Segmentation of Acute Lymphoblastic Leukemia

Piyamas Suapang^{a,*}, Kathtaliya Pumjaroen^b, Prasong Tosranon^c

^a Program in Medical Instruments and Operating Room Technology, Department of Health Technology, Faculty of Sciences and Health Technology, Navamindradhiraj University, Bangkok 10300, Thailand

^{b,c} Department of Industrial Physics and Medical Instrumentation, King Mongkut's University of Technology North Bangkok, 10800, Thailand

*Corresponding Author: piyamas.sua@nmu.ac.th

Abstract

The various steps involved in acute lymphoblastic leukemia detection are pre-processing, segmentation of white blood cells, feature extraction and then classification. The segmentation procedure, a novel simple algorithm, is proposed for localization of white blood cells and the different cell components are separated with automatic thresholding. Features extracted from the segmented nucleus are motivated by the visual cues of shape, color and texture. The results shown that signification feature to classified normal and acute lymphoblastic leukemia cell, such as area, perimeter, orientation, eccentricity, mean and standard deviation in RGB color space and L*a*b color space, number of vacuole, number of nucleoli, nucleus compactness, number of nucleus pixels to cytoplasm pixels ratio. The backpropagation neural network for implementation and uses the different combinations of feature sets. The results presented here are based on trials conducted with normal cells. For training the classifiers, a library set of 532 patterns is used. The tested data consists of 477 samples and produced correct classification rate in the range from 74.24% to 96.98 %, 88.48% for an average sensitivity and 78.58% for an average specificity.

Keywords: Acute Lymphoblastic Leukemia, Segmentation, Backpropagation Neural Network.

1. Introduction

Leukemia is a disease that affects blood forming cells in the body. Early detection of the disease is necessary for proper treatment management. Leukemia is a cancer of the

bone marrow and blood and is classified into four main groups according to cell type and rate of growth: acute lymphocytic (ALL), chronic lymphocytic (CLL), acute myeloid (AML), and chronic myeloid (CML). Symptoms may include fatigue, paleness, weight loss, repeated infections, fever, bruising easily, and nosebleeds or other hemorrhages. In acute leukemia, these signs can appear suddenly. Chronic leukemia typically progresses slowly with few symptoms and is often diagnosed during routine blood tests. Patients with CLL may experience swollen lymph nodes or pain in the upper left abdomen due to an enlarged spleen. Leukemia can be difficult to diagnose early because symptoms often resemble those of other, less serious conditions. When a physician does suspect leukemia, diagnosis can be made using blood tests and a bone marrow biopsy.

Acute Lymphocytic Leukemia (ALL), also known as acute lymphoblastic leukemia is a cancer of the white blood cells, characterized by the overproduction and continuous multiplication of malignant and immature white blood cells (referred to as lymphoblast or blasts) in the bone marrow. It is fatal if left untreated due to its rapid spread into the bloodstream and other vital organs. In order to classify the abnormal cells in their particular types and subtype of leukemia, an expert operator will observed some cells under a light microscopy looking for the abnormalities presented in the nucleus or cytoplasm of the cells. This classification is very important to determine which treatment should be given to the patient⁽¹⁾. However, this analysis suffers from slowness and it presents a not standardized accuracy since it depends on the operator's capabilities and tiredness. Regardless of advanced techniques i.e. flow cytometer, immune phenotyping, molecular probing etc, microscopic

examination of blood slides still remains as a standard leukemia diagnosis technique. Hence microscopic examination is the most economical way for initial screening of leukemia patients. Manual examination of the slides are subjected to bias i.e. operator experience, tiredness etc resulting with inconsistent and subjective reports. So there is always a need for a cost effective and robust automated system for leukemia screening which can greatly improve the output without being influenced by operator fatigue.

Fig. 1 Examples of five major groups of human leukocytes in peripheral blood: (a) Healthy lymphocytes cell from non-ALL patients. (b–d) Lymphoblasts from ALL patients where (b), (c), and (d) are L1, L2, and L3, respectively.

Leukemia is a disease that affects blood forming cells in the body. Early detection of the disease is necessary for proper treatment management. Abnormal white blood cells or blasts play important role for hematologists in their diagnostic process. The diagnosis and treatment of diseases have been significantly simplified by the differential counting of white blood cells as it furnishes critical information required by pathologists. The practice of manual counting of the white blood cells suffer from the obvious disadvantages associated with human errors. Naturally an automated process would be the ideal solution to such a problem. Digital image processing technique could help them by enhancing the visibility in their analysis and diagnosis. White blood cell segmentations are an important research issue in Hematology. Our research proposes a segmentation of nucleus and cytoplasm of white blood cell slides for acute lymphoblastic leukemia detection.

