Published by :
http://www.ijert.org

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 5 Issue 01, January-2016

Microscopic Digital Image Processing of Acute Leukemia

Alaa Shakir Ahmed
AL Mansoura University Faculty of Engineering
Electronics & Comm. Dept.
Mansoura - Egypt.

Dr. M. Morsy
AL Mansoura University Faculty of Engineering
Electronics & Comm. Dept.
Mansoura - Egypt.

Prof. Mohy Eldin A. Abo-Elsoud
AL Mansoura University Faculty of Engineering
Electronics & Comm. Dept.
Mansoura - Egypt.

***Abstract*—The goal of this research is to identify the characteristics of Acute lymphoblastic leukemia (ALL). This research was conducted on a set of images of microscope blood samples that have been obtained from the "Oncology Center - Faculty of Medicine - Mansoura University Hospital - Egypt" is made up of 50 microscope image samples of blood infected , 50 microscope image of the blood samples is not infected. The microscope blood images are exposed to series of pre-processing steps which include resize image such as 512*512, 256*256, and contrast enhancement. By executing K-means clustering on the resultant images, the nuclei of cells under consideration are obtained. Shape features, texture features, color features and hausdorff dimension are then extracted for classification. A back propagation neural network was employed for classification. The results show that final performance evaluation for 100 microscope blood image to find out any infected images and non-infected images.**

Keywords — (Digital Image Processing , Acute Leukemia , Contrast Enhancement , K-Means Clustering , Features Extraction , Classification , Back Propagation Neural Network)

I. INTRODUCTION

Acute lymphoblastic leukemia (ALL), also known as acute lymphocytic leukemia, or acute lymphoid leukemia, is an acute form of leukemia, or cancer of the white blood cells, distinguished by the overproduction and cumulation of cancerous, immature white blood cells, known as lymphoblasts [1]. Leukemia is a group of hematological Tumors which usually affects blood, bone marrow, and lymph nodes. It is distinguished by proliferation of abnormal white blood cells (leukocytes) in the bone marrow without responding to cell outgrowth inhibitors [2].

II. DIGITAL IMAGE PROCESSING (METHODS)

The procedure for Acute lymphoblastic leukemia (ALL) classification in microscopic blood images consists of pre-processing (resize and contrast image), segmentation using (k-means clustering), feature extraction (shape – texture – color – HD) and classification (using back propagation neural network). The proposed system is shown in Fig. 1.

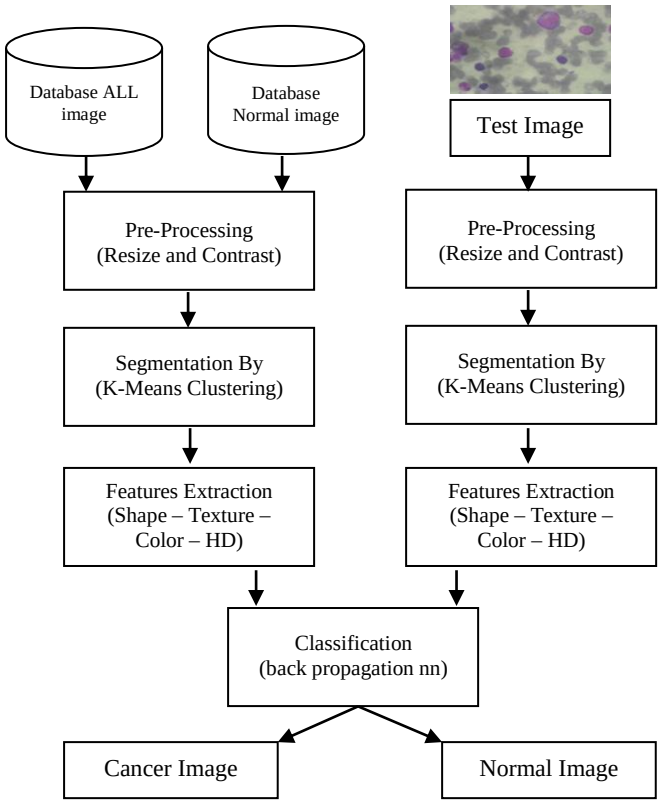


Fig.1. System Overview

A. *Pre-processing*

Pre-processing methods can be splitted into the two groups according to the goal of the processing:

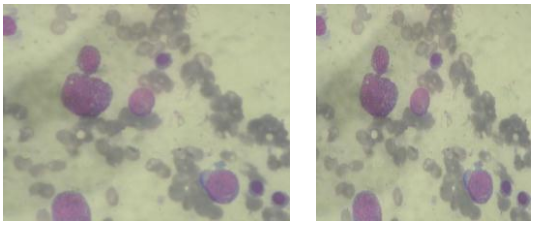
First: - Image resizing such as 512*512, 256*256 and 128 * 128.

