Volume 8, No. 5, May-June 2017
ISSN No. 0976-5697
International Journal of Advanced Research in Computer Science
REVIEW ARTICLE
Available Online at www.ijarcs.info

Automated Detection of Leukemia

Rupinder Kaur Sandhu
PhD student , Computer Engineering
Punjabi University, Patiala
Patiala, India

Dr. Raman Maini
Professor, UCOE
Punjabi University, Patiala
Patiala, India

Abstract: Blood is a very important component of our body. The main components in the blood are RBCs (Red Blood Cells), WBCs (White Blood Cells), plasma and platelets. Blood serves as an indicator for identifying many diseases namely anemia, malaria, thalassemia, Leukemia etc. Abnormality in the shape and count of white blood cells causes different diseases. Leukemia is one of them. This paper deals with the identification of Leukemia. Leukemia is the cancer of blood. Many methods have been proposed for the same in the literature. Complete review on the papers on Leukemia detection is done in the literature survey. For proper diagnosis of Leukemia image preprocessing has to be done which includes contrast enhancement and image segmentation techniques. Finally the extracted features are compared with those stored in the database and finally classified as AML and ALL

Keywords: Leukemia, AML, ALL, CML, CLL, Lymphoid, Artificial Neural Network

INTRODUCTION

Leukemia is the cancer of blood. It can be classified as acute or chronic. Acute leukemia is further classified into acute lymphocytic leukemia (ALL) and acute myelogenous leukemia (AML). Abnormal white blood cell known as blast consists of cytoplasm and nucleus. Both need to be separated properly for the proper diagnosis of leukemia. Identification of lymphoid or myeloid series can be done by the shape of the nucleus. Myeloid has a regular shape that is it has circular nucleus whereas lymphoid doesn't have a regular nucleus. Nucleus of lymphoid is sparse whereas that of myeloid is dense. Image segmentation is done to separate WBC from the image. Contrast enhancement techniques ease the classification of leukemia into ALL and AML. Bright stretching, dark stretching and partial contrast are the techniques to enhance the contrast of the image. Partial contrast can be combined with bright or dark stretching to improve the results. Before any diagnosis can be done the counting area has to be identified. Counting area refers to the area where cells do not interfere with each other and are properly separated. Once the area is identified then extraction of the area of interest can be performed.

LITERATURE SURVEY

Identification of counting area in hematology is very important and slide maker devices are very expensive to be afforded by medium sized labs. To facilitate the identification of counting area an approach is proposed in [1]. Many papers can be found in literature for counting red blood cells. In [2] comparison of two algorithms i.e. 'Connected component labeling' and 'Back projection of Artificial Neural Network' is done. Parameters utilized are number of hemoglobine, packed cell volume and red blood cells count. The results were benchmarked using a hematological analyzer. Hybrid Flexible Neural Tree Approach have been used for classifying the Leukemia Cancer. Two types of sets are used in the model namely hybrid flexible neuron instructor for joining the non-leaf node's subtree and terminal set for inputting the leaf node instructor [3]. For proper identification of diseases it is very important to identify which information is relevant and which is irrelevant. 'Ask expert' approach is used in [4]. Proposed approach cannot deal with the unknown diseases and the development cost is also high. In [5] an open source based system is developed for the retrieval, analysis and interpretation of leukemia images. Quad Tree was applied to represent successive subdivisions of the image due to its capability of handling the disjoint structures. Open source software 'Imgseek' was used to evaluate the performance of retrieving images. The proposed system clearly marks a distinction between AML and CML images whereas 'Imgseek' cannot. Vacuoles are used to identify a specific type of AML but it was not considered in the proposed system in order to reduce the complexity (only cytoplasm, nucleus and nucleolus were considered). In [6] another approach for classification of haematalogic diseases viz malaria and thalassemia is proposed using artificial intelligence. Before preparing dataset for training and testing various image processing tasks were performed including size normalization, color normalization, specifying ROI, extracting ROI, edge erosion etc. Three layer and four layer neural networks are employed for classification and the three layer structure has the better performance. [7] presents literature based on granulometry. It also considers agglutination and edge fracture as the side effects of pre-processing.