2. Methodology

2.1 Contrast Enhancement

During image acquisition and excessive staining, the images will be disturbed by noise. The noise may be due to illumination or shadows that make region of interest (ROI)

appear as blurred image region. Background will be excluded since ROI will be white blood cells. Image enhancement was done as the contrast enhancement technique⁽²⁾.

2.2 Segmentation

First, the captured image file was split into its three component bands (red green and blue as shown in figure 2). The result was three grayscale files one for each of the red green and blue components of the image captured by the camera. Histogram analysis was used to examine three grayscale components (corresponding to the red, green and blue bands) of 30 images covering all five basic white blood cell types. It was found that the green component was consistently a better discriminator between the purple nuclear material and the rest of the image.

Fig. 2 (a) Raw captured image and (b, c, d) 3 single color components.

We did utilize thresholding to produce a binary bitmap image from the green band bitmap for each image. An automatic threshold value was selected by minimum value to discriminate between nuclear and nonnuclear pixels, as every white blood cell has a nucleus. We used only the green band for this nuclear thresholding, as, in this band, the nuclear material is much darker than either the cytoplasm (with the exception of basophilic cytoplasm) or the background. Experimentation showed that an automatic threshold value of 95 (on a scale of 0-256) gave an acceptable discrimination between nuclear and non-nuclear pixels, leaving nuclear pixels black on a white background as shown in figure 3.

Due to edge effects on the captured images (the camera produced noticeable dark bands at the very edges of the images, as seen in Figure 2 and 3), there were many nonnuclear pixels around the perimeter of the image. An edge erosion filter, which simply set all of the pixels within 3 pixels of the edge of the image to white, was developed and used. This removed the dark bands around the edge of the image as shown in figure 4.

Fig. 3 Green component image and thresholded bitmap.

As a first step towards identifying a region of interest, an algorithm was developed to identify blobs (continuous connected groups of black – presumably nuclear – pixels) within a bitmap file and print out their details, including number of pixels in the blob and centroid (arithmetic mean of the x and y position values) of the blob. The x value of the centroid was found by summing the x values of all pixels in a blob and dividing the sum by the number of points in the blob. The y value of the centroid was found in a similar fashion, summing the y values of all pixels in the blob and dividing by the number of pixels. The (x, y) location represented by this centroid was used as the center of the blob of nuclear material. The center (c_x , c_y) and number of points were recorded for each blob found.

A Region of interest (ROI) was defined around the center (c_x , c_y) of each blob. This region of interest was defined to be a rectangle 110 by 110 pixels in size. The size of the ROI was originally chosen to preserve the aspect ratio of the camera's image, be large enough to accommodate the largest leukocyte with some headroom. The rectangular aspect ratio was later found to be unnecessary, a square ROI would have been simpler. This process produced a series of smaller images like the one in figure 5.

Most white blood cells are roughly circular in shape,

though some (monocytes, in particular) may deviate significantly from the circular). For the purposes of obtaining a good feature set, representative of the features of the cell only, we chose to try to find the largest circular area entirely within the cell. This process produced finding circle as shown in figure 6.

Fig. 4 The results of (a) thresholding and (b) erosion.

Fig. 7 Extracted image of white blood cell.

Fig. 6 Finding circle.

2.3 Overview of System for Acute Lymphoblastic Leukemia Detection

The segmented circular cell region was processed on a pixel by pixel basis and statistical information (mean, standard deviation, maximum value and minimum value). The signification feature classified normal and acute lymphoblastic leukemia cell, such as area, perimeter, orientation, eccentricity, mean and standard deviation in RGB color space and L*a*b color space, number of

vacuole, number of nucleoli, nucleus compactness, number of nucleus pixels to cytoplasm pixels ratio.

The backpropagation neural networks are processing structures “consisting of many interconnecting processing elements (neurons).” These artificial neurons are connected together to form neural networks. An extremely simple example of such a network is shown in Figure 7. In this example, a number of inputs are each connected to each of a number of neurons in an intermediate layer. The neurons in the hidden layer are each connected to all the output neurons (one in this case). This is an example of a fully connected feed forward neural network. Feed forward networks of this type can be trained by back propagation. This is a procedure that trains the network by making small adjustments to the weights of each neuron in the direction that reduces the error at that neuron’s output. The input layer used white blood cell feature 15 features. The hidden layer was designed by 5 nodes. Finally, the results of output layer were equal the number of white blood cells in different classes and determined from the probability of class membership. The whole process can be schematized for training process and testing process.

Fig. 7 Flow Chart for Acute Lymphoblastic Leukemia Detection.

3. Results and Discussion

Fig. 8 shows that the discrimination between nuclear and nonnuclear pixels is selected after an automatic threshold value. The test applied these leukocytes features are carried out and the results are shown in Table 1. The results presented here are based on trials conducted with normal cells. For training the classifiers, a library set of 532 patterns is used. The tested data consists of 477 samples and

produced correct classification rate in the range from 74.24% to 96.98 %.

