International Journal of Current Engineering and Technology
E-ISSN 2277 – 4106, P-ISSN 2347 – 5161

Available at http://inpressco.com/category/ijcet

Research Article

Automated Screening System for Acute Myelogenous Leukemia Detection using Layer Subtraction

Anitta K Varghese†* and Nisha J S†

†Department of Electronics and Communication Engineering, St. Joseph's College of Engineering and Technology, Palai, India

Accepted 30 Sept 2015, Available online 02 Oct 2015, **Vol.5, No.5 (Oct 2015)**

Abstract

Leukemia is one type of blood cancer. It mainly attacks the blood, bone marrow etc. One of the subtype of acute leukemia is Acute myelogenous leukemia (AML). Leukemia is detected by the microscopic analysis of peripheral blood smear. But this microscopic assessment is time consuming and is governed by haematologists. To overcome this drawback automatic screening system for leukemia detection arises. For automatic detection of leukemia in peripheral blood samples, an efficient methodology such as pre-processing, nuclei segmentation, feature extraction, classification is proposed. The advantage of this technique is its simplicity, classification of complete blood smear images and helps to segment and detect nucleated cells. Fourty microscopic blood images were tested and the proposed method obtains 97.56 percentage accuracy for the localization of cells and to separate it from the complete images.

Keywords: *Acute myelogenous leukemia (AML), classification, feature extraction, segmentation.*

1. Introduction

Medical imaging is an important interpretation and visualization methods in biology, medicine and is one of the fastest growing fields in medicine, clinical setting and research and development. Analyzing the images helps to gather information, detection of diseases, diagnosis diseases, control and therapy, monitoring and evaluation. The disorders in the blood are identified through visual inspection of microscopic images of blood cells. This identification helps to classify certain diseases related to blood.

Cancer is a generic term that describes a group of malignant diseases with cells displaying uncontrolled and invasive growth along with metastasis. It is occurred in blood, bone, lymph node, skin, colon, breast, or nerve tissue. Abnormal growth of white blood cells produced by bone marrow causes leukemia and it comes from the Greek word “leukos” meaning “white” and “aim” meaning “blood”.

Leukemia can be chronic or acute. At earlier stage in chronic leukemia, leukemic cells can make tasks such as normal white blood cells. Gradually they will become severe chronic leukemia. In acute leukemia, leukemia cells cannot make tasks like normal white blood cells. Many leukemia cells will grow rapidly and become severe in a short time. Most acute leukemia patients are referred to specialist units for evaluation and treatment. Leukemia can be diagnosed by bone marrow tests and blood tests based on the fact that neutrophils, platelets are decreased and white blood cell count are increased with immature blast cells (C. Haworth, *et al*, 1981). Therefore, haematologists routinely examine blood smear under microscope for proper identification and classification of blast cells. The significant symptom of leukemia is the excess number of blast cells in peripheral blood. Currently the best available treatment is chemotherapy, which unfortunately also kills normal body cells along with the cancerous ones.

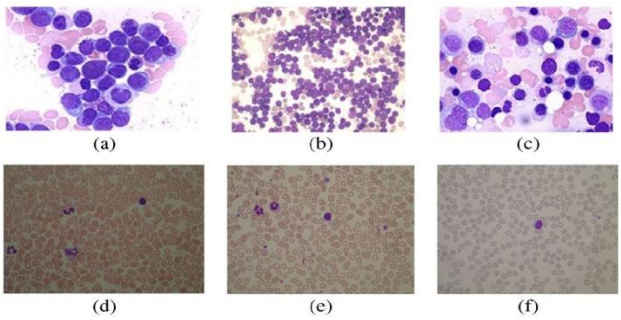

Fig.1 (a),(b),(c) depicting unhealthy cells from AML patients and (d),(e),(f) depicting healthy cells from non-AML patients.