Second: - Contrast enhancement by employ the "*fspecial*" function with the "*unsharp masking*" filter has the influence of making edges and fine detail in the image more crisp [3], then apply this mask filter on image by using "imfilter" function with Boundary Option "replicate" that Input array values outside the bounds of the array are assumed to equal the nearest array border value [4].

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 5 Issue 01, January-2016

Algorithm: Pre-processing using unsharp (Filtering)

- Read input image (X).
- Use Y = imresize (X, [512*512]).
- Create new variable color filtered for having same attribute that of input image using unsharp masking filter.
- Apply boundary option replicate filter algorithm on s Merge all three planes together eparated RGB planes.
- Merge all three planes together.
- Output (Y) as shown in Fig. 2.



(X) Input Image (Y) Resize 512*512 & Contrast Image

Fig. 2.

B. Segmentation

Segmentation is performed in two stages for extracting WBC nucleus from the blood microscopic images using color based clustering. Initial segmentation are completed by K-means clustering followed by nearest neighbor classification in L*a*b* space. K-means is a semi supervised clustering technique which is employ to create K clusters from n observations. It is aim to achieve partition such that objects within each cluster are as near to each other as possible, and as far from objects in the other clusters as possible [5]. Each pixel of an object is classified into four clusters based on corresponding a* and b* values in L*a*b* color space as shown in Fig. 3. The four clusters represents four regions i.e. RBC, WBC nucleus, cytoplasm and background stain. It was observed that WBC cytoplasm and RBC are classified into same cluster. In order to overcome the undesirable overlapping of regions, a second stage segmentation is performed using nearest neighbor classification. In the second stage we choose a sample region randomly from each of the four clusters acquired using K–means. The mean color of the each sample regions are calculated in a*b* space and those values act as color indicators. here each pixel in the L*a*b* space is distributing into any of the four classes by computing the Euclidean distance between that pixel and each color index. Each pixel of the whole image will be labeled to a specific color depending on the minimum distance from each index. The nucleus segmented RGB image is reconstructed from the labeled image. We have only considered the cluster which contains blue nucleus as it is required for feature extraction and hence leukemia detection. Few left out holes in the nucleus creates problem during texture extraction and hence they are filled using morphological reconstruction [6].

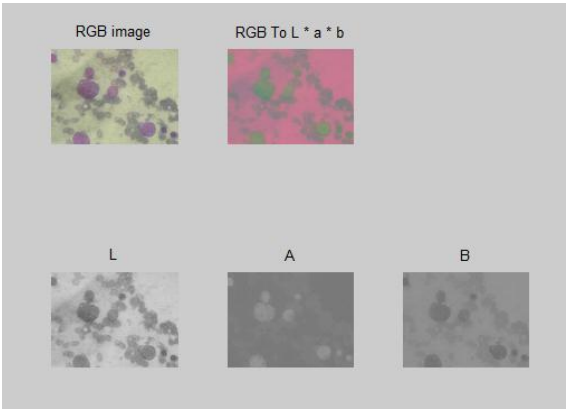
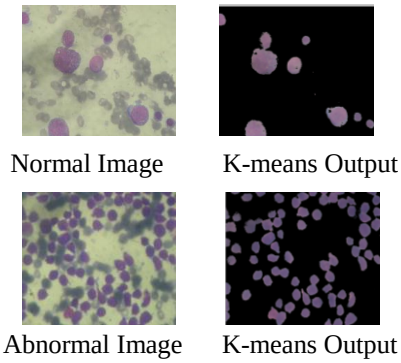


Fig. 3.

