

Identification of Types of Leukemia Influenced Cells and it's Classification

Akshata M Mudakavi¹, Akshitha A R², Dolly Muthamma M R³, Eshwari H M⁴, Prof. Mangala C N⁵

¹Student, EWIT, India, akshata.m1195@gmail.com

²Student, EWIT, India, akshitharangerajan30@gmail.com

³Student, EWIT, India, dollymr2010@gmail.com

⁴Student, EWIT, India, nagashree12.96@gmail.com

⁵Professor EWIT, India, mangalacn@ewit.edu

ABSTRACT

Image processing and analysis paves the way to extract and quantify objects and patterns in image data. Microscopy images in biology are often found to be complex and noisy. The diagnosing of blood disorders plays a very crucial role in the timely detection and curing of disease in the medical field and thereby protecting and saving human life. We discuss in this paper, a proposed framework for segmenting the white blood cell using Otsu's global threshold method, and then by routing it to morphological minimum filter the darkest part of the image is identified as 'Nucleus'. Then the image is subjected to the process of feature extraction to obtain its geometric features information such as the size and shape of a nucleus, area, perimeter, eccentricity, circularity and solidity and the classification of cells into types and subtypes of leukemia are explained.

Keywords: Eccentricity; Global threshold ; Minimum filter ; Segmentation

1. INTRODUCTION

White blood cell, also termed as leukocyte or white corpuscle, is a cellular component of the blood that lacks hemoglobin, has a nucleus, is capable of motility, and immune the body against infection and disease by absorbing external materials and cellular debris, by terminating infectious agents and cancer cells, or by producing antibodies[1].

1) Normal Blood Cells:

The various normal blood cells are: Erythrocytes, Leucocytes, and Thrombocytes [2].

(i) Erythrocytes:- The erythrocytes are the most numerous blood. They are termed as red cells too.

(ii) Leucocytes:- Leukocytes, or white cells, are involved in the resistance of the organism. In

the blood, they are much less in quantity than red cells.

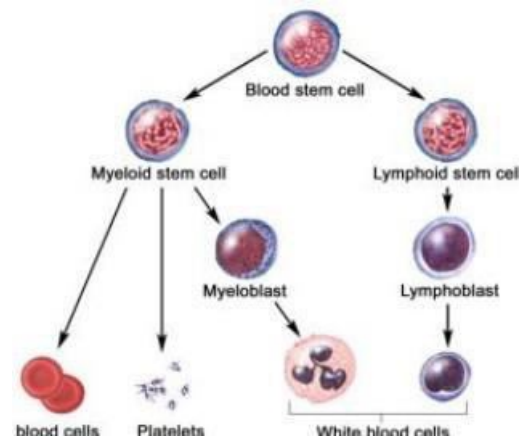


Fig 1: White Blood Cell Evolution[2]

(iii) Thrombocytes:- The main function of platelets, or Thrombocytes, is to impasse the loss of blood from wounds (hematostasis).

2) Leukemia Blood Cells:

The term "leukemia" refers to a group of cancers of the blood cells. In leukemia, white blood cells turn odd, split and multiply in an unrestrained way. Leukemia is clinically and pathologically categorized into a range of large groups. Leukemia is further typified as acute or chronic. In case of acute leukemia, undeveloped blood cells multiply exponentially affecting the generation of fit blood cells by the bone marrow. In chronic leukemia, there is an excessive build up of relatively mature, but abnormal, blood cells. Chronic leukemia is typically take months or years to evolve. Leukemia is further characterized based on the usual detection of abnormal blood cell. When leukemia infects lymphoid cells, it is termed as lymphocytic leukemia. When it affects myeloid cells, it is termed as myelogenous leukemia[3]. Leukemia is further categorized as

(i) Acute lymphocytic leukemia (ALL), which is also known as acute lymphoblastic leukemia. Children are usually affected by ALL and it is also found in adults aged 65+.

(ii) Acute myelogenous leukemia (AML) is more common in adults than in children.

(iii) Chronic lymphocytic leukemia (CLL) occurs most often in adults over the age of 55.

(iv) Chronic myelogenous leukemia (CML) affects grownups, also few children inflicted with it [4].

2. RELATED WORKS

Through reviewing the existing work in this regard and already published literature, it is observed that numerous researchers utilized diverse Segmentation strategies like Thresholding strategy, Region based methodologies, edge detection approach, Clustering approaches, Artificial Neural Network, Fuzzy procedure, watershed algorithm and so on for the segmentation of leukemia images and attempted to discover better result.

