Volume 5, Issue 3, March 2015 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Detection of Acute Myelogenous Leukemia in Blood Microscopic Images

Liya Varghese, Jucy Vareed
Cse & Calicut, Kerala,
India

Abstract— Acute myelogenous leukemia (AML), also known as Acute myeloid leukemia or acute nonlymphocytic leukemia (ANLL), is a cancer of the myeloid line of blood cells, characterized by the quick growth of abnormal white blood cells that pile in the bone marrow and interfere with the production of normal blood cells. In this paper, a simple technique that automatically detects and segments AML in blood smears is presented.

Keywords— AML detection, Classification, Feature Extraction, Segmentation.

I. INTRODUCTION

Acute myelogenous leukemia (AML) or Acute Myeloid Leukemia is a fast growing cancer in blood and bone marrow. In AML, the bone marrow makes many cancerous cells called leukemic blasts. Normal blasts develop into white blood cells that fight infection. In AML, the leukemic blasts do not develop properly and cannot fight infections. These leukemic blasts grow quickly and piles out the bone marrow, preventing it from making the normal red blood cells, white blood cells, and platelets. AML can affect people of any age, but it is most common in adults.

The word "acute" in acute myelogenous leukemia denotes the disease's quick progression. It's called myelogenous leukemia because it affects a group of white blood cells called the myeloid cells, which normally develop into the various types of mature blood cells, such as red blood cells, white blood cells and platelets. Without treatment AML can quickly be fatal. Since it is Acute, this type of leukemia can spread quickly to the blood and to other parts of the body. Here comes the importance of automatic detection for AML. It helps to diagnose the presence of AML in blood sample which helps for advanced testing or further treatments.

II. OVERVIEW

Most of the AML detection involves segmentation, feature extraction and classification. First step involves image enhancement and segmentation where the pre-processed images ensure perceptual uniformity. Also important features are extracted for further processes. The next step which is an important one, is feature extraction. It is a technique of redefining a large set of redundant data into a set of features of reduced dimensions. It greatly influences the classifier performance. Based on the features selected the classifier segment classifies the images to different classes.

III. DIFFERENT TYPES OF AML DETECTION SYSTEMS

Fabio Scotti et .al in [1], the first module is a Single-cell Selector module which enhances the input image and identifies the single cells. Secondly, the White-cells Identifier module selects the white cells present into the image by separating them from others blood's components. The third module, the Lymphocyte Identifier which recognizes a lymphocyte with respect to the other selected white cells. Then the next sub-system has to recognize if a lymphocyte is blast or normal. The Feature Extraction module processes a sub-image containing a lymphocyte coming from the Lymphocyte Identifier module and it produces in output a set of morphological indexes. The classification module processes those indexes in order to classify the cell as blast or normal.

N. Sinha et .al in [2], a method for segmentation of cells from color images of blood smears is introduced. It involves locating each of the white blood cells and identifying the regions corresponding to nucleus and the cytoplasm, in the given blood smear. The input will be a RGB image which will be converted to its HSV equivalent. Each pixel is treated as a vector of the three dimensions namely Hue, Saturation, and Value. K means clustering is performed on the 3-D feature vectors. The centroids and the variances obtained in the K-means step are used to initialize Gaussian parameters for Expectation Maximization (EM) algorithm. The most important feature of this technique is that there are no parameters to be tuned by the user.

C.Reta et .al in [3], a two-step method to analyze the morphology of abnormal leukocytes images for the classification of acute leukemia subtypes. In the first step a segmentation algorithm is proposed which uses color and texture information in order to extract leukocytes and their respective nucleus and cytoplasm from bone marrow images. In the second step feature extraction is performed to the regions segmented and uses these attributes to classify the cells into leukemia subtypes. With respect to the cell classification, the use of descriptive features of nucleus and cytoplasm of the cells improved their representation.

Piuri et .al in [4], paper focuses on the problem of identification and classification of white blood cells by microscope images. This paper presents a methodology to achieve a fully automated detection and classification of leucocytes by microscope color images identifying the following classes: Basophil, Eosinophil, Lymphocyte, Monocyte and Neutrophil. The proposed system firstly individuates in the blood image the leucocytes from the others blood cells. secondly it extracts morphological indexes and finally it classifies the leucocytes by a neural classifier.

