

Acute Lymphocytic Leukemia Detection from Blood Microscopic Images

Sulaja Sanal

M. Tech student, Department of CSE.
Sree Budhha College of Engineering for Women
Elavumthitta, India

Lashma. K

Asst. Prof., Department of CSE.
Sree Budhha College of Engineering for Women
Elavumthitta, India

Viji Balakrishnan

Asst. Prof., Department of CS.
College of Applied Science (IHRD)
Mavelikkara, India

Abstract— The classification of blood cells is important for the evaluation and diagnosis of many diseases in medical diagnosis systems. Acute Lymphocytic Leukemia (ALL) is a type of childhood blood cancer and is mostly seen in children below 7-8 years. It can be dangerous if left untreated and causes death. Detection of ALL can be done through the analysis of white blood cells (WBCs) also called as leukocytes. Usually the analysis of blood cells is performed manually by skilled operators. This manual techniques have numerous drawbacks, such as slow analysis and a non-standard accuracy. It all depends on the skill of the operator. Hence many automated systems are using in order to analyze and classify the blood cells, but most of which produces only partial results. The main steps of this work are image preprocessing, WBC extraction, separation of adjacent WBCs, feature extraction and classification. Image preprocessing is done by converting RGB images into Lab color space images. It is done to enhance the visual appearance of the image and to reduce the memory requirements. Then the WBCs are identified by using fuzzy C means clustering algorithm. Adjacent WBCs are a major challenge while performing the feature extraction in the later stages. For avoiding that, separation of adjacent leukocytes is done by using Marker based watershed segmentation. For feature extraction, the features of WBC such as area, energy, entropy etc. are considered. To detect whether the patient is leukemic or not, a neuro-fuzzy classifier is used.

Keywords— ALL, Fuzzy C Means Clustering, Watershed segmentation, feature extraction, classification

I. INTRODUCTION

Leukemia is also called blood cancer, which is related with white blood cells. When abnormal white blood cells begins to continuously replicate itself, this bone marrow disorder arises. Acute leukemia is a rapid progressing disease that affects mostly cells that are not matured. Acute leukemia is classified according to the French-American- British (FAB) classification into two types: Acute Lymphocytic Leukemia (ALL) and Acute Myelogenous Leukemia (AML). ALL progresses in a faster manner by replacing healthy cells that produce functional lymphocytes with leukemia cells

which can't mature properly. It is usually difficult to diagnose, because the exact cause of ALL is still unknown. In addition, the symptoms of the disease are very similar to flu or other common diseases, such as fever, weakness, tiredness, or aches in bones or joints. Its detection can be done with a complete blood count [1]. The figure shown below are the blood microscopic images of cancerous and non-cancerous patients respectively.

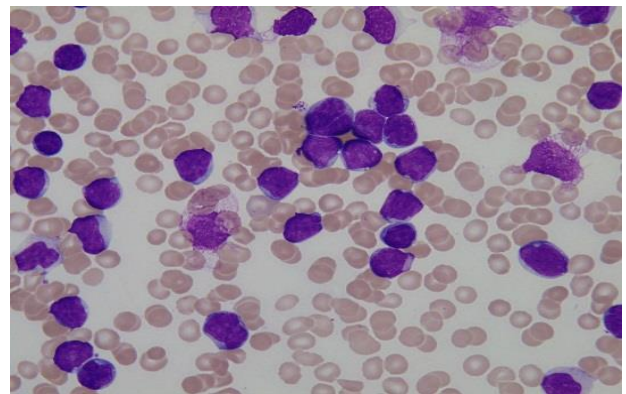


Figure 1. Blood microscopic images of cancerous patients.

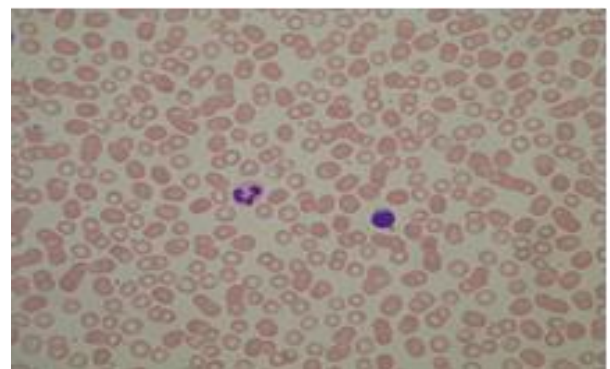


Figure 2. Blood microscopic images of non-cancerous patients.

