



AUTOMATED SCREENING SYSTEM FOR ACUTE MYELOGENOUS LEUKEMIA DETECTION IN BLOOD MICROSCOPIC IMAGES

Abirami.V 1* , B. Muthu Kumar 2 , Kavitha 3 , S.Prabhakaran 4

1, 3Department of Computer Science, Sakthi College of Arts and Science for Women, TN, India

2 Professor, Department of CSE, Syed Ammal Engineering College, Ramanathapuram, TN, India

4 Department of Information Technology, Kathir College of Engineering

*Corresponding Author: mailto: abiramibca10@gmail.com

ABSTRACT

The pervasive among the adults are Acute Myelogenous Leukemia (AML) is a sub-type of acute leukemia. The person has AML mostly at the age of 65 years. The current method implements the manual examination of the blood smear during the diagnosis. It needs for the detection of leukemia [11]. It is favor for the operators' ability, accuracy and time consumption. AML is presented a small technique in blood smear. The proposed method includes 1) this approach developed for simplicity 2) classification of complete blood smear images as opposed to sub-images. 3) this algorithm is used to segment and detect nucleated cells. The computer simulation is under the testers comparing the impact of Hausdorff [7] dimension on the system before and after the influence of local binary pattern. The sub-images and the whole images performance and the results are compared to the existing system and proposed system. Our aim is to achieve 98% accuracy for the localization of the lympho [14] blast cells. Already the microscopic blood images are tested.

Keywords: Acute Myelogenous Leukemia (AML), classification, segmentation.

1. INTRODUCTION

Our goal is to a) illustrate that the classification of peripheral blood smear images containing multiple nuclei can be fully automated b) the hold-out cross validation method is used to validate and segment the images c) this method estimates around a 50 set of images (25 abnormal samples and 25 normal samples) from American Society of Hematology. The segmentation and classification of AML based on complete microscopic blood images is our proposed method. The term classification [13] of leukemia is mostly finding the problem accurately.

In our proposed system, it can extract the shape based approaches like lines, sheets from the 3-D biomedical images. The ellipsoidal model is considered as thin structures. In existing system, Gaussian filter method are implemented to simplify and for getting segmentation results. Using CIELAB color space and various edge detection



algorithms, the segmentation is performed. The edge detection methods consolidated are Sobel, Prewitt, Roberts, Laplacian of Gaussian and Canny edge detectors^[17]. The best detection technique is canny edge detectors.

The canny implements two different threshold values to identify the strong and weak edges. The weak edges are identified only if they are connected to strong edges. A new method is implemented for segmentation is Gabor filters for extraction and segmentation of tagged cardiac images. The designs of adaptive filters for different local regions are implemented by Gabor filters. Because the wavelets^[12] are like local filters in spatial domain. The advantage of Gabor filter is they can combine with Gaussian envelopes. The minimum space bandwidth is achieved by these filters. Gabor filters are used widely in image processing applications like texture segmentation and edge detection. Gabor filters are used widely in image processing applications like texture segmentation and edge detection. [5] proposed a system which uses intermediate features of maximum overlap wavelet transform (IMOWT) as a pre-processing step. The coefficients derived from IMOWT are subjected to 2D histogram Grouping. This method is simple, fast and unsupervised. 2D histograms are used to obtain Grouping of color image. This Grouping output gives three segmentation maps which are fused together to get the final segmented output. This method produces good segmentation results when compared to the direct application of 2D Histogram Grouping. IMOWT is the efficient transform in which a set of wavelet features of the same size of various levels of resolutions and different local window sizes for different levels are used. IMOWT

is efficient because of its time effectiveness, flexibility and translation invariance which are useful for good segmentation results.