Mainly there are four types of leukemia, Acute Myelogenous Leukemia, Acute Lymphoblastic Leukemia, Chronic Myeloid Leukemia, and Chronic Lymphocytic Leukemia. AML, a serious illness caused

*Corresponding author: **Anitta K Varghese**

due to the abnormal development and growth of nongranular white blood cells. It starts in the bone marrow blast cells which develop into granulocytes, i.e. white blood cells containing granules. As the blast cells grow they will hamper the body's natural ability to fight against infection and stop bleeding. Therefore, this disease requires immediate treatment. Some of the symptoms are weakness, fever, aches in bones. If these symptoms are present, blood tests, such as a renal function, full blood count, electrolytes, liver enzyme and blood count have to be done (F. Scotti,2005). Since the precise cause of AML is unknown it is very difficult to diagnose.

Rest of the paper is organized as follows: Section 2 summarizes the system model. Section 3 describes the methodology and design in detail. Results and discussion in section 4. Experimental Results are discussed in section 5 and paper is concluded in section 6.

2. System Model

The block diagram of system model is given in Fig: 2. The AML images generated by digital microscopes are usually in RGB color space. It is given as input for segmentation.

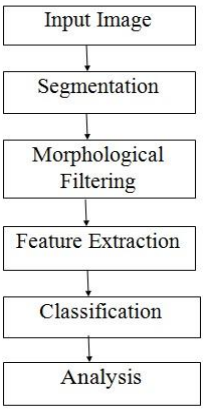

Fig.2 Block diagram of system model

Segmentation is done using K-mean clustering and layer subtraction technique. Since using K-mean algorithm the blue nuclei region was not extracted properly. To improve the segmentation of blue nuclei a new method known as layer subtraction is introduced. From the segmented image various features such as shape feature, GLCM features, color feature, hausdroff dimension with and without LBP is extracted. This feature plays an important role in classification. Support vector machine is used for classification. The classifier accuracy is analyzed using true positive rate, false positive rate. Then comparison is made with the K-mean and layer subtraction segmentation and found that layer subtraction gives better accuracy and image quality than K-mean segmentation.

3. Methodology and Design

3.1 Image Acquisition

The data images used in this experiment are accessed from the American Society of Hematology. The database consists of 100 images, 50 from AML patients and 50 from non-AML patients. The resolution used for classification is 184 × 138 pixels.

3.2 Segmentation

The process of separating a digital image to multiple parts is known as segmentation. Using this technique a label is assigned to every pixel. The pixels having the same label have certain characteristics. In this system segmentation is performed for extracting the nuclei from the AML image. Here segmentation is done using K-mean segmentation and layer subtraction segmentation

3.2.1 Segmentation using K-mean clustering

According to measured intrinsic characteristics or similarity a formal algorithms for clustering objects is cluster analysis. K-mean is one of the clustering techniques. A given data set is classified into certain number of clusters fixed a priori using this technique. and it is one of the simplest unsupervised learning algorithms. Here the CIEL*a*b* image is given as input to the K-mean clustering and obtain three clusters.

K-mean algorithm requires three user-specified parameters: the number of clusters k, cluster initialization, and distance metric. So initially the image is converted into the CIEL*a*b* color space. Based on the corresponding *a and *b values in the L*a*b color space each pixel is classified into clusters. It corresponds to nucleus, background and other cells. Here the cluster that contains the blue nucleus, which is required for the feature extraction, is considered.

3.2.2 Segmentation using layer subtraction

The RGB image is given as input for segmentation. Each layer of the rgb image is separated independently. Then the blue layer is subtracted from the red layer. The output of layer subtraction is converted into binary image and threshold value is applied on it. Each layer of the input image gets multiplied with the binary value. Thus the blue nuclei get segmented. Thus by using this segmentation image quality and accuracy gets increased.

3.3 Morphological Filtering

To improve the perceptibility and visibility of these regions morphological filtering such as sobel, canny, dilation, hole filling is applied. This action is performed in the segmented output.

3.4 Feature Extraction

Feature extraction is a technique used to transform the input data into set of features and is a form of

dimensionality reduction. The features set will extract the important information from the input data if the features are extracted correctly.