But the main difficulty in the manual blood test is that it mainly depends on the operator's skill. The automated detection of ALL from blood microscopic images usually avoids the problems of manual testing of blood smear and the accuracy is also increased. It reduces the computational time and thus increases the efficiency. Automated Acute Lymphocytic Leukemia detection consist of the following steps:-

1. Image preprocessing
2. WBC identification
3. Feature selection
4. Classification

II. RELATED WORKS

Scotti et al. in [2] uses the same procedure such as identification of WBC, extraction of nuclei from that, feature extraction and classification. The main difference is that, in [2], at first the leukocytes are separated as sub images from the whole image and then identifies the nucleus from the sub images and it classifies the presence of leukemia using neural network.

Lorenzo Putzu et al. in [3] use ALL-IDB image database for getting the blood microscopic images for processing. The main step in this are identification of WBC, identification and separation of grouped WBCs, image cleaning, nucleus and cytoplasm selection, feature extraction and classification.

Sos Agaian et al. in [4] uses the same steps. But different methods are applied in each step. For image preprocessing, RGB to lab color space conversion is used. For WBC identification, K-means clustering is used. For classification, SVM is used. The main advantage of [4] is that the system is applied to complete blood smear images containing multiple nuclei. Many other systems process only sub images and it requires more computational time and memory. Two new features, such as Local Binary Pattern and Hausdorff dimension (HD), have been used. This result is then compared with the results of other existing models. Many other works are described in [5].

III. PROPOSED SYSTEM

The major steps in this work are image pre-processing, leukocyte identification, separation of adjacent leukocytes, feature extraction and classification. The overview of proposed system is shown in the figure .3.

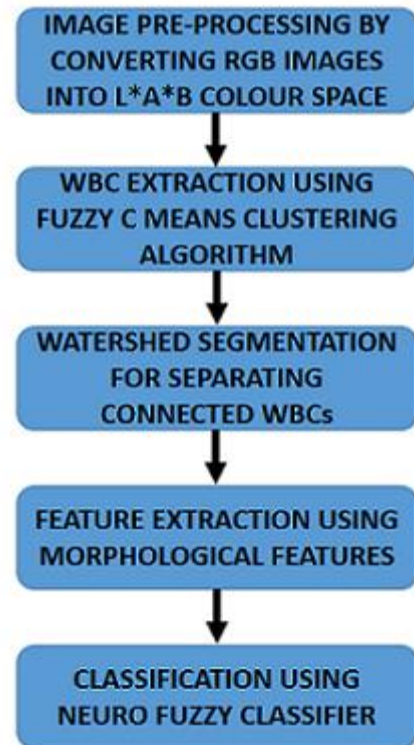


Figure 3. System overview.

A. Image aquisition

Fabio Scotti provided a public supervised image datasets (ALL-IDB) [6], [7] for study purpose. It is used to test and compare cell segmentation algorithms and the classification of the Acute Lymphocytic Leukemia. Two types of datasets are available. The ALL-IDB1 is used both for testing, segmentation capability of algorithms, the classification systems and image pre-processing methods. In ALL-IDB2, segmented WBCs are present to test the classification of blast cells.

B. Image preprocessing

The main aim of image preprocessing is to enhance the visual appearance of images. In this, image preprocessing is done by converting the RGB image into L^*a^*b color space. L in Lab color space represents lightness of the color, a represents position between red or magenta and b represents position between yellow and blue. RGB to Lab color space conversion makes the device dependent images to device independent.

C. WBC extraction

For WBC extraction, segmentation is used. The main aim of image segmentation is to extract important information from an input image. In this work, for the extraction of WBC, Fuzzy C means clustering algorithm is used. It is a modified K-Means clustering method. The algorithmic steps of fuzzy C means clustering is given below.

1) We are to calculate the c center vector $\{V_{ij}\}$ for each step.

$$V_{ij} = \frac{\sum_{k=1}^n (\mu_{ik})^m x_{kj}}{\sum_{k=1}^n (\mu_{ij})^m} \quad (1)$$

2) Calculate the distance matrix $D_{[c,n]}$.

$$D_{ij} = \left(\sum_{j=1}^m (x_{kj} - v_{ij})^2 \right)^{1/2} \quad (2)$$

3) Update the partition matrix for the r^{th} step, $U^{(R)}$ as

$$\mu_{ij}^{r+1} = \left(1 / \sum_{j=1}^c (d_{ik}^r / d_{jk}^r)^{2/m-1} \right) \quad (3)$$

Fuzzy C means clustering algorithm is more accurate than K means clustering. In K means, the data points exactly belong to one cluster. But sometimes data points are not much far to identify. But this situation can be easily handled in FCM, because in this the data points may belong to more than one cluster.