Tobias Bergen, Dirk Steckhan, Thomas Wittenberg and Thorsten Zerfab [5] proposed a method for the exact segmentation of leukocytes and erythrocytes pixel-wise classification with template matching to locate erythrocytes and used a level-set approach in order to get the exact cell contours of leukocyte nucleus and plasma regions as well as erythrocyte regions. The limitation here was separation of leukocyte clusters was not possible and hence could not detect abnormal leukocytes.

Zahra Khandan Khadem Alreza and Alireza Karimian [6] proposed a model for segmenting leukocytes groups using division of watershed conversion and image cleanup is done and all leukocytes available on the edge of images and abnormal components are removed. Zach algorithm is used for thresholding which does not eliminate the background objects. The segmentation of leukocyte clusters are not very accurate (60%).

Krishna Kumar Jha, Biplab Kanti Das and Himadri Sekhar Dutta [7] proposed two phase algorithm where phase 1 involves identification of different types of nucleus. Phase 2 involves counting abnormal nucleus from WBC. The proposed method uses morphological filter to reduce noise from input image and edge detection to highlight the feature of interest. It helps in easy analysis of actual problem in fast and effective manner. The edge detection has a problem with corner and junction. The corner pixel look for wrong direction for their neighbor, leaving open

ended edges and missing junctions. Thus reducing the accuracy in diagnosis. Its accuracy is only 65%.

N.H.Abd Halim, M.Y.Mashor, A.S.Abdul Nasir, N.R.Mokhtar and H.Rosline [8] proposed a Global Contrast Stretching (GCS) to enhance the area of interest and segmentation of nucleus is done using HSI (Hue, Saturation, Intensity) color space. It is fast and effective segmentation procedure which accelerates diagnosis of leukemia disease. The proposed method uses only Saturation scale which may contain wide range of lightness. It may progress from white to green which is combination of both saturation and lightness. This results in limitation in the proposed method, which decreases accuracy.

R.G Bagasjvara, Ika Candradewi, Sri Hartati and Agus Harjoko [9] proposed a technique for automatic detection and classification of Acute Leukemia where for segmenting White Blood Cells a dual-threshold method is used and this approach used two threshold values which obtained from golden section search algorithm, where two threshold values were applied in different color space: gray-scale image and HSV color space which gives the accurate count of cells. The limitation here is it cannot classify subtypes of acute cells and chronic cells.

3. METHODOLOGY

PROBLEM STATEMENT:

To identify the leukemia influenced cells in the given input blood smear image and classify them into types and subtypes of leukemia like Acute Lymphocytic Leukemia(ALL), Acute Myeloid Leukemia(AML), Chronic Lymphocytic Leukemia(CLL) or Chronic Myeloid Leukemia(CML).

a. IMAGE ACQUISITION

The blood cell images of humans are taken as the input in the RGB format (24 bits/pixel). The images may be of both acute and chronic type. The RGB image is converted to grayscale image using RGB to gray conversion process. The output is the grayscale image (8 bits/pixel). The input RGB image can be also converted into Red channel, Green channel and Blue channel image. For medical image processing, it is preferred to select green channel image for further processing.

b. IMAGE ENHANCEMENT

The image consists of invisible noises which degrades the quality of image. So we need to apply an image enhancement algorithm which increases the contrast of the image. This process is called as Image enhancement. The image is enhanced by using Contrast Limited Adaptive Histogram Equalization (CLAHE) for better visualization and to enhance the contrast of the image.

As a result of CLAHE, the dark area in the input image that was badly illuminated will become brighter in the output image while the side that was highly illuminated remains or reduces so that the whole illumination of the image is same.

A. Contrast Limited Adaptive Histogram Equalization (CLAHE)

1. Divide the image into sub-images.
2. For each sub-image, compute the histogram.
3. Apply point transformation function to pixel to form a local window.
4. Boundaries of images are removed by CDF.
5. Bulk of image is bi-linearly interpolated and contrast will be enhanced.

• IMAGE SEGMENTATION AND MORPHOLOGICAL OPERATIONS

Binarization is a process where each pixel in an image is converted into one bit and you assign the value as '1' or '0' depending upon the mean value of all the pixel. If greater than mean value then its '1' otherwise its '0'. Image Binarization converts an image of up to 256 gray levels to a black and white image. Frequently, Binarization is used to convert the grayscale image into binary image with the aim of making the WBC and platelets regions in white color and the background regions in black color. The WBC cells and platelets can be segmented and shown from the image. The simplest way to use Image Binarization is to choose a threshold value, and classify all pixels with values above this threshold as white, and all other pixels as black. For selecting the threshold, we use Otsu's global thresholding algorithm.

