Segmentation and Analysis of Cancer Cells in Blood Samples

Arjun Nelikanti

Assistant Professor, NMREC, Department of CSE
Hyderabad, India
anelikanti@gmail.com

Abstract

Blood cancer is an umbrella term for cancers that affect the blood, bone marrow and lymphatic system. Acute Lymphoblastic Leukemia (ALL) is one of the kinds of blood cancer which can be affected at any age in the humans. The analysis of peripheral blood samples is an important test in the procedures for the diagnosis of leukemia. In this paper the blood sample images are used and implementing a clustering algorithm for detection of the cancer cells. This paper also implements morphological operations and feature extraction techniques using MATLAB for the analysis of cancer cells in the images.

Keywords*:* Blood Cancer, Acute lymphoblastic leukemia(ALL), Clustering, MATLAB.

1. Introduction

Anyone can get a blood cancer at any age. Many people, from babies to grandparents, are diagnosed with blood cancer every year. Most of these cancers start in your bone marrow where blood is produced. Stem cells in your bone marrow mature and develop into three types of blood cells: red blood cells, white blood cells, or platelets. In most blood cancers, the normal blood cell development process is interrupted by uncontrolled growth of an abnormal type of blood cell. These abnormal blood cells, or cancerous cells, prevent your blood from performing many of its functions, like fighting off infections or preventing serious bleeding.

Acute lymphoblastic leukemia (ALL) or acute lymphoid leukemia is an acute form of leukemia, or cancer of the white blood cells, characterized by the overproduction of cancerous, immature white blood cells—known as lymphoblasts. This is one of the kinds of blood cancer found more commonly. They can then spread to other parts of the body, including the lymph nodes, liver, spleen, central nervous system (brain and spinal cord), and testicles (in males). Other types of cancer also can start in these organs and then spread to the bone marrow, but these cancers are not leukemia.

In medicine, the digital image processing techniques are used to process bio-medical images. For any digital image processing applications there are certain steps to be followed as, segmentation, morphological operations and feature extraction etc,. Clustering techniques are also used on these bio-medical images so as to differentiate cancer tissues from the other textures. Clustering can be considered the most important unsupervised learning problem; so, as every other problem of this kind, it deals with finding a structure in a collection of unlabeled data. The adaptive K-means clustering algorithm [1] is used in this paper. And in morphological operations and feature extraction the image processing techniques are used.

2. Methodology

The overview of the proposed methodology is shown in the fig.(1). The proposed system start with reading a true or color image.

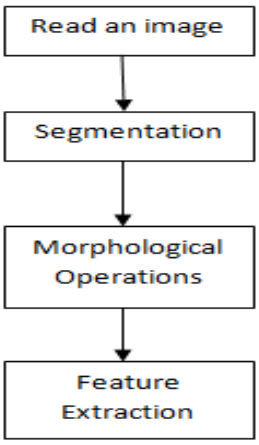

Fig. 1 Overview of flowchart.

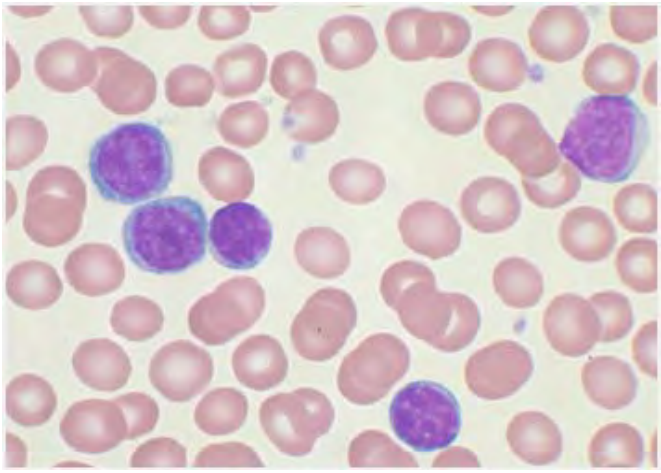

Fig. 2 Input Image.