Some of the features are Hausdorff dimension with and without LBP, Shape features, Gray level co-occurrence features, Color feature.

3.4.1 Hausdorff dimension

A quantity that gives an indication of how completely a fractal appears to fill space is the fractal dimension D. Theoretical fractal dimensions are the packing dimension, the HD, and the Rényi dimension. Due to their ease of implementation, the box-counting dimension is widely used in practical cases. In this, the number of boxes covering the point set is a power-law function of the box size. The exponent of such power law is estimated as D. All fractal dimensions are real numbers that characterize the roughness of the objects. Myeloblast can be differentiated using perimeter roughness of the nucleus. The procedure for HD measurement using the box counting method is described below:

- binary image in obtained from the gray-level image of the blood sample.
- edge detection technique is employed to trace out the nucleus boundaries.
- edges are superimposed by a grid of squares.
- the HD may then be defined as follows:

$$HD = \frac{\log(R)}{\log(R(s))} \quad (1)$$

Where R is the number of squares in the superimposed grid, and R(s) is the number of occupied squares or boxes. Higher HD signifies higher degree of roughness.

3.4.2 Local Binary Pattern

The LBP method has proved to outperform many existing methods, including the linear discriminant analysis and the principal component analysis. In this algorithm the local neighbourhood is defined as a set of sampling points evenly spaced on a circle centered at the pixel to be labelled allows any radius and number of sampling points. M represents the number of samples that determines the number of points that are taken uniformly from the contour of the circle. If needed, these points are interpolated from adjacent pixels. Each grayscale pixel P is compared with these sample points one by one. If the centre point P is larger than the current neighbourhood sample point I, the result is a binary zero; otherwise, the result is a binary one. When doing this operation, for example, clockwise from a certain starting point, the result will be a binary pattern with length M.

Here the segmented images were extracted using K-mean clustering and layer subtraction, and then the LBP operator was applied on them before calculating the HD. Two sets of values were extracted: first, HD of the images without applying LBP, and second, HD of the images after applying LBP. When comparing these values, it was observed that the LBP operator enhanced the overall performance by a very high margin.

3.4.3 Shape features

One of the shape features that help to classify the AML and NON-AML image is the compactness. For shape analysis of nucleus, region- and boundary-based shape features are extracted. These features are extracted from the binary-equivalent image of the nucleus where the nucleus region is represented by the nonzero pixels. Some of the shape features are:

- Area: It is determined by counting the total number of non zero pixels within the image region.
- Perimeter: Measured by calculating distance between successive boundary pixels.
- Compactness: It is defined as the measure of nucleus.

$$\text{Compactness} = \frac{\text{Perimeter}^2}{\text{Area}} \quad (2)$$

- Solidity : The ratio of actual area and convex hull area is known as solidity and is also an essential feature for blast cell classification.

$$\text{Solidity} = \frac{\text{Area}}{\text{ConvexArea}} \quad (3)$$

- Eccentricity : This is used to measure how much a shape of a nucleus deviates from being circular. It's an important feature since lymphocytes are more circular than the blast.

$$\text{Eccentricity} = \frac{\sqrt{a^2 - b^2}}{a} \quad (4)$$

- Elongation : Abnormal bulging of the nuclei is also an feature which signifies towards leukemia. Hence the nucleus bulging is measured in terms of a ratio called elongation. This is defined as the ratio between maximum distance R_{max} and minimum distance R_{min} from the center of gravity to the nucleus boundary.

$$\text{Elongation} = \frac{R_{max}}{R_{min}} \quad (5)$$

- Form factor : It is a dimensionless parameter which changes with surface irregularities.

$$\text{Formfactor} = \frac{4 * pi * \text{Area}}{\text{Perimeter}^2} \quad (6)$$

3.4.4 GLCM features

GLCM stands for gray-level co occurrence matrix. A function of the spatial variation in pixel intensities is known as texture. It is one of the image analysis techniques. Second-order statistics helps to describe the gray-level pixel distribution. This is depicted in 2-D gray-level co-occurrence matrices that can be computed for various orientations and distance. Haralick defined some of the statistical measures to extract textual characteristics from the information contained in the GLCM. Some of them are the energy, contrast, entropy, correlation.