D. Foran et .al in [5], a distributed, clinical decision support prototype for classifying among hematologic malignancies is designed. The prototype developed consists of a distributed telemicroscopy system operating in concert with an intelligent image database. The system enables consulting physicians and scientists to engage in interactive telemicroscopy sessions. Then automatically search through databases of consensus-graded medical cases based upon the visual content of constituent pathology image records in order to obtain reliable decision support in detecting and discriminating among hematologic malignancies and for computer-assisted diagnosis.

N. Sinha et .al in [6], there distinguishes between the five classes of mature WBC's, namely lymphocyte, monocyte, eosinophil, basophil and neutrophil and automates the process of differential blood count. The input to the system is a digital image of blood smears and the output is the differential count of the cells. The main stages of the proposed system are: (i) Acquisition (ii) Segmentation (iii) Feature Extraction and (iv) Classification. WBC segmentation is a two-step process carried out on the HSV-equivalent of the image, using K-Means clustering followed by EM-algorithm. Features extracted from the segmented cytoplasm and nucleus, are motivated by the visual cues of shape, color and texture. Various classifiers have been explored on different combinations of feature sets.

Niranjan Chatap et .al in [7], the main phases system goes through are Image Segmentation, Image preprocessing, Image feature Extraction and Image Classification. The features analyzed here are color, texture, geometry and statistical analysis. The classification technique used here are Nearest Neighbor and Support Vector Machine.

Bakht Azam et .al in [8], the RGB image which is the input image will be smoothened first for making it clearer. Then image quantization will be done for reducing color levels. Then the image undergoes Binarization process which resembles gray scaling an image. Further it undergoes some morphological operations. The operations involve Morphological Filling and Morphological opening. The last step of segmentation is Labeling and Counting. Since the count of leucocyte plays a vital role in diagnosing the hematological diseases.

Nagabhushana R M et .al in [9], the main steps involved are Leukocyte Nucleus Segmentation, Feature Extraction and Classification. The main theme of the paper is WBC nucleus segmentation of stained blood smear images followed by relevant Feature Extraction for leukemia detection. The classification method used here is SVM.

Ms. Minal D. Joshi1 et .al in [10], the system follows steps such as White Blood Cell Identification and segmentation, Feature Extraction and Classification. The features extracted are Area, Perimeter and Circularity. The classification method used in this system is kNN Classifier.

IV. METHODOLOGY OF AUTOMATIC AML DETECTION

A. Preprocessing and Segmentation

Preprocessing is done for improving the quality of input images. Thus it will be more suitable for the application which is concerned of. This preprocessed image will be served as input image for the following modules. The method used for this can vary accordingly. From the preprocessed image, segmentation can be done. The goal of image segmentation is to extract important information from an input image. This will enhance the clarity for further processes. Segmentation process also depends upon the underlying application.

Fabio Scotti et .al in [1], there involves a Single-cell Selector module which enhances the input image and identifies the single cells. It has been composed by adaptive pre-filtering and segmentation algorithms. Then there is White-cell Identifier which selects the white cells present in the image by separating them from other blood's components. Lymphocyte Identifier recognizes a lymphocyte with respect to the other selected white cells.

N. Sinha et .al in [2], presents a method for segmentation of cells from color images of blood smears. It involves locating each of the white blood cells and identifying the regions corresponding to nucleus and the cytoplasm, in the given blood smear. The input will be a RGB image which will be converted to its HSV equivalent. Each pixel is treated as a vector of the three dimensions namely Hue, Saturation, and Value. K means clustering is performed on the 3-D feature vectors. The centroids and the variances obtained in the K-means step are used to initialize Gaussian parameters for Expectation Maximization (EM) algorithm. The EM algorithm iterates between segmentation and parameter estimation till convergence.

C.Reta et .al in [3], firstly a segmentation algorithm is proposed which uses color and texture information to extract leukocytes and their respective nucleus and cytoplasm from bone marrow images with heterogeneous staining.