The blood sample input image is shown in fig.(1) above. The detailed version of methods as follows.

A. Segmentation:

The blood sample images are segmented using the clustering algorithm. We used Adaptive k-means clustering algorithm [1] for segmentation. This algorithm can be applied for both color and gray images. It is based on k-means clustering which suffer from the following disadvantages. 1) Prior prediction of the number of clusters or the K value, which may raise an empty clusters for higher values. 2) It is slow and scales poorly on the time. 3) It may find worse local optima and 4) Different initial partitions can result in different final clusters, which leads to inconsistent output for the same image. In this paper we have used an Adaptive K-means algorithm to find clusters. This algorithm is adaptive in nature and proceeds as follows:

Starting with random selection of K initial seeds from the input image. These random K elements form the seeds of clusters and each element in the having them also form the properties of cluster.

The function uses Euclidean distance from centroid to quantify how close two objects are to each other. If the distance is within the threshold then the element is merged into the closest cluster by this centroid may change. In addition, the distance of the affected cluster from every other cluster, as well as the minimum distance between any two clusters and the two clusters that are closest to each other. Based on the distance clustering is done. By repeating these steps until all the elements have been clustered. When all elements in the input are clustered, algorithm stops.

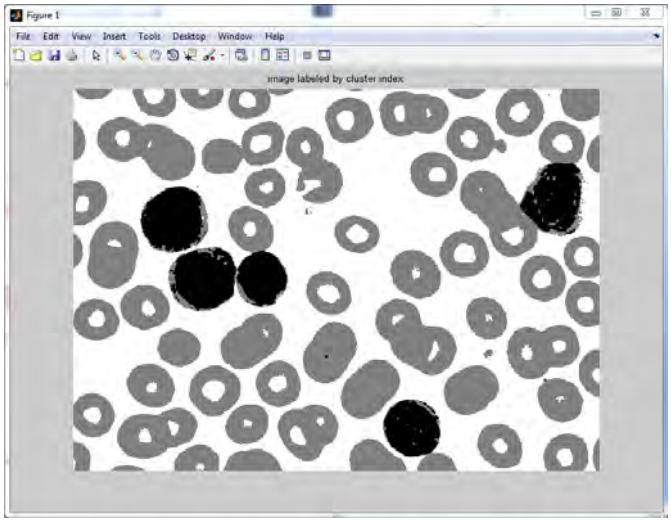

Fig. 3 Cluster Index.

The above fig.(3) shows the result of the clustering algorithm which returns an cluster centre corresponding to a cluster and label every pixel in the image with its cluster index.

Using cluster index we can separate the objects as shown in fig.(4) below.

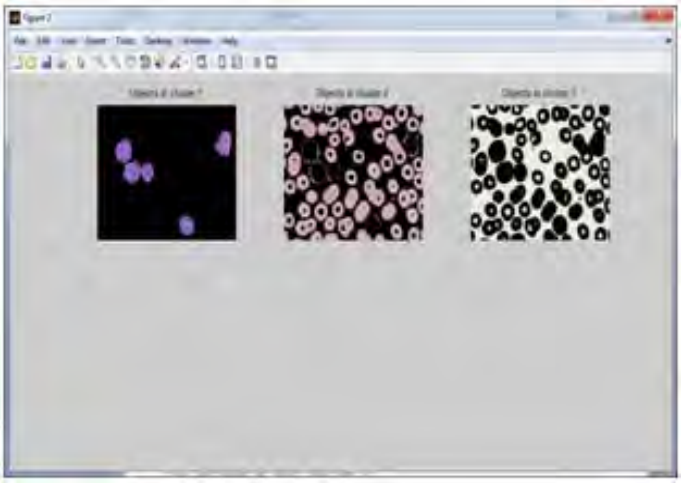

Fig. 4 Objects in Clusters.