Algorithm Segmentation using K-means clustering

- Read Enhancement image (Y).
- [l a b] = Convert To L*A*B space
- A_B = Merge A*B space
- [r c] = get size (A_B)
- A_B_new = reshape(A_B, r*c , 2)
- Apply K-means clustering and output as shown in Fig. 4.



Normal Image K-means Output

Abnormal Image K-means Output

Fig. 4.

C. Sub Imaging

Sub images including single nucleus per sub image are obtained using bounding box technique [7]. Using image morphology [8] only those sub images are selected which contains only lymphocytes. The nucleus sub images of neutrophils, eosinophil's, and basophils are not considered for feature extraction as they are not associated with lymphocytic leukemia.

Fig. 5. shows the major steps and examples of input/output images:

Step one) Input Image

Step tow) Sobel edge enhancing: - It's enhances the borders of the membranes [9].

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 5 Issue 01, January-2016

Step three) Structured image dilation: - The morphological operator named dilation [10] has been employed to better connect to the separated points of the membrane border and make the perimeter of cell as a connected item (thicker more than one pixel).

Step four) Hole filling: -This step include of filling internal holes of the connected element with the largest area in the processed image [10, 11].

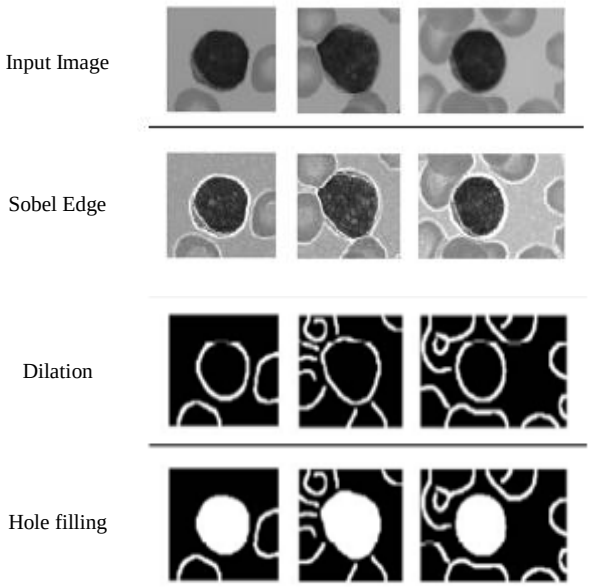


Fig. 5

Fig. 6. Separeted Nucleus Sub Images using Bounding box technique

D. Feature Extraction:
Feature extraction in image processing involves reducing the amount of resources required to describe a large set of data. In the present paper broadly four types of features are extracted (shape features, texture features, fractal dimension. In addition also color features are extracted from the nucleus image).

A: Shape Feature:- *According to the hematologist the shape of the nucleus is an essential feature for distinguish of blasts. boundary based shape features and Region are extracted for shape analysis of the nucleus. All the features are extracted from the binary equivalent image of nucleus with nonzero pixels represents the nucleus region. for each nucleus we make a quantitative evaluation by using the extracted features under two classes. region based and boundary based. This features are as follows:*

- *Area:* It was determined by computation the total number of nonezero pixels within the image region.
- *Perimeter:* the perimeter was measured by computation distance between the successive boundary pixels.
- *Compactness:* Compactness or roundedness is the measure of a nucleus as defined in (1).

$$Compactness = \frac{Perimeter^2}{Area} \quad (1)$$

- *Solidity:* The solidity is the ratio of actual area and the convex hull area and is also an essential feature for classification a blast cell. This measure is defined in (2).

$$Solidity = \frac{Area}{ConvexArea} \quad (2)$$

- *Eccentricity:* This parameter 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. To measure this a relation is defined in (3).

$$Eccentricity = \frac{\sqrt{a^2 - b^2}}{a} \quad (3)$$

where "a" is the major axis and "b" is the minor axis of the equivalent ellipse representing the nucleus region.

- *Elongation:* Abnormal bulging of the nucleus It's also an feature which indicates 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 and is given by (4).

$$Elongation = \frac{R_{max}}{R_{min}} \quad (4)$$

where R_{max} and R_{min} are maximum and minimum radii respectively.