Piuri et .al in [4], enhancement of image and identification of white blood cells are being done initially. The identification of leukocytes in the blood image is based on an adaptive pre-filtering and segmentation. The pre-filter helps to enhance only the leukocytes. It separates the others blood components with respect to the gray level intensities in order to achieve a very accurate leukocyte segmentation. After segmentation center of the sub-images will be located in the centroid position of the segmented nuclei. The dimension of the squared sub-image has been fixed to 3D.

D. Foran et .al in [5], the segmentation of images was accomplished by mapping red, green, and blue intensity values of the imaged specimens into color space and utilizing the fast nonparametric clustering method.

N. Sinha et .al in [6], the input to the system is a digital image of blood smears of healthy subjects and the output is the differential count of the cells. WBC segmentation is a two-step process carried out on the HSV equivalent of the image, using K-Means clustering followed by EM-algorithm.

Niranjan Chatap et .al in [7], image preprocessing involves enhancing image quality and deletes the overlapped blood cells. The segmentation part involves series of steps. The algorithm requires priori information about blood smear. It yields better results.

Bakht Azam et .al in [8], the input RGB image will be forwarded to Image Smoothing. Smoothing is done for making the input image clearer. Image Quantization will reduce color levels. Binarization is the next step after Image Quantization. Then morphological operations such as Morphological Filling and Opening will be done. And finally Labeling and Counting will be done which is the last step in segmentation.

Nagabhushana R M et .al in [9], Leukocyte Segmentation involves series steps. Initially the color image is converted to gray image. Then Histogram Equilization is the process followed. After that Contrast Stretching will be done. Arithmetic operations will be done on the processed image. By thresholding the resultant image will be converted to binary image. Finally Morphological Operations is done which is the last step in segmentation.

Ms. Minal D. Joshi1 et .al in [10], the Image Segmentation step involves a series of steps. Initially color image will be converted to gray scale image. Then enhancement will be done by applying Histogram Equilization process. After that linear contrast stretching will be done. Then the system undergoes some filtering processes and applies global thresholding for converting the image into a binary one. Finally morphological operations will be done such as morphological opening.

B. Features Extraction

Feature extraction is a crucial step in image processing which redefines a large set of redundant data into a set of features of reduced dimension. The process of transforming an input data into set features of reduced dimension is known as feature extraction. It greatly influences the classifier performance.

Fabio Scotti et .al in [1], the input will be a sub-image containing a lymphocyte coming from the Lymphocyte Identifier module and it produces in output a set of morphological indexes. Six sub-images of lymphocyte will be produced from the previous step. Three images containing portions of the original gray-level image, three black and white sub-images which represent only the borders of the membrane, the cytoplasm and the nucleus. The first three sub-images can be used to extract features regarding the gray-level intensity pattern of the image. From the three black- and-white sub-images we can easily measure morphological features such as the perimeter, the area, the momentums of the image, etc.

C.Reta et .al in [3], identification of leukocytes is done by the recognition of their nucleus and cytoplasm. From the regions obtained in segmentation processes the shape color and spatial relationship with respect to other regions are analyzed for checking nucleus or leukocyte. The features used for recognizing cellular elements are circularity, color and eccentricity.

Piuri et .al in [4], from the segmented images most important indexes need to be extracted from sub-images and their biological relevance in the leukocytes classification. The features extracted are Area, Perimeter, Convex Area, Solidity, Major Axis Length, Orientation, Filled Area, Eccentricity. In addition, ratio between the cell and nucleus areas, the nucleus' "rectangularity", the cell "circularity, the number of lobes, and also the solidity, area and mean gray-level intensity of the cytoplasm is processed.

D. Foran et .al in [5], features extracted are Multi-resolution Texture, Area, Color, Overall Similarity Measure and Retrieval Performance.

N. Sinha et .al in [6], features for discriminating between different cell classes are arranged based on domain knowledge of human experts. The features considered are based on Shape, Color, and Texture. Binary masks of nucleus and cytoplasm is used to compute these features. The advantage of using this method is that the contour information is not lost as would have been if one used morphological operations of open-close, for de-clustering. Color features are extracted from the segmented nucleus and cytoplasm. The average value of each of the color components of the nucleus and those of the cytoplasm are calculated. And texture features are computed only for the cytoplasm. The texture of cytoplasm is visually different across various classes, but this is not in the case of nucleus. The texture features used are based on computations of Gray-Level co- occurrence matrix, (GLCM) and Autocorrelation matrix. The features based on GLCM are energy, entropy and correlation. The features based on autocorrelation matrix are coarseness and busyness.