Fig. 8 The nucleus segmentation.

Table 1 The results of testing by leukocytes features set.

Classification	No.	True Positive	False Positive	False Negative	Sensitivity (%)	Positive Predictivity (%)
Acute Non-Lymphocytic Leukemia	440	330	8	110	97.63	75.00
Acute Lymphoblastic Leukemia (L1)	66	49	9	17	84.48	74.24
Acute Lymphoblastic Leukemia (L2)	26	25	5	1	83.33	96.15
Classification	No.	True Negative	False Negative	False Positive	Sensitivity (%)	Negative Predictivity (%)
Lymphocyte	41	38	12	3	76.00	92.68
Eosinophil	8	6	2	2	75.00	75.00
Neutrophil	30	28	5	2	84.85	93.33
Others	398	386	106	12	78.46	96.98

4. Conclusions

The segmentation procedure is proposed for localization of white blood cells and the different cell components are separated with automatic thresholding.

Features extracted from the segmented nucleus are motivated by the visual cues of shape, color and texture. The results shown that signification feature to classified normal and acute lymphoblastic leukemia cell, such as area, perimeter, orientation, eccentricity, mean and standard deviation in RGB color space and L*a*b color space, number of vacuole, number of nucleoli, nucleus compactness, number of nucleus pixels to cytoplasm pixels ratio. The backpropagation neural network for implementation and uses the different combinations of feature sets. The results presented here are based on trials conducted with normal cells. For training the classifiers, a library set of 532 patterns is used. The tested data consists of 477 samples and produced correct classification rate in the range from 74.24% to 96.98 %, 88.48% for an average sensitivity and 78.58% for an average specificity. Classification error rate increased by the unclearly image. Therefore, the image acquisition process was considered significant for the accuracy of the classification.

Acknowledgment

This work is partially supported by Program in Medical Instruments and Operating Room Technology (MIORT), Department of Health Technology, Faculty of Sciences and Health Technology, Navamindradhiraj University and Department of Industry Physics and Medical Instrument at King Mongkut's University of Technology North Bangkok. The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

References

- (1) Fabio Scotti, "Automatic Morphological Analysis for Acute Leukemia Identification in Peripheral Blood Microscope Images", CIMSIA 2005 IEEE International Conference on Computational Intelligence for Measurement Systems and Applications Giardini Naxos, Italy, 20-22 July 2005
- (2) Fauziah Kasmin, Anton Satria Prabuwono, Azizi Abdullah, "Detection of leukemia in human blood sample based on microscopic images: a study", Journal of Theoretical and Applied Information Technology, 31st December 2012. Vol. 46 No.2.
- (3) Fauziah Kasmin, Anton Satria Prabuwono, Azizi Abdullah, "Detection of leukemia in human blood sample based on microscopic images: a study", Journal of Theoretical and Applied Information Technology, 31st December 2012. Vol. 46 No.2.
- (4) N. M. Salem, "Segmentation of white blood cells from microscopic images using K-means clustering," in *Proceedings of the 31st National Radio Science Conference (NRSC '14)*, pp. 371–376, Cairo, Egypt, April 2014.
- (5) P. K. Mondal, U. K. Prodhan, M. S. Al Mamun et al., "Segmentation of white blood cells using fuzzy C means segmentation algorithm," *IOSR Journal of Computer Engineering*, vol. 1, no. 16, pp. 1–5, 2014.
- (6) S. Mohapatra and D. Patra, "Automated Cell Nucleus Segmentation and Acute Leukemia Detection in Blood Microscopic Images," in *International Conference on Systems in Medicine and Biology*, India, 2010.
- (7) N. H. A. Halim, M. Y. Mashor, A. S. Abdul Nasir, N. R. Mokhtar, and H. Rosline. 2011. Nucleus segmentation technique for acute leukemia, in *Proceedings of the IEEE 7th International Colloquium on Signal Processing and Its Applications (CSPA '11)*, pp. 192–197, March 2011.
- (8) Donida Labati, R., Piuri, V., Scotti, F. ALL-IDB: the Acute Lymphoblastic Leukemia Image DataBase for image processing, 2011.
- (9) S. Mohapatra and D. Patra, "Automated leukemia detection using hausdorff dimension in blood microscopic images," in *Proceedings of the International Conference on IEEE Robotics and Communication Technologies (INTERACT '10)*, pp. 64–68, Chennai, India, December 2010.
- (10) B. C. Ko, J.-W. Gim, and J.-Y. Nam, "Automatic white blood cell segmentation using stepwise merging rules and gradient vector flow snake," *Micron*, vol. 42, no. 7, pp. 695–705, 2011.