A. Otsu's Global Thresholding Method

The division of an image into meaningful structures, image segmentation, is often a step forward in image analysis, object representation, visualization, and a bunch of other image processing tasks [8]. Thresholding is the most commonly used image segmentation method, it is used to create binary images from a grayscale image. Otsu's thresholding method applies iterative process with all probable threshold values and computing the degree of spread for the pixel levels each side of the threshold, i.e. the pixels that either falls in foreground or background. The aim is to obtain the threshold value where the sum of foreground and background spreads is at its minimum [9].

1. Compute histogram and probabilities of each intensity level
2. Set up initial $\omega_i(0)$ and $\mu_i(0)$ and t
3. Step through all possible thresholds $t=1, \dots$ maximum intensity.

Update ω_i and μ_i .

4. Compute $\sigma^2_b(t)$.

5. Desired threshold corresponds to the maximum $\sigma^2_b(t)$.

After converting the image into binarized image, we can remove all the pixels smaller than WBC by using morphological operations. In the morphological processed image, we can apply morphological min filter to highlight the darkest nucleus in WBC.

B. Morphological Min filter

Morphological image processing is a technique used to introduce operations for transforming images in a unique way which takes into account the image content too. The two existing general morphological operations are minimum (often termed as dilation) and maximum (erosion) filters. The minimum filter facilitates the extension of object boundaries, while the maximum filter erodes shapes on the image. In morphological filters, each pixel is updated based on its comparison against surrounding pixels in the running window. The running window is defined as the image area formed around an existing pixel with defined radius. For example, if we specify the radius = 1, the running window will be a 3x3 square around the target pixel. The transformation results in the replacement of central pixel by the darkest one in the running window [10].

• FEATURE EXTRACTION

Feature extraction means to categorize the input data into different set of features. This paper defines 3 characteristics of lymphocyte cells, area, perimeter and circularity since shape of the nucleus is the prime factor for differentiation of blasts. After detecting each WBC area, the area can be calculated for each WBC nucleus. First we have to count number of pixel inside nucleus of WBC then the area of WBC nucleus can be calculated [11].

i. Area: The area was ascertained by totalling the number of none zero pixels within the image region.

ii. Perimeter: By calculating the distance between each adjoining pair of pixels around the border of the region.

iii. Circularity: This is a dimensionless parameter which changes with surface irregularities and is defined in equation (1):

$$\text{Circularity} = 4 * \pi * \text{Area} / \text{Perimeter}^2 \quad (1)$$

iv. Eccentricity: This parameter is used to measure how much a shape of a nucleus deviates from being circular [13]. It's an important feature since normal blood cells are more circular than the blast [14]. The value is between 0 and 1. An ellipse whose eccentricity is 0 is actually a circle. Eccentricity given by the equation (2):

$$\text{Eccentricity} = (\sqrt{a^2 - b^2}) / a \quad (2) \text{ Where } a \text{ Major axis, } b \text{ Minor axis.}$$

v. Solidity: Solidity is used to find out the density of a

component. Solidity is calculated by equation (4). If the solidity value is less than one then it is having irregular boundaries [14]. It is given in equation (3).

$$\text{Solidity} = \frac{\text{area}}{\text{convex area}} \quad (3)$$

Where convex area is returns a scalar that specifies the number of pixels in the hull filled in.

• CLASSIFICATION

Based on the feature values of the input image, we are going to classify that the given input image is of which type. It may be Acute Lymphocytic Leukemia(ALL), Acute Myeloid Leukemia(AML), Chronic Lymphocytic Leukemia(CLL) or Chronic Myeloid Leukemia(CML). For classification, we can use SVM(Support Vector Machine) Algorithm. SVM is a promising nonlinear, nonparametric classification technique, which already showed good results in the field of medical diagnostics. Further, the SVM is a potent algorithm with robust theoretical foundations based on the Vapnik–Chervonenkis theory and it possess sound generalization with regard to the new data. In the case of SVMs, a data point is considered as an m -dimensional vector, and the query is that whether such points can be parted away with a hyper plane. There are plenty of hyper planes classifying the data. The common purpose of SVM is to pinpoint the best hyper plane which represents the largest separation or margin between the two classes. In fact, SVM creates a decision surface in the feature space by diverse kernel functions; linear or nonlinear such as quadratic, polynomials, and radial basis functions (RBF) [15].