B. Morphological Operations

After the clustering is done the following operations are performed on the object in the cluster which is closed to the required data. First the image is converted into binary image. The output image replaces all pixels in the input image with luminance greater than level with the

value 1 and replaces all other pixels with the value 0.

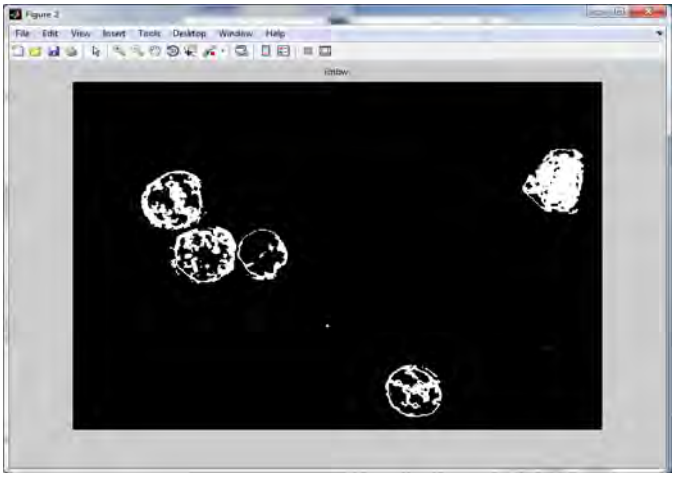

Fig. 5 Binary Cluster Object.

Second in the binary image holes are filled using the function imfill in the MATLAB[2]. A hole is a set of background pixels that cannot be reached by filling in the background from the edge of the image.

Now the image is enhanced by applying two dimensional median filter[3]. A media filter is more effective than convolution when the goal is simultaneously reduce noise and preserve edges. Each output pixel contains the median value in the (8-by-8) neighborhood around the corresponding pixel in the input image. medfilt2 pads the image with 0s on the edges, so the median values for points within one-half the width of the neighborhood ([m n]/2) of the edges might appear distorted. This is done to trace the boundaries of the cells in the image.

$$f'(m, n) = med\,\{(-k \le u, v \le k)\ \{f(m+u, n+v)\}. \qquad (1)$$

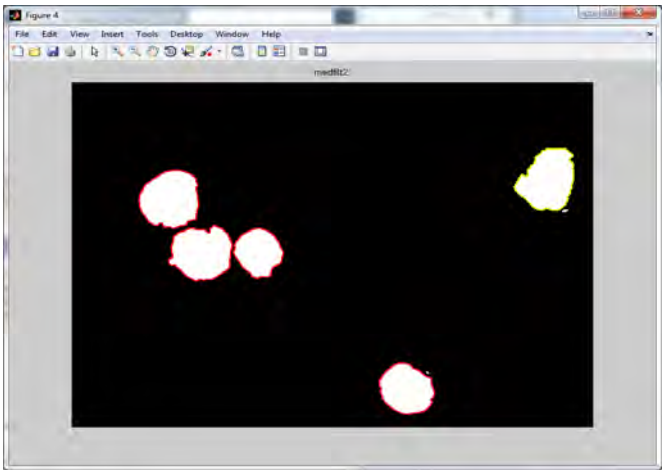

Fig. 6 After Filling Holes and Filtering.

The fig.(6) shows the image after the operations performed and tracing the boundaries of cells. bwboubdaries is used to trace the exterior boundaries of regions.

C. Feature Extraction

After performing all the operations on the image the features[4] like mean, maximum and minimum intensity and the area of the cells are calculated. The total number of cancer cells in the image are also calculated. These are done using the regionprops function in the MATLAB. The cell having maximum area is colored in yellow as shown in fig.(7) and the remaining cells are colored in red.

Using image processing techniques[5] the extraction is done.