3.4.5 Color features

The color-based feature is cell energy. It is also known as measure of uniformity. We define feature δ to be:

$$\delta = \sum_i \sum_j P^2(i, j) + (\sqrt{-1})\left(\frac{\sqrt{\sum_{i=1}^{n}(x_i - x')^2}}{n-1}\right) \quad (7)$$

Where $x' = \sum_{i=1}^{n} \frac{x_i}{n}$, P (i,j) represents the normalized GLCM element for the i^{th} row and j^{th} column, and $\sum_i \sum_j P^2(i, j)$ represents the ASM.

3.5 Classification

The selection of a classifier for classification is a challenging problem. Here support vector machine (SVM) classifier is used for making a decision surface for bisecting the two categories, i.e. AML and NON-AML, and also for maximizing the margin of separation between two classes. It is primarily a two-class classifier and it can be nonlinear or linear. Here, a linear SVM two-class classifier is used; because it is inexpensive, and provides a good performance.

3.6 Analysis

To ensure the effectiveness of the classifier and segmentation certain parameters are calculated. For calculating the effectiveness of classifier certain parameter such as sensitivity, precision, specificity and f-measure are calculated. These are defined in relation to the possible outcomes of the classifier system. The four possible outcome of the classifier is: true positives (TP), when AML are correctly identified; false positives, when NON-AML cells are identified as AML; true negatives, when NON-AML are correctly identified; and false negatives, when AML are identified as NON-AML.

Table 1 Performance Evaluvation Parameters

SI.No	Parameters	Formulae
1	Sensitivity	$\frac{TP}{(TP+FN)}$
2	Specificity	$\frac{TN}{(TN+FP)}$
3	Precision	$\frac{TP}{(TP+FP)}$
4	F-measure	$\frac{2*Precision*Sensitivity}{Precision+Sensitivity}$

For calculating the image quality of segmentation using K-mean and layer subtraction certain parameters such as AMBE and Entropy are calculated.

- *Absolute Mean Brightness Error*

Absolute Mean Brightness Error is defined as absolute difference between the mean input and output image. It is defined as:

$$AMBE = |E(X) - E(Y)| \quad (8)$$

where X and Y are input and output image. E(X) represents mean of input image and E(Y) represents the mean of output image. Small AMBE value indicates that brightness is preserved.

- *Entropy*

Entropy is used to measure the richness of details in the output images. It is defined as:

$$Entropy(p) = -\sum_{k=0}^{L-1} p(X_k) \log_2 p(X_k) \quad (9)$$

Higher entropy value indicates richness of details.

4. Results and Discussions

Simulation is done using MATLAB. MATLAB stands for matrix laboratory. For technical computing, it is a high-performance language. Image quality of segmentation is assessed using AMBE and entropy.

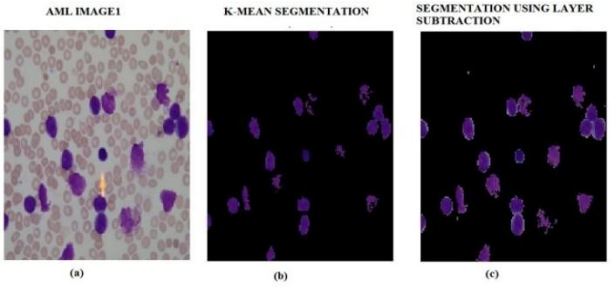

Fig.3 (a) AML image, (b) K-mean segmentation and (c) Segmentation using Layer subtraction