4. EXPERIMENTAL RESULTS

The proposed method is made functional by using the Image Processing Tools of MATLAB R 2013a. The associated figures provided in Fig 2 and Fig 3 represents to the outcomes acquired from the proposed structure calculation. The sample images are collected from hospital labs and also from the dataset. Initially, in the proposed method the gray scale image is taken as the input. To enhance the image the pre-processing step is performed. As the next step segmentation of the white blood cell is done using Otsu's global thresholding method. The darkest nucleus was highlighted using the morphological min filter. In feature extraction step the geometric feature of the resultant image is calculated. The Table I represents geometric feature extraction of the leukemia infected blood cell image. The important information is contained in the geometrical shapes and parameters associated with them. In the Table I the circularity value of all the three Nucleus1, 2 & 3 is less than 0.90, which indicates that the nucleus is not circular and also denotes the distortion in the shape of the nucleus. If eccentricity value is near to one indicates that the nucleus is a blast. Considering the solidity, if its value is below one then it seems to possess irregular boundaries which again substantiates that the input image contains a blast

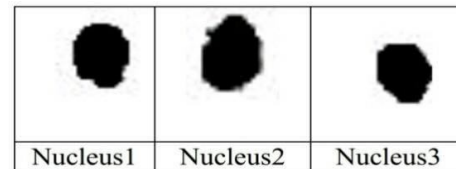


Fig 2: Separated Nucleus Sub-images

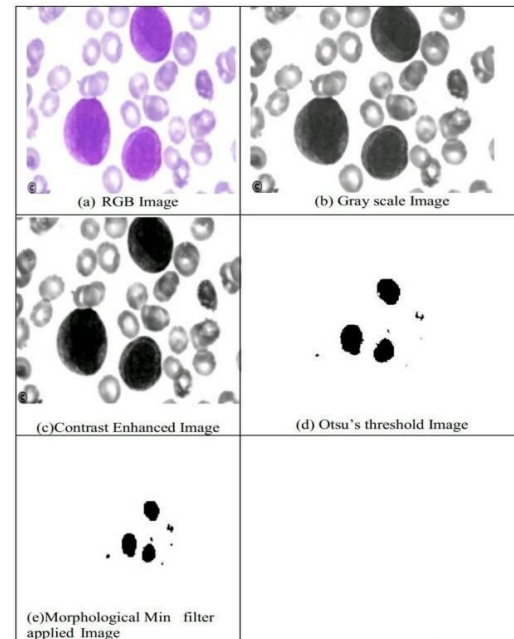


Fig 3: Classification Process

TABLE I THE RESULT OF GEOMETRICAL FEATURE EXTRACTION

Geometric Features	Nucleus 1	Nucleus 2	Nucleus3
Area	2093	1109	1571
Eccentricity	0.8095	0.6395	0.7906
Perimeter	206.3087	176.8944	198.5513
Solidity	0.9045	0.8698	0.7906
Circularity	0.6176	0.4451	0.5005

5. CONCLUSION

The prime objective of this proposed method is to diagnose the types of leukemia from microscopic images by analyzing its geometric features. Through this study it can be concluded that the shape of the nucleus is more important than the shape of cytoplasm for accurate detection of the leukemia blast. Hence classification is deep and accurate. In future, the arrived result can be subjected to classification process to identify different subtypes of types of acute and chronic Leukemia such as ALL is divided into L1, L2 and L3

6. ACKNOWLEDGEMENT

We wish to offer our sincere gratitude to our principal Dr. Prof .Channakeshvala for providing me an opportunity to be a part of the East West Institute of Technology, Bangalore under the graduate program and moreover to allow me work on this final year project on the respective domain.

We would like to offer gratitude to the H.O.D of Computer Science and Engg ,East West Institute of Technology ,Dr. Arun Biradar for his continuous support and encouragement towards our project.

We sincerely thank our guide Prof. Mangala .C.N , for the guidance and encouragement in carrying out this work and rectifying the mistakes in the same for the betterment of our work.

REFERENCES

[1].<https://www.britannica.com/science/white-blood-cell>.

[2]. International Journal of Electrical and Electronics Engineers ISSN- 2321-2055 (E) IJEEE, Vol: 7, Jan- June 2015.

[3]. <http://serc.carleton.edu/Woburn>.