2. RELATED WORKS

SEGMENTATION OF 4D CARDIAC MRI: AUTOMATED METHOD BASED ON SPATIO-TEMPORAL WATERSHED CUTS

J. COUSTY; L. NAJMAN

A new automated and fast procedure method is proposed to segment the left ventricular myocardium in 4D MRI sequences. The quantitative and qualitative evaluations^[20] are provided. The discrete mathematical morphology is used and the time efficiency is high. The ability of the method to compute reliable characteristics of the LV (ejection fraction and left ventricular mass); the temporal continuity of the resulting automated segmentation and the time-efficiency (about 3' to segment a sequence of 25 3D-images on a low-end computer) of the proposed method is checked for the accuracy. The cardiac segmentation^[3] method is difficult to design. For that, the watershed^[1] cuts, a notion of watershed in edge-weighted graphs which is optimal in a sense equivalent to minimum spanning trees is introduced. The watershed-cuts in 4-dimensional spaces are proposed. Our second goal is to improve the ability of both the spatial and temporal gradient of the images. The segmentation produced the spatially as well as temporal consistent.

The proposed method can be fairly used to assess model-based segmentation schemes and the results are compared with the non-model based segmentation. The proposed method is used to register generic physiological models of the heart to real patient specific cardiac images. It can be



easily register a model to a binary segmentation than directly to images. The successive segmentations obtained over the time take into account spatio-temporal properties of the images. The proposed scheme does not permit the direct derivation of deformation parameters. The accuracy is improved by registration with other modalities such as delayed enhanced MRI and CT scan.

A REVIEW OF SEGMENTATION METHODS IN SHORT AXIS CARDIAC MR IMAGES

The fully and semi-automated methods performing segmentation in short axis images using a cardiac cine MRI ^[4] sequence is discussed. To tackle the ventricle segmentation in cardiac MRI, an image-driven is proposed. These methods require either minimal or user intervention. The image based and the pixel classification-based frameworks are incorporating strong prior, straightforward extensions of deformable models. The next section presents strong prior for heart segmentation. It can be generated by manually segmenting an image or by integrating information from multiple segmented images from different individuals. Strong prior based methods overcome the segmentation problems. We have proposed a categorization for these methods, highlighting the key role of the type of prior information used during segmentation, and has distinguished three levels of information:

- i. No information is used.
- ii. Weak prior i.e, low level information.
- iii. Strong prior information.

CARDIAC LV AND RV SEGMENTATION USING MUTUAL CONTEXT INFORMATION

In this paper, a graph cut based method is implemented to segment the cardiac right ventricle (RV) and left ventricle (LV) by using mutual context information. It includes a 'context penalty' for the RV by learning its geometrical relationship with respect to the LV. Similarly, the RV provides geometrical context information for LV ^[2] segmentation. The smoothness cost is formulated for the learned context and captures the geometric relationship between the RV and LV. The mutual context information method is to segment the RV and LV. The geometric relationship is used for the shapes in the form of relative orientations. The RV and LV are manually segmented and they are learned from a set of training images. The learned information is encoded on a graphical model of the image as weights between pixel nodes and terminal nodes. Graph cuts are used to find the final labels in an iterative fashion. This paper makes the following contributions:

- 1) We encode context information for each pixel as well as between pixels. The Graph cuts do not need the sub-modular graphs. The distance information is not needed.
- 2) The RV and LV is used to segment the individual organs.

The RV and LV are used to improve segmentation accuracy. Context information was modeled in the form of relative orientation of the two organs from a set of training images. The two penalties and smoothness cost leads to improve the segmentation performance.



3. CLUSTERING BASED SEGMENTATION

Clustering means measurement of points or patterns are grouped together. This technique implements in data of n-dimension. This n represents arbitrary number i.e it can be two, three or more. The cluster technique is a best suit for sparse type of images. This technique includes methods like k-means, fuzzy ^[8] c-means etc. SubrajeetMohapatra & DeeptiPatra proposed an automated nucleus segmentation method. For segmentation two steps are performed to segment a WBC nucleus ^[10] from the rest of the image objects. The first step, the segmentation is performed by executing a semi-supervised k-means clustering. The second step of segmentation is performed by nearest neighbor classification in L^*a*b space.