D. Separation of adjacent leukocytes

Marker based watershed segmentation is useful for separating connected components. Algorithmic steps of Marker-based watershed segmentation.

- Step1: Read the Color Image and Convert it to Grayscale.
- Step2: Use the Gradient Magnitude as the Segmentation Function.
- Step3: Mark the Foreground Objects.
- Step4: Compute Background Markers.
- Step5: Compute the Watershed Transform of the Segmentation Function.
- Step6: Visualize the Result.

At first, the RGB image is converted into grayscale image. Next step is to compute a segmentation function. In this sobel edge mask is used as segmentation function. Direct watershed transform results in over segmentation. In order to avoid this, additional image enhancement techniques such as erosion and dilation are done.

E. Feature extraction

It is defined as the transformation of input image into a set of features. Feature selection mainly influences the performance of classifier. Therefore, a correct choice of features is a very crucial step [4]. The features such as energy, entropy and area are used in this work for feature extraction.

F. Classification

Classification is the task of assigning a label from one of the known classes to the unknown test vector. For classification, Adaptive Neuro Fuzzy System (ANFIS) is used. To incorporate the best characteristics of neural networks and Fuzzy systems, ANFIS is used [8].

IV. RESULTS AND DISCUSSION

In this work blood microscopic images from ALL-IDB is taken as dataset. In this 15 cancerous and 15 non-cancerous images are used for training purpose.

At first the images in RGB format are converted into Lab color space format as a part of image pre-processing step. The figure.4 shown below represents the RGB image and figure 5. shows its corresponding Lab image.

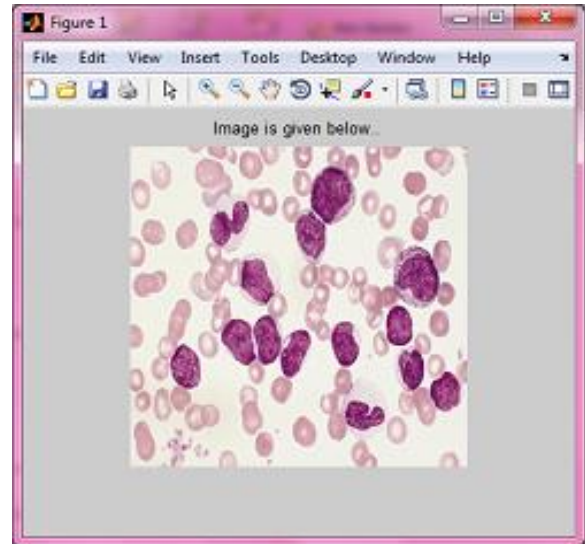


Figure 4. RGB image.

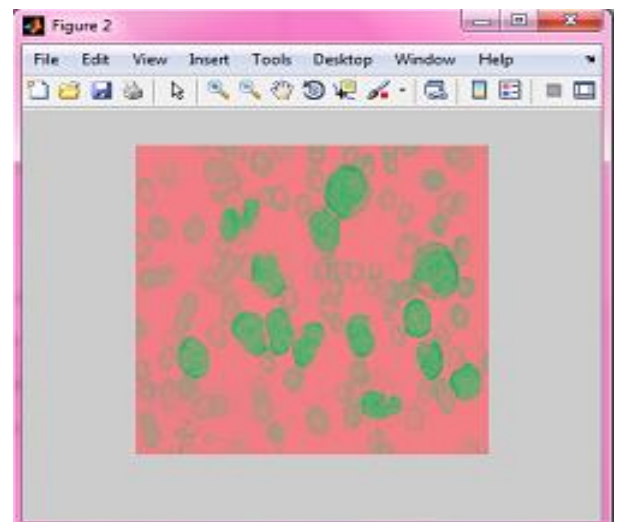


Figure 5. Lab color space image.

The next step is to identify the White Blood Cells from the Lab color space image. For that Fuzzy C Means clustering algorithm is used. The resultant image after the application of this algorithm is shown in the figure. 6.