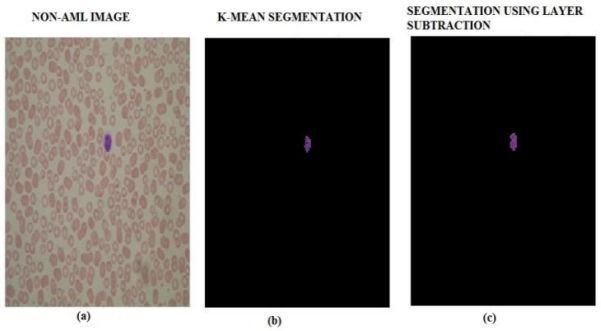

Fig.4 (a) NON-AML image, (b) K-mean segmentation and (c) Segmentation using Layer subtraction

Fig: 3 and Fig: 4 show the segmented output of AML and NON-AML image using K-mean and layer subtraction. Table 3 and 4 shows the image quality assessment of K-mean and Layer subtraction segmentation of the AML and NON-AML input image**.**

4.1 Image Quality Assessment of AML and NON-AML image using Layer subtraction and K-mean segmentation

From the Table 2 and 3 it is found that entropy increase for layer subtraction and AMBE decrease. Thus average intensity and richer details are obtained from proposed method.

Table 2 Image Quality Assessment of AML image

Sl.No	Methods	Segmentation using Layer Subtraction	K-mean Segmentation
1	Entropy	1.0204	0.6095
2	AMBE	137.1834	140.7218

Table 3 Image Quality Assessment of NON-AML image

Sl.No	Methods	Segmentation using Layer Subtraction	K-mean Segmentation
1	Entropy	0.0307	0.0236
2	AMBE	141.0935	141.1440

5. Experimental Results

From Table 4 it is understood that the accuracy of the proposed method is 97.56 percentage with one NON-AML are misclassified.

Table 4 Experimental Results

Sl.No	Methods	Sensitivity	Specificity	Accuracy
1	Existing Method	0.9500	0.9000	0.9268
2	Proposed Method	1	0.9497	0.9756

For training of AML and NON-AML 60 images are taken. For testing of AML and NON-AML 20 images are taken. Using the existing system 1 AML image is misclassified as NON-AML and 2 NON-AML images are misclassified as AML image. But by using the proposed system only 1 NON-AML image is misclassified as AML-image. Therefore the accuracy of the proposed system increased from 92.68 percentage to 97.56 percentage.

Conclusions

This paper has reported the design, development, and evaluation of an automated screening system for AML in blood microscopic images. The presented system performs better segmentation of the nucleated cells, feature extraction, classification and analysis than the k-mean clustering techniques. Features such as shape, texture, color is constructed to obtain all the information required to perform efficient classification. HD with LBP and color feature presents a good demarcation between AML and NON-AML cells.

References

R. C. Gonzalez and R. E. Woods,(2002), Digital Image Processing, *Prentice Hall,* Vol. 2nd edition.

Agaian.S, Madhukar. M and Chronopoulos A.T,(2014), Automated Screening System for Acute Myelogenous Leukemia Detection in Blood Microscopic Images, *Systems Journal, IEEE,* Vol.8,No.3, pp.995 – 1004.

M. Subrajeet, D. Patra,(2011), Automated Leukemia Detection in Blood Microscopic Images using Statistical Texture Analysis, *in Proc.Int. Conf. Commun., Comput. Security,* pp. 184–187.

C. Haworth, A. Hepplestone, P. Jones, R. Campbell, D. Evans, and M. Palmer,(1981), Routine bone marrow examination in the management of acute lymphoblastic leukemia of childhood, *J. Clin. Pathol*., Vol. 34, No. 5, pp. 483–485.

F. Scotti, (2005), Automatic morphological analysis for acute leukemia identification in peripheral blood microscope images, *in Proc. CIMSA,* pp. 96–101.

Medline Plus: Leukemia.National Institutes of Health. [Online].Available:http://www.nlm.nih.gov/medlineplus/ency/article/001299.htm

ASH Image Bank: American Society of Hematology. [Online].Available:http://imagebank.hematology.org/Default.aspx.