- *Formfactor:* This is a dimensionless parameter which changes with surface irregularities and is defined as (5).

$$Formfactor = \frac{4 * pi * Area}{Perimeter^2} \quad (5)$$

B: Texture Feature-:The nucleus texture measurements were performed on a gray scale version of the nucleus images. These features were computed from the co-occurence matrices for each nucleus image. This includes:

- *Homogeneity*: It is a measure of degree of variation.
- *Energy*: The energy are used to measure uniformity.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 5 Issue 01, January-2016

- *Correlation*: This represents the correlation between the pixel values and its neighborhood.
- *Entropy*: It is Usually used to measure the randomness.
- *Contrast:* The contrast is a measure of the intensity contrast between a pixel and its neighbor over the entire image.

C: Color Feature:- Since color is an important feature that human perceiv ewhile visualizing it is considered for extraction from nucleu sregions. Hence for each nucleus image the mean color values in RGB color spaces are obtained. [13].

D: Fractal Dimension:- Fractals have been used in medicine and science previously for several quantitative measurement [14] [15]. the important measure that decided whether a particular nucleus represents a lymphoblast or a mature lymphocyte is the Perimeter roughness of nucleus. the more convenient way to parameterize the cell boundary surface in comparison to Euclidean geometry is fractal geometry. Hausdorff dimension (HD) is a main feature for fractal geometry and will be a main quantitative measure for cell boundary roughness measurement. a procedure for Hausdorff Dimension (HD) measurement using box counting method [16]. The Hausdorff Dimension *HD* may then be obtaned as in (6).

$$HD = \frac{Log(N)}{Log(N(s))} \qquad (6)$$

where, N the number of squares in the superimposed grid, and N(s) the number of occupied squares or boxes (box count). Higher HD denote to higher degree of roughness.

E: ***Classification****:-*

Classification is the task of specifing a label from one of the known classes to the unknown test vector,. Image classification analyzes the numerical properties of various image features and organizes data into categories. two phases of processing typically employ to Classification algorithms: training and testing.. In the initial training phase, the distinguishing properties of typical image features are isolated and based on these a unparalleled description of each classification category, training class is created. In the subsequent testing phase, these partitions of feature-space are used to classify the image features.

In this paper Backpropagation(BP), an abbreviation for "backward propagation of errors", is a One of the most popular method of training artificial neural networks used *in conjunction* with an optimization method such as gradient descent. The method computes the gradient of the loss function with respect to all weights in the network. The gradient is fed to the optimization method which in turn uses it to update the weights, in aim to minimize the loss function.

Backpropagation(BP) requires a known, desired output for the each input value in order to compute the loss function gradient. For this usually considered to be a supervised learning method, although it is as well used in some unsupervised networks such as auto encoders. It is a generalization of the delta rule to multi-layered feedforward networks, made possible by emploing the chain rule to iteratively calculate gradients to each layer. Backpropagation requires that the activation function used by the artificial neurons (or "nodes") be differentiable.

To reduce this training time effectively proper algorithm should be chosen so that tha system provide the system fastest intelligence. In this research we have used Scaled Conjugate Gradient(SCG).

III. EXPERIMENTAL RESULT

The proposed technique has been applied on 100 blood smear images obtained from “Oncology Center - Faculty of Medicine - Mansoura University Hospital - Egypt” is made up of 50 microscope image samples of blood infected , 50 microscope image of the blood samples is not infected.

A. Experiment

The experiment work of proposed system consist several steps, all image are preprocessed by MATLAB to contrast enhancement and resize be defined as in Table I. The system segments all dataset image (cancer and normal image) ,and we get separated nucleus sub Images using Bounding box technique to extract all features. The system classify uses SCG algorithm in neural network by use train dataset for all features.