In blood microscopic ^[9] images and the white blood cells ^[16] are segmented by using a cluster method. In our proposed system, the first step to convert the images into Lab color space. In fuzzy, k-means clustering is to divide the images into three clusters. At the same time, automatic histogram is performed on the Lab color image. The results of clusters and the reference image are compared, selected and performs a logical AND operation.

In the proposed approach first the RGB image is converted to Smith's HSI transformation. Then a membership function is assigned to each color pattern. After obtaining the membership degree for color patterns, pixel classification is carried out for segmentation purposes. A fuzzy approach ^[19] for leukemia detection is proposed.

3.1 INPUT CELL IMAGE

The input is taken as blood cell image. Basically the data are collected from the outside sources. If the blood cell is visible size then the refraction effects are almost negligible. The disc function is used to approximate the real images. Comparing to MRI images, the disc approximation is enough for testing images. The 3-D geometric information is used to identify in-focus subjects by getting the local blood cell ^[15] images for testing samples.

3.2 EVALUATE CELL FEATURES

In image processing, feature extraction ^[6] method is used to redefine a large set of redundant data and a set of features from reduced dimension. The input data are transformed to the set of features is called feature extraction. This method influences the classifier performance and it is very crucial to select the correct one. To give an effective feature set we have to analyze several published articles and methods are absorbed. The features are widely used for a good classification and in our proposed method implements the features on whole images. It is used to boost the classifier performance. This technique is implemented on a segmented image to identify the location and size of complete non-overlapping cells in a microscopic image.

3.3 SEGMENTATION USING k-MEANS ALGORITHM

The least squares partitioning method that divide a collection of objects is called k-means or k-groups. The algorithm follows the below steps:

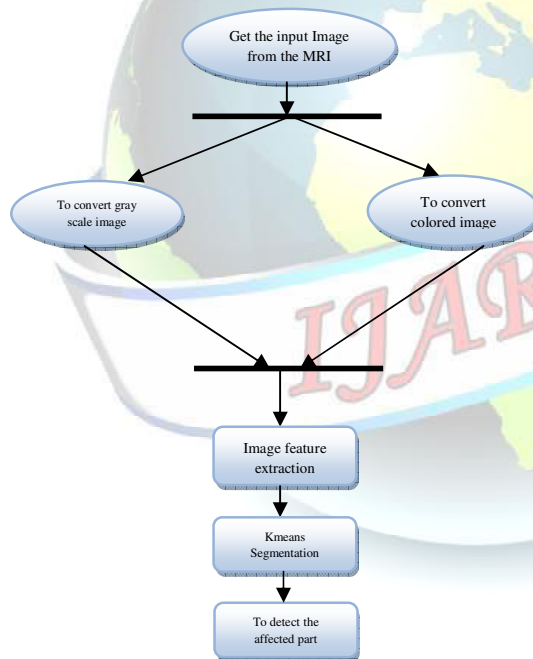
1. Compute the mean of each cluster.
2. Compute the distance of each point from each cluster by computing its distance



from the corresponding cluster mean. The each nearest point is assigned to the cluster.

Initially the clusters are used to assign randomly. This iteration implements to minimum the sum, overall groups, squared within the group errors. These are the distance of the points to the respective group means. Convergent occurs when it reach the residual sum of squares cannot be lowered any more. The roots obtained the geometrically compact as possible around their respective means.

3.4 ARCHITECTURE



3.5 FEATURE CLASSIFICATION

The classification algorithm includes the identification of images and the assumption of images in depicts one or more features for classifying the system or spectral regions in the

case of remote sensing. This feature belongs to several distinct and exclusive classes. The classes are specified on a priori by an analyst on in supervised classified and unsupervised [18] classification. The set of prototype classes are specified the number of desired categories.