DCC Technique along with IFS is used to compare set of images for object recognition. Performance of the proposed technique is compared with the existing methods for the three parameters: space, time and execution efficiency. Performance for space and execution efficiency is higher for the proposed method but it takes longer time to execute [8]. In [9] cancer and blood disorders are predicted using MLP with BP learning algorithm. [10] discuses automatic diagnosis of leukemia using a Hybrid Fuzzy SVM Classifier and concludes that filtered gene classifier more accurately classifies as compared to the unfiltered classifier since redundant information is excluded. [11] uses a simple method for estimation of hemoglobin in blood by first taking the photograph of the thumb then a rubber band is used to increase the flow of blood in the thumb and again it is photographed and both images are

then compared. Both images have to be taken under same standard conditions. In [12] the Hb levels in human blood were estimated using the BPN Architecture consisting of five layers and was compared with the standard cyanmethemoglobin method. It was also verified that the results do not offer more accuracies with more than three layers but it took more time to train the network. [13] discuses Flow Cytometry in the form of a game for the recognition of cancer. Technique determines two diseases closer to the goal then it is the task of the doctor to diagnose the exact disease between the two results. In [14],[15],[16],[17] detection of Acute Lymphocytic Leukemia is done and classification is done using Support Vector Machine. Color image segmentation is done and RGB Color space is converted to L*a*b*before clustering. Two stage segmentation is done first using K means clustering then by nearest neighbor classification. Images with many lymphocytes are considered in the proposed scheme. [18] proposed three contrast enhancement techniques namely bright stretching, dark stretching and partial contrast for the color images(using RGB components) and concludes partial contrast as the best technique. For improving the accuracy of segmentation a combination of image segmentation (HSI based) and contrast enhancement techniques is proposed in [19]. Combination of partial contrast and bright stretching achieves fully segmented blast whereas partial contrast and dark stretching obtains fully segmented nucleus. [20] also does color segmentation of WBC's to obtain acute leukemia images. It combines color segmentation (HSI based) and linear contrast technique. Segmentation of WBC is done based on H and S components of the image. Finally median filter and region growing are applied on the resultant images thus improving the segmentation accuracy. Comparison of segmentation based on HSI and RGB color space is done in [21] and concluded HSI as a better approach for segmenting color images. Subtyping of leukemia is done in [22] using novel compressive sensing based approach. Leave-one-out is used for cross validation for evaluating the accuracy of detection. 100% accuracy is achieved for ALL group but due to the difference in sample size misclassification occurs in AML group. [23] illustrates that SVM Model gives highest accuracy for the classification of leukemia cancer. Diagnosis of leukemia from the bone marrow blood cell images is done in [24]. Detection rate is below the usable value for the promyelocyte and is identified as the further work to be done.

METHODOLOGY

Algorithm:

1) Identify the counting area (area where cells do not interfere with each other and are properly separated).
2) Perform color normalization on the image.
3) Separate white blood cells from the image by performing image segmentation.
4) Specify the region of interest.
5) Extract the region of interest.
6) Apply edge erosion filter to deal with the fractured edges.
7) Feature vector analysis is performed which includes extraction and selection of features of the cell images including size of cell, circularity of cells, count of cells, area of cytoplasm and nucleus, nucleus/cytoplasm ratio etc.
8) Initial classification into lymphoid or myeloid series is done by the shape of the nucleus.
9) The extracted features are then compared with those stored in the database and classification is done into AML or ALL categories.
10) K nearest neighbor classifier is used for the classification.

RESULTS & DISCUSSION

Table I. **Percentage of Cell Segmentation Accuracies for Different Methods**

S. No.	Technique used	Result (%)
1	Watershed Algorithm Optimal Thresholding [**Error! Bookmark not defined.**]	99.85 %
2	Scale space filtering Watershed clustering [25]	98.9 %
3	Morphological operations Watershed transform [26]	95.09 %

Table II. **Results of segmentation of lymphocytes using different techniques**

S. No.	Technique used	Result of segmentation	
		Nucleus (%)	Cytoplasm (%)
1	Morphological analysis using snake algorithm Pixel intensity thresholding [27]	92 %	78 %
2	Watershed algorithm Optimal Thresholding [28]	99.92 %	99.63%
3	Image arithmetic Automatic thresholding using contrast stretching [29]	98.81 %	85-98 %

CONCLUSION AND FUTURE WORK

Color feature must be added to improve the classification rate of the leukemia images. It is difficult to segment the images in RGB color space. So, it is better to convert the images into L*a*b* or HSI color space. If in addition to cytoplasm, nucleus and nucleolus; vacuoles are also considered then a specific type of AML could be determined though the complexity will increase.