B. Result Analysis

The experimental result has been developed by taking the entire test image. The entire test images are gone through the Preprocessing – Segmentation – Features Extraction and simulate training in neural network.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 5 Issue 01, January-2016

Table.1

Pre-processing (contrast – resize)	Time	
	Cancer 45 image	*Normal 45 image*
512*512	14.7313	14.778
256*256	13.412	13.7233
128*128	12.8222	12.2991
600*400	14.9824	145964
300*200	13.2879	13.1211
200*180	12.8286	12.8089

Table.2

Segmentation K-Means	Time	
	Cancer 45 image	***Normal 45 image***
512*512	54.4611	52.575
256*256	14.5826	14.9145
128*128	8.1589	6.2723
600*400	49.5682	48.9681
300*200	13.5127	13.5717
200*180	12.5690	12.3620

Table.3

Sub – Image	Time	
	Cancer 45 image	***Normal 45 image***
512*512	9.47817	10.1307
256*256	3.8499	3.67331
128*128	3.42657	2.94794
600*400	7.92125	7.48196
300*200	3.88147	3.99509
200*180	3.7890	3.6025

Table.4

Features Extraction	Time	
	Cancer 45 image	***Normal 45 image***
512*512	16.898	17.0685
256*256	5.21939	5.31412
128*128	2.42911	2.84163
600*400	29.3403	29.1306
300*200	8.20683	8.09558
200*180	3.6898	3.45067

Table.5

Size	Time
	Test (All Stages)
512*512	1.92556
256*256	0.546619
128*128	0.308521
600*400	2.09599
300*200	0.678705
200*180	0.4888692

Table.6

Size	Time
	Test (Train)
512*512	0.454771
256*256	0.446777
128*128	0.405953
600*400	0.524658
300*200	0.433325
200*180	0.418521

Table.7

Size	Time
	Test (Validation)
512*512	0.024298
256*256	0.0193219
128*128	0.0150367
600*400	0.024692
300*200	0.035152
200*180	0.017608

Table.8

Size	**Accuracy**
Old Result by SVM	95%
New result by pb-SCG nn with size image 128*128	99.74%

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 5 Issue 01, January-2016

IV. CONCLUSION

This paper has present segmentation (K-means Clustering) technique and features extraction (Shape – Texture – Color – HD) and BP – SCG neural network for classify. The system was evaluated in MATLAB 2014 and using data base 45 infected images and 45 non – infected images. The system is less computational requirement this make system well suited for low cost hardware implementation, the system achieved better Accuracy rate (99.74%) and training time.

REFERENCES

[1] K. Breden, T. Schorr, J. B. Schorr, "Blood" Colliers Encyclopaedia, Vol.4, 1978

[2] Children's Hospital of Wisconsin Website. http://www.chw.org

[3] Website http://www.mathworks.com/help/images/filter-images-using-predefined-filters.html

[4] Website http://www.mathworks.com/help/images/ref/imfilter.html

[5] K. S. Ravichandran and B. Ananthi. Color Skin Segmentation using K–Means Cluster. International Journal of Computational and Applied Mathematics, 4(2):153 – 157, 2009.

[6] S. Mohapatra and D. Patra, "Automated Cell Nucleus Segmentation and Acute Leukemia Detection in Blood Microscopic Images ," International Conference on Systems in Medicine and Biology. P(51). 16-18 December 2010, IIT Kharagpur, India

[7] R. C. Gonzalez and R. E. Woods. Digital Image Processing. Addison Wesley, 2nd edition, 1992.

[8] A. K. Jain, Fundamentals of Digital Image Processing. Pearson Education, 1st Indian edition, 2003.

[9] J.S. Lim "Two dimensional signal and image processing" Prentice Hall 1990

[10] R. C. Gonzalez, R. E. Woods, S.L. Eddins, "Digital Image Processing Using MATLAB", Pearson Prentice Hall Pearson Education, Inc., New Jersey, USA, 2004

[11] H. J. A. M. Heijmans (1994): Morphological Image Operators, Academic Press, New York.

[12] Fabio Scotti," Automatic Morphological Analysis for Acute Leukemia Identification in Peripheral Blood Microscope Images", IEEE International Conference on Computational Intelligence for Measurement Systems and Applications Giardini Naxos, Italy,P98. 20-22 July 2005.

[13] The Wikipedia the Free Encylopedia Website. http://en.wikipedia.org

[14] B. B. Mandelbrot. How long is the coast of Britain? Statistical self similarity and fractional dimension. Science, 156:636 – 638, 1967.

[15] B. T. Milne. Measuring the fractal geometry of landscapes. Applied Mathematics and Computation, 27:67 – 79, 1988.

[16] A. P. Pentland. Fractal based description of natural scenes . IEEE Transactions on Pattern Analysis and Machine Intelligence, 6:661 – 674,1984.