Niranjan Chatap et .al in [7], the features to be extracted are Area, Perimeter and Circularity.

Nagabhushana R M et .al in [9], the features to be extracted are Area, Perimeter, Compactness, Solidity, Eccentricity and Elongation. The quantitative evaluation of each nucleus is done using extracted features.

Ms. Minal D. Joshi1 et .al in [10], three set of features are to be extracted. They are Area, Perimeter and circularity of the segmented image.

C. Classification

Fabio Scotti et .al in [1], classification module processes indexes in order to classify the cell as blast or normal. If the system finds a blast cell, the blast counter is increased; otherwise a new lymphocyte will be processed. Classification is done by three types of classifier methods such as KNN classifier, Linear and Feed-Forward Neural Network.

C.Reta et .al in [3], different types of classification methods are being used to distinguish between ALL and subtypes of ALL. The AML subtypes behavior is analyzed by performing binary and multi class classification. The classification was carried out using instance based classifiers, decision trees, regression functions as well as meta classifiers such as k Nearest neighbors, Random Forest, Simple Logistic, SMO and Random Committee. The evaluation of the classification model is done using 10 cross validations.

Piuri et .al in [4], the capability of selected features in separating the 5 classes can be qualitatively evaluated by plotting the classes with respect to the three most relevant features such as cell area, nucleus area and gray intensity of the

cytoplasm. The most relevant features have been found by applying a feature selection technique called forward selection based on nearest neighbor classifier evaluated with Leave One Out method.

D. Foran et .al in [5], the performance of the system is compared with three classes of individuals such as those in training, those who perform screening of specimens for detecting abnormalities, and those responsible for rendering the differential diagnosis. Each of the participants in the study will be shown one digitized specimen at a time on a high resolution screen with no other distractor displayed. It will be asked to classify the cell as belonging to one of five classes as mantle cell lymphoma, follicular cell lymphoma, chronic lymphocytic leukemia, normal or other.

N. Sinha et .al in [6], according to classification rule of thumb, the number of training patterns for each class must be 5 to 10 times the dimensionality of the feature vector. But due to unavailability of sufficient data, classification is done with smaller training set. The training data consists of 50 samples and test data consists of 34 samples with fair representation from each class.

Niranjan Chatap et .al in [7], the classifier module classifies the lymphocyte cells as blast or normal cells. The kNN decision rule based classification is being done in this paper. It is very simple and an effective method of classification.

Nagabhushana R M et .al in [9], based on the extracted features classification is done by classifier as blast or normal by analyzing the lymphocyte cells. Support Vector Machine is applied for classification.

Ms. Minal D. Joshi1 et .al in [10], according to the extracted features the classifier classifies the lymphocyte cells among blast or normal cells. The kNN decision tool is used as the classifier here. The value for k is given as one in this system.