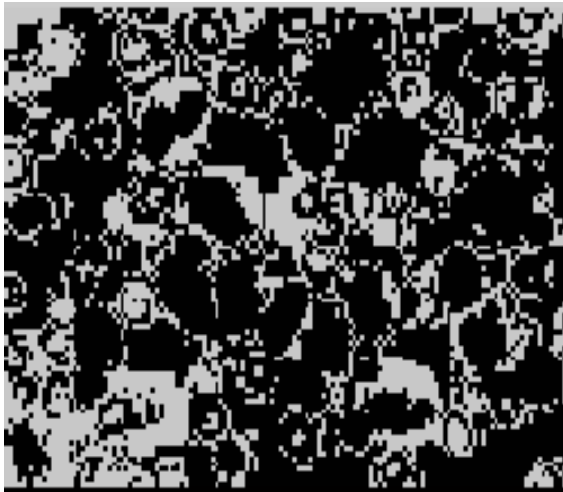


Figure 6. Output of FCM.

But the presence of adjacent leukocytes causes problems in the next stage, feature extraction. Hence Marker based watershed segmentation is used to separate adjacent leukocytes. The figure 7. represents the result of Marker based watershed segmentation.

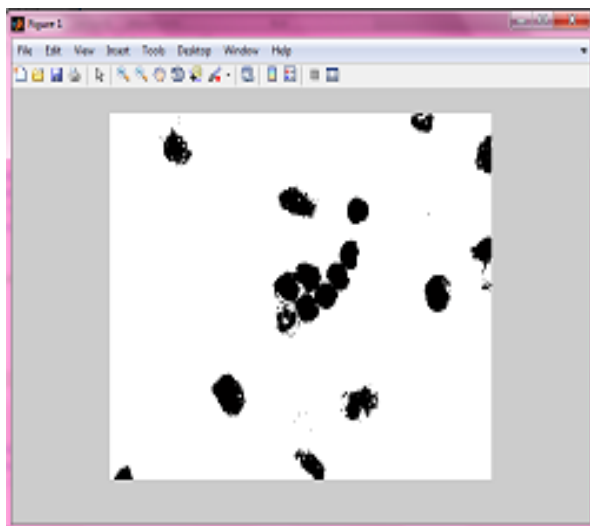


Figure 7. Marker based watershed segmentation.

The next step is to extract features from the resultant image. For feature extraction, features such as energy, entropy and area are considered. After considering these features, training is done. Training and testing are done using neuro-fuzzy classifier. The result of training is shown in figure 8.

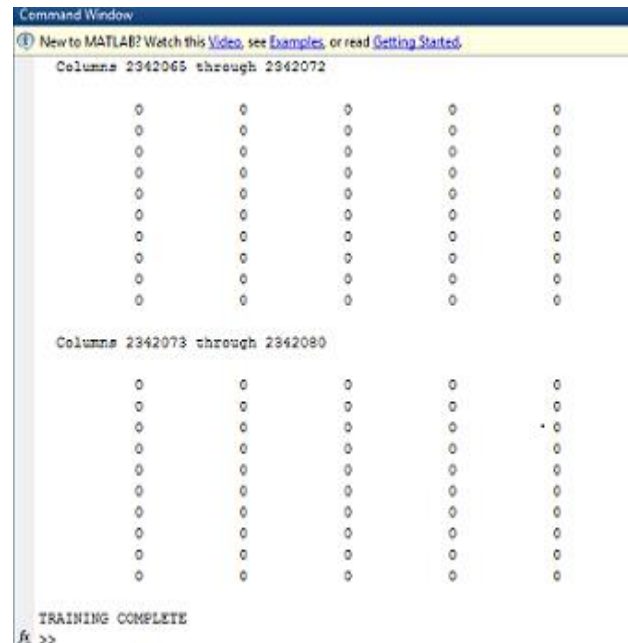


Figure 8. Result of training.

The result of feature extraction phase is shown in figure 7. for a particular image. Its energy, entropy and area values are calculated.