REFERENCES

[1] Rupp, S., Schlarb, T., Haslmeyer, E., Zerfas, T. "Fully automated detection of the counting area in blood smears for computer aided hematology", Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE, p. 7759 – 7762, 2011.

[2] Nasution, A.M.T.; Suryaningtyas, E.K.**, "**Comparison of Red Blood Cells Counting using two Algorithms: Connected Component Labeling and Backprojection of Artificial Neural Network **",** PhotonicsGlobal@Singapore , 2008. IPGC 2008., p.1-4, 2008.

[3] Lei Zhang, Yuehui Chen, Ajith Abraham, Zhenxiang Chen, "Hybrid flexible neural tree approach for leukemia cancer classification", Information and Communication Technologies (WICT), p.32 – 35, 2011.

[4] Benjamins, V.R.; Albuquerque, A.R.P.L., "Artificial Intelligence Approaches For Cytological Image Interpretation", Ieee International Conference On Systems, Man, And Cybernetics, 1994. 'Humans, Information And Technology'., 1994 Volume: 3 , P. 2306 - 2310, 1994.

[5] Rajendran, S., "Image Retrieval Techniques, Analysis and Interpretation for Leukemia Data Sets", 12th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD) , p. 19-23, 2011.

[6] Hirimutugoda, Y.M.; Wijayarathna, G., "Artificial Intelligence-Based Approach for Determination of Haematalogic Diseases", 2nd International Congress on Image and Signal Processing, CISP '09, p. 1-5, 2009.

[7] Saberi, A.; Kheiri, K., "Granulometry deficiencies in blood smears particles specification", International Conference on Signal and Image Processing (ICSIP), p. 271 – 274, 2010.

[8] J. Ferdin Joe, T. Ravi, A. Natarajan, S. Prakash Kumar, "Object recognition of Leukemia affected cells using DCC and IFS", International Conference on Computing Communication and Networking Technologies (ICCCNT), p. 1 –6, 2010.

[9] Vahid Aeinfar, Hoorieh Mazdarani, Fatemeh Deregeh, Mohsen Hayati, Mehrdad Payandeh, "Multilayer Perceptron Neural Network with supervised training method for diagnosis and predicting blood disorder and cancer", IEEE International Symposium on Industrial Electronics, p. 2075 – 2080, 2009.

[10] Perez, M., Rubin D.M., Marwala TM, Scott LE, Stevens W, "A Hybrid Fuzzy-SVM classifier, applied to gene expression profiling for automated leukaemia diagnosis", IEEE 25th Convention of Electrical and Electronics Engineers in Israel, p. 041 – 045, 2008.

[11] K S Srinivasan, D Lakshmi, H Ranganathan, N Gunasekaran, "Non-Invasive Estimation of Hemoglobin in Blood Using Color Analysis", First International Conference on Industrial and Information Systems, p. 547 – 549, 2006.

[12] Ranganathan, H.; Gunasekaran, N., "Simple Method for Estimation of Hemoglobin in Human Blood Using Color Analysis", IEEE Transactions on Information Technology in Biomedicine, Volume: 10 , Issue: 4 p. 657– 662, 2006.

[13]Atefeh Torkaman, Nasrollah Moghaddam Charkari, Mahnaz Aghaee Pour, Marjan Bahrololum, Esmerdis Hajati, "An expert system for detection of leukemia based on cooperative game theory", IEEE 35th Annual Northeast Bioengineering Conference, p. 1-2, 2009.

[14] Mohapatra, S.; Patra, D. , "Automated cell nucleus segmentation and acute leukemia detection in blood microscopic images ", International Conference on Systems in Medicine and Biology (ICSMB), p. 49-54, 2010.

[15] Mohapatra, S.; Patra, D., "Automated leukemia detection using hausdorff dimension in blood microscopic images ", International Conference on Emerging Trends in Robotics and Communication Technologies (INTERACT), p. 64-68, 2010.