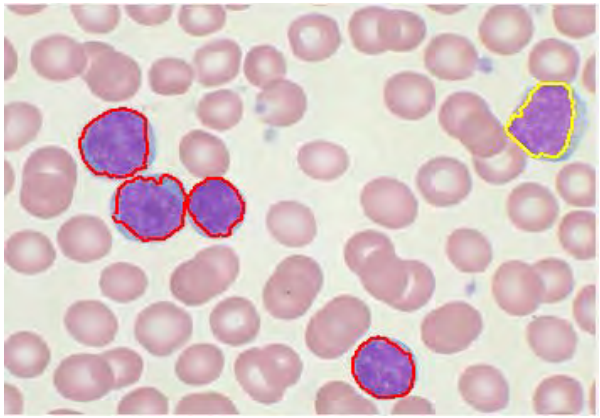

Fig. 7 Boundary Tracing.

For extraction of regions found in the boundaries in the fig.(7), the values we get are shown in the below fig.(8).

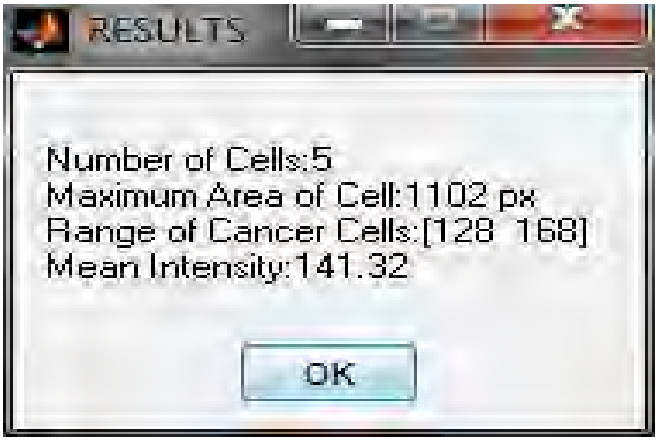

Fig. 8 Feature Extraction.

All the result figures are of image01. The results obtained are the total number of cancer cells in the image are 5 and the cancer cell having maximum area(yellow) is 1102 pixels, the range of intensities of the cancer cells is [128 168] and the mean intensity is 141.32 as shown in fig.(8).

3. Results

For the analysis of the steps proposed for feature extraction of the images is shown in the below table (1). This shows the image number, range and mean intensity of the cancer cells.

Table 1: Result Values

Image#	Range	Mean Intensity
Image01	[128 168]	141.32
Image02	[128 177]	143.16
Image03	[128 165]	142.27
Image04	[128 161]	141.99
Image05	[130 159]	145.26
Image06	[128 161]	142.90
Image07	[132 148]	143.43
Image08	[128 138]	132.77
Image09	[128 177]	142.06
Image10	[128 154]	143.86

These values are useful for further analysis of the cancer. We get the above results.

4. Conclusions

Blood cancer is an major problem which affect the production and function of the blood cells. It is important to detect in early stages. In this paper we proposed a framework to read and analysis the blood sample images which helps in detection. These images are segmented using Adaptive K-means clustering algorithm for the analysis. And we have calculated the mean intensities of the blood cancer cells which is useful for further analysis.

References

[1] Arjun Nelikanti, Narasimha LV Prasad, and Naresh M. Goud. "Colorectal Cancer MRI Image Segmentation Using Image Processing Techniques."International Journal on Computer Science and Engineering 6.7 (2014): 280.
[2] Matlab user manual – Image processing toolbox, MathWorks, Natick, 1999.
[3] Lim, Jae S., "Two-Dimensional Signal and Image Processing", Englewood Cliffs, NJ, Prentice Hall, 1990, pp. 469-476.
[4] Mark Nixon, Alberto S. Aguado, "Feature Extraction & Image Processing for Computer Vision", Second Edition, 2008
[5] Gonzalez and Woods, 2008. “Digital Image processing” 3 Edition, Prentice Hall.