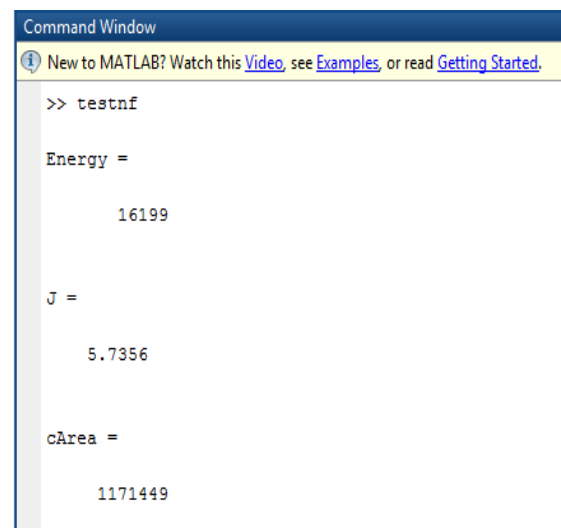


Figure 9. Result of feature extraction.

Testing is done after training process to check whether the patient is leukemic or not. The figure 10 and 11 shows the training result of a cancerous and non-cancerous patient respectively.

```

a =
-9.6555e+04

e =
1.0011

j =
0.9887

Cancer Detected
fx >>

```

Figure 10. Test result of cancerous patient.

```

a =
0.9622

e =
3.0616e-05

j =
1.0636e-04

No Cancer Detected
fx >>

```

Figure 11. Test result of a non-cancerous patient.

V. CONCLUSION

Manual examination of blood microscopic images under microscope remains an important screening procedure for ALL. But sometimes, it is not sufficient to make the correct diagnosis of acute lymphocytic leukemia. Such human visual evaluation is also time consuming and inconsistent when compared to various automated technique available nowadays. These automated techniques uses different algorithms which are more accurate and fast. In this, an automated system for the detection of acute lymphocytic leukemia is implemented. There are mainly five phases such as image pre-processing, segmentation to extract leukocytes from the blood microscopic images, separation of adjacent leukocytes, feature extraction and classification. RGB images are converted to Lab color space as part of image pre-processing. This is mainly done to make the device dependent image to device

independent. Various other color space conversions are available. Among that lab color space conversion is most suitable in this work, because it is invariant to illumination changes. Then the next step is to extract leukocytes from the images. For that, segmentation is used. Fuzzy C Means clustering algorithm is used in this phase for segmentation. But the presence of adjacent leukocytes will make the subsequent steps more difficult. This can be solved by applying Marker based watershed segmentation algorithm. The next is feature extraction phase. For this, features such as energy, entropy and area are used. Energy and entropy are texture features and area is a shape feature. To detect whether the patient is leukemic or not, a classifier based on neuro fuzzy system is used.

The main advantage of this work is that the entire phases are carried out by considering the image as a whole. Hence it is faster than the other automated systems.

REFERENCES

- [1] Hayan T. Madhloom, Sameem Abdul Kareem, Department of Artificial Intelligence, Faculty of Computer Science and Information,, "A Robust Feature Extraction and Selection Method for the Recognition of Lymphocytes versus Acute Lymphoblastic Leukemia", International Conference on Advanced Computer Science Applications and Technologies, 2012.
- [2] F. Scotti, "Automatic morphological analysis for acute leukemia identification in peripheral blood microscope images", in Proc. CIMSA, 2005, pp. 96-101.
- [3] Lorenzo Putzu, Giovanni Caocci, Cecilia Di Ruberto, "Leucocyte classification for leukaemia detection using imageprocessing techniques", Artificial Intelligence in Medicine, 2014.
- [4] Sos Agaian, Monica Madhukar, and Anthony T. Chronopoulos, Senior Member, IEEE, "Automated screening system for acute myelogenous leukemia detection in Blood Microscopic Images", IEEE Transactions on Image Processing, Vol. 23, No.2, February 2014.
- [5] Sulaja Sanal, "Automated Detection of Acute Lymphocytic Leukemia-A survey" International Journal of Engineering Research and General Science, Volume 3, Issue 3. May-June 2015.
- [6] Donida Labati R, Piuri V, Scotti F, "ALL-IDB: the acute lymphoblastic leukemia image Database for image processing", In Macq Benot, Schelkens Peter, editors. Proceedings of the 18th IEEE ICIP international conference on image processing, September. Brussels, Belgium: IEEE Publisher; 2011.
- [7] <http://www.dti.unimi.it/fscotti/all>
- [8] D.Wermeser. G. Haussmann. And C.E. Liedtke, "Segmentation of Blood Smears by Hierarchical Thresholding", Computer Vision, Graphics, and image Processing, Vol25, 1984, pp. 151-168.