute megakaryoblastic leukemia [20].

VIII. EXPEERIMENTAL RESULTS

A dataset, which consists of 150 images provided by the American Society of Hematology (ASH) are used for doing this experiment. The microscopic images have been sent to the AML detection system and the system gives the subsequent images as the result.

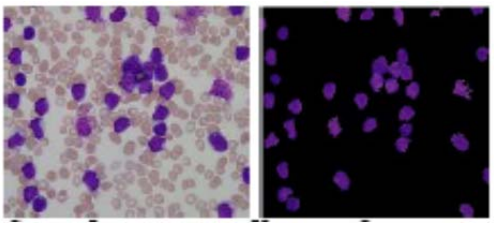

Fig. 3. An AML image and image after k-means clustering

Figure 3 shows the original image and the image after *k*-means clustering. Then after the classification by using SVM, we get the outcome as shown in Figure 4. Similar images as the outputted image are retrieved from the image database.

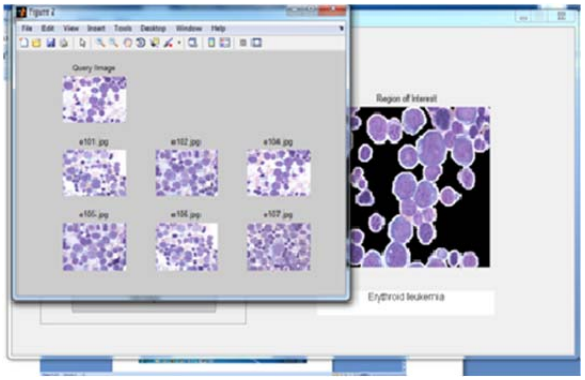

Fig. 4. Erythroid Leukemia Detected

IX. CONCLUSION AND FUTURE WORK

This experiment was performed to classify Acute Myelogenous Leukemia into its subtypes by the implementation of medical imaging approaches. Image segmentation is mainly done by a technique based on k-means clustering technique. The experiments have been carried out in MATLAB. The propounded approach was tested with blood microscopic images acquired by the help of microscopy, thus precisely detecting the affected cells. DRLBP and DRLTP, which is based on LBP, are extracted to obtain all the information required to perform efficient classification. The classification is done with the help of Linear Support Vector Machine (SVM), which helps us to classify AML into its subtypes. Image Retrieval is also included which will retrieve similar images of the output image.

Further research will focus on improving the classification of AML subtypes by considering more images for the dataset.

ACKNOWLEDGMENT

We are extremely grateful to the American Society of Hematology for providing a high quality online image database.

REFERENCES

[1] Sos Agaian, Monica Madhukar, and Anthony T. Chronopoulos, "Automated Screening System for Acute Myelogenous Leukemia Detection in Blood Microscopic Images".

[2] F. Scotti, "Automatic morphological analysis for acute leukemia identification in peripheral blood microscope images," in Proc. CIMSA, 2005, pp. 96–101.

[3] V. Piuri and F. Scotti, "Morphological classification of blood leucocytes by microscope images," in Proc. CIMSA, 2004, pp. 103–108.

[4] M. Subrajeet, D. Patra, and S. Satpathy, "Automated leukemia detection in blood microscopic images using statistical texture analysis," in Proc. Int. Conf. Commun., Comput. Security, 2011, pp. 184–187.

[5] H. Ramoser, V. Laurain, H. Bischof, and R. Ecker, "Leukocyte segmentation and classification in blood-smear images," in Proc. IEEE EMBS, 2006, pp. 3371–3374.

[6] C. Reta, L. Altamirano, J. A. Gonzalez, R. Diaz, and J. S. Guichard, "Segmentation of bone marrow cell images for morphological classification of acute leukemia," in Proc. 23rd FLAIRS, 2010, pp. 86–91.

[7] G. Ongun, U. Halici, K. Leblebicioglu, V. Atalay, M. Beksac, and S. Beksac, "Feature extraction and classification of blood cells for an automated differential blood count system," in Proc. IJCNN, 2001, vol. 4, pp. 2461–2466.

[8] S. Mohapatra and D. Patra, "Automated leukemia detection using hausdorff dimension in blood microscopic images," in Proc. Int. Conf. Emerg. Trends Robot Commun. Technol., 2010, pp. 64–68.

[9] S. Mohapatra, S. Samanta, D. Patra, and S. Satpathi, "Fuzzy based blood image segmentation for automated leukemia detection," in Proc. ICDeCom, 2011, pp. 1–5.

[10] S. Mohapatra, D. Patra, and S. Satpathi, "Image analysis of blood microscopic images for acute leukemia detection," in Proc. IECR, 2010, pp. 215–219.

[11] L. Gupta, S. Jayavanth, and A. Ramaiah. Identification of different types of lymphoblasts in acute lymphoblastic leukemia using relevance vector machines. In Proceedings of the 31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society, USA, 2009.

[12] Piuri V, Scotti F. Morphological classification of blood leucocytes by microscope images. In: Proceedings of the IEEE international conference on computational intelligence for measurement systems and applications, 14–16 July. Boston,MA, USA: IEEE Publisher; 2004. p. 103–8.

[13] Halim NHA, Mashor MY, Hassan R. Automatic blasts counting for acute leukemia based on blood samples. Int J Res Rev Comput Sci 2011;2(August(4)).

[14] E. Wharton, K. Panetta, and S. Agaian, "Logarithmic edge detection with applications," J. Comput., vol. 3, no. 9, pp. 11–19, Sep. 2008.

[15] J. Hu, J. Deng, and M. Sui, "Color space conversion model from CMYK to LAB based on prism," in Proc. IEEE GRC, 2009, pp. 235–238.

[16] Acute Myeloid Leukemia, National Cancer Institute, Bethesda, MD, USA, 2006.

[17] S. Nercessian, K. Panetta, and S. Agaian, "A non-reference measure for objective edge map evaluation," in Proc. IEEE Int. Conf. Syst., Man Cybern., 2009, pp. 4563–4568.

[18] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," IEEE Trans. Image Process., vol. 19, no. 6, pp. 1635–1650, Jun. 2010.

[19] C. C. Chang and C. J. Lin, "LIBSVM: A library for support vector machines,"ACM Trans. Intell. Syst. Technol., vol. 2, no. 3, p. 27, Apr. 2011.

[20] Bennett JM, Catovsky D, Daniel MT, Flandrin G, Galton DA, Gralnick HR, Sultan C (1976). "Proposals for the classification of the acute leukaemias. French-American-British (FAB) co-operative group". Br J Haematol 33 (4): 451–8.