3.5.1 PROCESS

The classification of algorithm employ the two phases of processing are training and testing. The training phase includes the typical image features are isolated and unique classification category. The motivating criteria for constructing training classes are that they are:

- *independent*, i.e. a change in the description of one training class should not change the value of another,
- *discriminatory*, i.e. different image features should have significantly different descriptions, and
- *reliable*, all image features within a training group should share the common definitive descriptions of that group.

3.6 CLASSIFICATION

This classification clear cuts the cancer and non-cancer blood cells to get the spitted part from k-means segmentation. Classification generates the result for k-means cluster and also in support vector machine.



4. METHODOLOGY

4.1 k-MEANS ALGORITHM

The k-means clustering algorithm was developed by J. MacQueen (1967) and then by J. A. Hartigan and M. A. Wong around 1975. This algorithm is to classify or to group the objects based on attributes/features into K groups. After an initial random assignment of data points to k clusters, the centers of clusters are computed. The data points are allocated to the clusters with the closest centers. This process continues until the cluster centers do not significantly change.

The data set of each cluster's architecture is updated. During the updation of clusters, data points are removed from one cluster and added to another. The updating clusters changes the values of the centroids. This change is a reflection of the current cluster data points.

When the cluster is not change, the training of the k-means algorithm is complete. The 'k' cluster centroids are created and it is ready to classify the patient's statuses as normal, hyperthyroid and/or hypothyroid function.

The process gives a easy way to classify the data set through a certain number of clusters (assume k clusters) fixed a priori. The aim is to define k centroids by assigning one to each cluster. These centroids should be placed in a crafty way to avoid location problem. The next step is to check with the given data set and associate it to the nearest centroid. It is necessary to re-calculate k new centroids.

After obtaining k new centroids, a new binding has to be done between the same data points and the nearest new centroid. The loop may change their location step by step until no more changes are to perform in the k centroids.

The objective function

$$J = \sum_{j=1}^k \sum_{i=1}^n \| \mathbf{x}_i^{(j)} - \mathbf{c}_j \|^2 \quad (1)$$

Where $\| \mathbf{x}_i^{(j)} - \mathbf{c}_j \|^2$ is a chosen distance measure between a data point $\mathbf{x}_i^{(j)}$ and the cluster center \mathbf{c}_j is an indicator of the distance of the n data points from their respective cluster centers.

The k-means algorithm can be run multiple times to reduce this effect. This algorithm is used to work for a randomly generated data points. It is used to find the locally minimal solution. The dissimilarity measure is represented as the Euclidean distance.

A set of n vectors \mathbf{X}_j , $j = 1 \dots n$, are to be partitioned into C groups G_i , $i = 1, \dots, C$. The cost function, based on the Euclidean distance between a vector \mathbf{x}_k in group j and the corresponding cluster center \mathbf{C}_i , can be defined by:

$$J = \sum_{i=0}^c J_i = \sum_{i=1}^c \left[\sum_{\mathbf{x}_k \in G_i} \| \mathbf{x}_k - \mathbf{C}_i \|^2 \right] \quad (2)$$

Where $J = \sum_{\mathbf{x}_k \in G_i} \| \mathbf{x}_k - \mathbf{C}_i \|^2$ is a cost function within group i . There are two problems that are inherent to k-means clustering algorithms. The first is determining the initial partition and the second is determining the optimal number of clusters.

If the cluster is in the centre of k-means algorithm is used to run several times with the



different set of initial cluster centre to improve performance.