IV. CONCLUSIONS

Most of the automatic detection systems follow same processes. The advantage of this automatic system is the accuracy of detecting AML from blood sample. Fabio Scotti et .al in [1], the system presents a methodology which is achievable and it offers remarkable classification accuracy. Whereas N. Sinha et .al in [2], an efficient automatic system for blood cell segmentation from color images of blood smears are developed. The advantage of this system is it requires no user interaction or parameter tuning. The system can be easily adapted for any given data set with a known magnification. The performance is good even in cases where the nucleus is multi lobed as in neutrophils. C.Reta et .al in [3], the method used can be applied to images that show heterogeneous color and texture staining and high cell population which is a desirable property when working with bone marrow smears. Experimental results show that this method achieves a segmentation accuracy of 95% when it is compared with a manual segmentation performed by an expert. Piuri et .al in [4], a methodology that achieves a fully automated detection and classification of leucocytes by microscope color images identifying the classes like Basophil, Eosinophil, Lymphocyte, Monocyte and Neutrophil. Experiments show that the final classification module implemented by means of a parallel classifier composed by feed-forward neural classifiers achieves an accurate solution with minor computational complexity than traditional nearest neighbor classifier. Results indicate that the morphological analysis of blood‘s white cells is achievable and it offers remarkable classification accuracy. Whereas D. Foran et .al in [5] focuses on remote access and the associated image transfer technology. This system offers a fully automated indexing and database management interface for recording salient visual and clinical information and provides clinical decision support. With the addition of temporal data, the system might be used as a tool for staging, disease management, and clinical outcome studies. The proposed system described by N. Sinha et .al in [6] is developing an automatic blood counting system from the blood smear. The main attraction of this system is it does not require any user interaction or parameter tuning. A prior knowledge of number of classes is essential. The feature extraction phase is very simple and no complex computation is required. And it offers better performance. The method defined by Niranjan Chatap et .al in [7] involves, detecting the types of WBC’s and RBC’s from the blood smear. Leukemia detection with proposed features is classified using kNN classifier and offers an accuracy of 93%. Whereas Bakht Azam et .al in [8] involves segmenting Leukocytes in the blood smear. This method is simple and assures fine accuracy with excellent reduction of time complexity. The aim of Nagabhushana R M et .al in [9] involves nucleus segmentation of blood smear for leukemia detection. Primary importance is given to Feature Extraction phase. It offers an accuracy of 88%. Ms. Minal D. Joshi1 et .al in [10], the main theme of the paper lies in WBC nucleus segmentation of blood smear by relevant feature extraction processes. This system assures an accuracy of 93%.

ACKNOWLEDGMENT

TheI express me sincerely thanks to Mrs.Jucy Vareed (Asst. Professor, CSE Department) for taking part in discussions and giving necessary advices and guidance, facilitating preparation of this paper.

REFERENCES

[1] Fabio Scotti University of Milan, Department of Information Technologies, via Bramante 65, 26013 Crema, Italy

[2] N. Sinha and A. G. Ramakrishnan, “Blood cell segmentation using EM algorithm,” in Proc. 3rd Indian Conf. Comput. Vis., Graph., 2002, pp. 445–450.

[3] C.Reta,L.Altamirano,J.A.Gonzalez,R.Diaz,andJ.S.Guichard,“Segmentation of bone marrow cell images for morphological classification of acute leukemia,” in Proc. 23rd FLAIRS, 2010, pp. 86–91.

[4] Piuri and F. Scotti, “Morphological classification of blood leucocytes by microscope images,” in Proc. CIMSA, 2004, pp. 103–108.

[5] D. Foran, D. Comaniciu, P. Meer, and L. A. Goodell, "Computer-assisted discrimination among malignant lymphomas and leukemia using immunophenotyping, intelligent image repositories, and telemicroscopy," IEEE Trans. Inf. Technol. Biomed., vol. 4, no. 4, pp. 265–273, Dec. 2000.

[6] N. Sinha and A. Ramakrishnan, "Automation of differential blood count," in Proc. Conf. Convergent Technol. Asia-Pac. Region, 2003, vol. 2, pp. 547–551.

[7] Niranjan Chatap and Sini Shibu "Analysis of blood samples for counting leukemia cells using Support vector machine and nearest neighbor", e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 16, Issue 5, Ver. III (Sep – Oct. 2014), PP 79-87.

[8] Bakht Azam, Rashid Jalal Qureshi, Zahoor Jan, and Taj Ali Khattak "Color Based Segmentation of White Blood Cells in Blood Photomicrographs Using Image Quantization", Research Journal of Recent Sciences, Vol. 3(4), 34-39, April (2014).

[9] Nagabhushana R M and Rajeshwari P "Geometrical Analysis of Leukocyte Nucleus to Detect Lymphoblast from Microscopic Blood Images", International Journal of Technology and Engineering Science [IJTES]TM Volume 2[6], pp: 1947-1951, June 2014.

[10] Ms. Minal D. Joshi1, Prof. A.H.Karode and Prof. S.R.Suralkar "Detection Of Acute Leukemia Using White Blood Cells Segmentation Based On Blood Samples", IJECET Volume 4, Issue 3, May – June, 2013, pp. 148-153.