[16] Subrajeet Mohapatra, Dipti Patra, Sanghamitra Satpathi, "Image analysis of blood microscopic images for acute leukemia detection", International Conference on Industrial Electronics, Control & Robotics (IECR), p. 215-219, 2010.

[17] Mohapatra Subrajeet, Sushanta Shekhar Samanta, Dipti Patra and Sanghamitra Satpathi, "Fuzzy Based Blood Image Segmentation for Automated Leukemia Detection ", International Conference on Devices and Communications (ICDeCom), p. 1-5, 2011.

[18] A. N. Aimi Salihah, Mohd Yusoff Mashor, Nor Hazlyna Harun and H. Rosline, "Colour image enhancement techniques for acute Leukaemia blood cell morphological features ", IEEE International Conference on Systems Man and Cybernetics (SMC) , p. 3677 – 3682, 2010.

[19] A.N. Aimi Salihah, M.Y. Mashor, Nor Hazlyna Harun, Azian Azamimi Abdullah, H. Rosline, "Improving colour image segmentation on acute myelogenous leukaemia images using contrast enhancement techniques ", IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES), p. 246 - 251 , 2010.

[20] Abdul Nasir, A.S. , M.Y. Mashor and H. Rosline, "Unsupervised colour segmentation of white blood cell for Acute Leukaemia images", IEEE International Conference on Imaging Systems and Techniques (IST) , p. 142 - 145, 2011.

[21] Harun Nor Hazlyna, M.Y. Mashor, N.R. Mokhtar, A.N. Aimi Salihah, Rosline Hassan, R.A.A. Raof, M.K. Osman, "Comparison of acute leukemia Image segmentation using HSI and RGB color space ", 10th International Conference on Information Sciences Signal Processing and their Applications (ISSPA), p. 749 - 752 , 2010.

[22] Wenlong Tang, Hongbao Cao, Yu-Ping Wang, "Subtyping of Leukemia with Gene Expression Analysis Using Compressive Sensing Method ", First IEEE International Conference on Healthcare Informatics, Imaging and Systems Biology (HISB), p. 76 – 80, 2011.

[23] Sung-Huai Hsieh, Zhenyu Wang, Po-Hsun Cheng, I-Shun Lee, Sheau-Ling Hsieh, Feipei Lai, "Leukemia cancer classification based on Support Vector Machine", 8th IEEE International Conference on Industrial Informatics (INDIN), p. 819 – 824, 2010.

[24] Sung-Huai Hsieh, Zhenyu Wang, Po-Hsun Cheng, I-Shun Lee, Sheau-Ling Hsieh, Feipei Lai, "Leukemia cancer classification based on Support Vector Machine", 8th IEEE International Conference on Industrial Informatics (INDIN), p. 819 – 824, 2010.

[25] Jiang, K., Liao, QM. & Xiong, Y., "A novel white blood cell segmentation scheme based on feature space clustering", Soft Computing, vol. 10, pp. 12-19, 01/01, 2006.

[26] Lim Huey Nee, Mohd Yusoff Mashor and Rosline Hassan, "White Blood Cell Segmentation for Acute Leukemia Bone Marrow Images", International Conference on Biomedical Engineering (ICoBE), Penang, Malaysia, pp. 357-361,27-28 February 2012.

[27] Sadeghian F, Seman Z, Ramli AR, Abdul Kahar BH, Saripan MI., "A framework for white blood cell segmentation in microscopic blood images using digital image processing", Biol. Proc. Online, vol. 11, pp. 196-206, 2009.

[28] Mohammed, Emad, Mostaja MA Mohamed, Christopher Naugler, and Behrouz H. Far., "Chronic lymphocytic leukemia cell segmentation from microscopic blood images using watershed algorithm and optimal thresholding", 26th IEEE Canadian Conference of Electrical and Computer Engineering (CCECE) , 2013.

[29] H. T. Madhloom, S.A. Kareem, H. Ariffin, A. A. Zaidan, H.O. Alanazi and B.B. Zaidan, "An automated white blood cell nucleus localization and segmentation using image arithmetic and automatic threshold", J. App. Sci., vol. 10, pp. 959-966, 2010.