5. RESULTS AND DISCUSSION

5.1 PROBLEM IDENTIFICATION

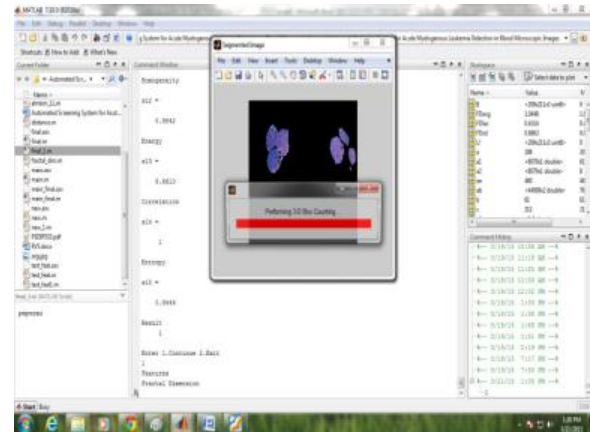
Detection accuracy is low so that segmentation task is complex. The major problem occurs are Image processing and pattern recognition problems. The LBP method performs the outsource of existing methods are the linear discriminant analysis and the principal component analysis. In this system performs classification of whole images and also includes the better performance for sub-images. The proposed technique is to achieve an automatic classification system to diagnosis the presence of the acute leukemia from blood microscope images. Comparing segmentation of nucleus to the segmentation of the entire cell is much easier. In the bone marrow, the WBC density is very high. A set of manually segmented images of the nucleus are used to decouple segmentation errors. A microscopic blood image of size 184×138 is considered for evaluation.

Drawbacks

- Accuracy is low
- Cannot provide optimized cancer detect
- Segmentation process have some trouble

5.2 PROBLEM SOLVING

The system implements on peripheral blood smear images obtained from two places. Here the method performs automated processing, including color correlation, segmentation of the nucleated cells, and effective validation and classification.



The feature set used to perform the efficient classification among the exploiting shape, color, and texture parameters of a cell. The impact of the LBP operator on the HD proved to be a promising feature for this analysis. A color feature called cell energy was introduced, and results show that this feature presents a good demarcation between cancer and non-cancer cells.



Advantages

- It provides high accuracy
- It provide optimized cancer detect
- No problem on this algorithm
- Here we used K-Means Algorithm for segmentation and classification
- Performance also high

6. CONCLUSION

Our proposed system performs automated processing includes color correlation,



segmentation of the nucleus, effective validation and classification. The shape, color and texture parameters of a cell are utilized from the feature set and to obtain the information and to perform efficient classification. The FD proved promising feature for clustering operator. A color feature called cell energy and they presented good demarcation between cancer and non-cancer cells. The affected part of an MR image is detected. The label information is used to improve the image registration accuracy. It provides useful information for clinicians in cancer disease diagnosis.