[4].Tejashri G. Patil, V. B. Raskar, "Blood Microscopic Image Segmentation & Acute Leukemia Detection", International Journal of Emerging Research in Management & Technology,ISSN:2278-9359,Vol:4, September 2015.

[5].Tobias Bergen, Dirk Steckhan, Thomas Wittenberg and Thorsten Zerfab,"Segmentation of leukocytes and erythrocytes in blood smear images",ISSN:978- 1-4244-1815,February 2008.

[6]. Zahra Khandan Khadem Alreza and Alireza Karimian , "Design a new algorithm to count white blood cells for classification Leukemic Blood Image using machine vision system",ISSN:978-1-5090-3586,October 2016.

[7]. Krishna Kumar Jha, Biplab Kanti Das and Himadri Sekhar Dutta,"Detection of abnormal cells on the basis of nucleus shape and counting of WBC",November 2015.

[8]. N.H.Abd Halim, M.Y.Mashor, A.S.Abdul Nasir, N.R.Mokhtar and H.Rosline , "Nucleus segmentation technique for acute leukemia",ISSN:978-1-61284-413,2011.

[9]. R.G Bagasjvara,Ika Candradewi,Sri Hartati and Agus Harjoko , "Automated Detection and Classification Techniques of Acute Leukemia using Image Processing: A Review",ISSN:978-1-5090-4357,2016.

[10]. MuhammadSajjad,SirajKhan, Muhammad Shoaib and Hazrat Ali,Zahoor,"Computer Aided System for Leukocytes Classification and Segmentation in Blood Smear Images",ISSN:978-1-5090- 5300,2016.

[11]. Sos Agaian, Monica Madhukar, and Anthony T. Chronopoulos,"Automated Screening System for Acute Myelogenous Leukemia Detection in Blood Microscopic Images",ISSN:1932-8184,2014.

[12]. Shaikh Mohammed Bilal N and Sachin Deshpande,"Computer Aided Leukemia Detection using Digital Image Processing Techniques",ISSN:978-1-5090-3704,2017.

[13].<http://www.cs.uu.nl/docs/vakken/ibv/reader/chapter10.pdf>.

[14].<http://www.labbookpages.co.uk/software/imgProc/otsuThreshhold.html>.

[15].www.graphicsmill.com/docs/gm/minimum-maximum-median-Filters.htm

[16].EstiSuryani, Wiharto, and Nizomjon Polvonov,"Identification and Counting White Blood Cells and Red Blood Cells using Image Processing Case Study of Leukemia", International journal of Computer Science & Network Solutions, Vol:2.No.6, ISSN:2345-3397, June 2014.

[17]. Nimesh Patel, Ashutosh Mishra, "Automated Leukemia Detection using Microscopic Images. Second International Symposium on computer vision and Internet (VISIONNET'15), Procedia Computer Science 58, pp:635 642, 2015.

[18].<http://in.mathworks.com/help/images/ref/regionprops.html>.

[19].<http://www.hematologyatlas.com/leukemia.html>.

20].Nagabhushana, R.M, Rajeshwari\ P,"Geometrical Analysis of Leukocyte Nucleus to Detect Lymphoblast from Microscopic Blood Images" Vol:2[6], pp: 1947-1951, June 2014.

AUTHORS BIOGRAPHY

[1] Ms. Akshata M Mudakavi is pursuing her 8th semester B.E in Computer Science & Engineering at East West Institute of Technology, Bangalore, India. Her area of interest includes Image Processing and machine Learning.

[2] Ms. Akshitha A R is pursuing her 8th semester B.E in Computer Science & Engineering at East West Institute of Technology, Bangalore, India. Her area of interest includes Image Processing and machine Learning.

[3] Ms. Dolly Muthamma M R is pursuing her 8th semester B.E in Computer Science & Engineering at East West Institute of Technology, Bangalore, India. Her area of interest includes Image Processing and machine Learning.

[4] Ms. Eshwari H M is pursuing her 8th semester B.E in Computer Science & Engineering at East West Institute of Technology, Bangalore, India. Her area of interest includes Image Processing and machine Learning.

[5] Mrs. Mangala C N received the B.E degree in Computer Science and Engineering from NCET, Bangalore, VTU in 2006 and got M.Tech degree in Computer Science from RVCE, Bangalore, India. She is currently working as Associate Professor in the Department of CSE, EWIT, and pursuing PhD in DSCE, Bangalore, India. Her area of interest includes Image Processing, Data Mining and Big Data.