REFERENCES

- [1] J.Cousty, L.Najman, M. Couprie, S. Clément-Guinaudeau, T. Goissen, And J. Garot, "Segmentation of 4D cardiac MRI: Automated method Based on spatio-temporal watershed cuts,"
- [2] M. Jolly, "Fully automatic left ventricle segmentation in cardiac Cine MR Images using registration and minimum surfaces," MIDAS.
- [3] M. Lorenzo-Valdés, G. I. Sanchez-Ortiz, A. G. Elkington, R. H. Mohiaddin and D. Rueckert, "Segmentation of 4D cardiac MR images using a probabilistic atlas and the EM algorithm,"
- [4] M. Lynch, O. Ghita, and P. F. Whelan, "Segmentation of the left ventricle Of the heart in 3-MRI data using an optimized Non-rigid Temporal model,"
- [5] Christo Ananth, A.S.Senthilkani, S.Kamala Gomathy, J.Arockia Renilda, G.Blesslin Jebitha, Sankari @Saranya.S., "Color Image Segmentation using IMOWT with 2D Histogram Grouping", International Journal of Computer Science and Mobile Computing (IJCSMC), Vol. 3, Issue. 5, May 2014, pp-1 – 7
- [6] G. Ogun, U. Halici, K. Leblebicioglu, V. Atalay, M. Beksac, and S. Beksac, "Feature extraction and classification of blood cells for an automated differential blood count system," in Proc. IJCNN, 2001, vol. 4, pp. 2461–2466.
- [7] S. Mohapatra and D. Patra, "Automated leukemia detection using haus-dorff dimension in blood microscopic images," in Proc. Int. Conf. Emerg. Trends Robot Commun. Technol., 2010, pp. 64–68.
- [8] S. Mohapatra, S. Samanta, D. Patra, and S. Satpathi, "Fuzzy based blood image segmentation for automated leukemia detection," inProc. ICDeCom, 2011, pp. 1–5.
- [9] S. Mohapatra, D. Patra, and S. Satpathi, "Image analysis of blood microscopic images for acute leukemia detection," inProc. IECR, 2010, pp. 215–219.
- [10] S. Mohapatra, D. Patra, and S. Satpathi, "Automated cell nucleus segmentation and acute leukemia detection in blood microscopic images," in Proc. ICSMB, 2010, pp. 49–54.
- [11] MedlinePlus: Leukemia.National Institutes of Health. [Online]. Available: <http://www.nlm.nih.gov/medlineplus/ency/article/001299.htm>
- [12] C. Lopez and S. Agaian, "A new set of wavelet- and fractals-based features for Gleason grading of prostate cancer histopathology images," inProc. SPIE, Image Process., Algorithms Syst. XI, 2013, vol. 8655, pp. 865516-1–865516-12.
- [13] S. Serbouti, A. Duhamel, H. Harms, U. Gunzer, J. Mary, and R. Beuscart, "Image segmentation and classification methods to detect leukemias," in Proc. Int. Conf. IEEE Eng. Med. Biol. Soc., 1991, pp. 260–261.
- [14] D. Foran, D. Comaniciu, P. Meer, and L. A. Goodell, "Computer-assisted discrimination among malignant lymphomas and leukemia using immune phenotyping, intelligent image repositories, and tele



microscopy," IEEE Trans. Inf. Technol. Biomed., vol. 4, no. 4, pp. 265–273, Dec. 2000.

[15] K. S. Kim, P. K. Kim, J. J. Song, and Y. C. Park, "Analyzing blood cell image do distinguish its abnormalities," in Proc. ACM Int. Conf. Multim., 2002, pp. 395–397.

[16] Q. Liao and Y. Deng, "An accurate segmentation method for white blood cell images," in Proc. IEEE Int. Symp. Biomed. Imaging, Atlanta, GA, USA, 2002, pp. 245–248.

[17] K. Panetta, S. Agaian, S. Nercessian, and A. Almunstashi, "Shape-dependent canny edge detector," Opt. Eng., vol. 50, no. 8, pp. 087008-1–087008-12, Aug. 2011.

[18] P. Bamford and B. Lovell, "Method for accurate unsupervised cell nucleus segmentation," in Proc. Eng. Med. Biol. Soc. Conf., 2001, vol. 3, pp. 2704–2708.

[19] W. Shitong, K. F. L. Chung, and F. Duan, "Applying the improved fuzzy cellular neural network IFCNN to white blood cell detection," Neurocom-puting, vol. 70, no. 7–9, pp. 1348–1359, Mar. 2007.

[20] M. Oberholzer, M. Ostreicher, H. Christen, and M. Bruhlmann, "Methods in quantitative image analysis," Histochem. Cell Biol., vol. 105, no. 5, pp. 333–355, May 1996.



Kavitha, Professor, Department of Computer Science, Sakthi College of Arts and Science for Women, TN, India



S. Prabhakaran, Department of Information Technology, Kathir College of Engineering, Now he is currently working in Ascox Techno Soft at Madurai, India. Him main research interests include Image Processing, Network Security, Cloud Computing.



V. Abirami received her MCA degree from RVS College of Engineering & Technology, Dindigul, India in 2015. Now she is currently working toward the M.Phil degree at Sakthi Arts & Science College for women, Oddanchtram, India. Her main research interests include Image Processing, Network Security.



B. Muthu Kumar, Professor, Department of CSE, Syed Ammal Engineering College, Ramanathapuram, TN, India