ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization Volume 3, Special Issue 5, July 2014

International Conference On Innovations & Advances In Science, Engineering And Technology [IC - IASET 2014]

Organized by

Toc H Institute of Science & Technology, Arakunnam, Kerala, India during 16th - 18th July -2014

# Fuzzy Local Information C Means Clustering For Acute Myelogenous Leukemia Image Segmentation

Meera V[1], Shammy Arun Mathew[2]

M.Tech Student (AE&I, Dept. of ECE), Lourdes Matha College of Science and Technology, Kuttichal, Trivandrum, Kerala, India[1]

Assistant Professor (Dept. of ECE), Lourdes Matha College of Science and Technology, Kuttichal, Trivandrum, Kerala, India[2]

**ABSTRACT**: Leukemia is a type of cancer that affects the blood cells and most commonly WBCs or leukocytes. There are two main types of acute leukemia: Acute lymbhoblastic Leukemia (ALL) and Acute Myelogenous Leukemia (AML). In this paper AML is only considered. Here microscopic blood smear images containing multiple nuclei are exposed to a series of preprocessing steps which includes color correlation and contrast enhancement. By performing FLICM clustering algorithm on the resultant images, the nuclei of the cells invested with cancer are obtained. The main objective is to demonstrate that the classification of peripheral smear images containing multiple nuclei can be fully automated and to validate the segmented images. The method has been evaluated using a set of 50 images (with 25 abnormal samples and 25 normal samples). The system robustly segments and classifies AML based on complete microscopic blood images. SVM is employed for classifying the nucleus images based on the extracted features in to healthy and leukemic. The developed system can be used as ancillary/backup service to the physician.

**KEYWORDS**: ALL, AML, preprocessing, color correlation, contrast enhancement, FLICM, SVM.

## I. INTRODUCTION

Leukemia is a type of cancer of the blood or bone marrow characterized by an abnormal increase of immature white blood cells called "blasts", which are unable to fight infection. Acute leukemia is a progressive disease which appears suddenly and need to be treated urgently. Leukemia that affects myeloid cells is called myelogenous leukemia. Thus AML is a cancer of the myeloid line of blood cells, characterized by the rapid growth of abnormal white blood cells that accumulate in the bone marrow and interfere with the production of normal blood cells.

Signs or symptoms of acute leukemia are similar to other common illnesses like weight loss or loss of appetite, excessive bruising or bleeding from wound. So this paper is about developing a cost effective and robust automated system for AML segmentation and classification. In the system, the microscopic blood images containing multiple nuclei are exposed to series of preprocessing steps which include color correlation, and contrast enhancement. By performing segmentation on the resultant images, the nuclei of the cells invested with cancer are obtained. Then features are extracted from the image and are classified using a classifier. The system will focus on WBC disease, leukemia because disease is dangerous and can lead to death.

ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization Volume 3, Special Issue 5, July 2014

International Conference On Innovations & Advances In Science, Engineering And Technology [IC - IASET 2014]

Organized by

Toc H Institute of Science & Technology, Arakunnam, Kerala, India during 16th - 18th July -2014

Despite of advanced techniques such as flow cytometer, immunophenotyping, molecular probing etc, microscopic examination of blood slides still remains as the standard leukemia diagnosis technique. This analysis suffers from time delays and it presents not a standardized accuracy since it depends on operator's capabilities and tiredness.The microscopic images are used to identify the types and maturity of blood cells but it is a time consuming and tedious job and is inadequate to identify the type of the cell. Also the equipment required is very costly and may not be exist in all hospitals and clinics. The practice of manual counting of the white blood cells suffer from the disadvantages associated with human errors. 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. Accurate segmentation of the white blood cells is very important to locate the cancer region.

The main objective of this paper is to a) demonstrate that the classification of peripheral blood smear images containing multiple nuclei can be fully automated, b) To segment the image more efficiently and to validate the segmented images.

This can be used as ancillary/ backup service to the physician inorder to improve efficiency and accuracy of clinical practice. It provide a software based cost effective and an efficient alternative in recognising and analysing blood cells. Here tedious human task is transformed in to a computer based process in which system outperforms the manual system.

## II. RELATED WORKS

From the review of related work and published literature, it is observed that many researchers used different segmentation techniques like thresholding method, region based approaches, edge detection approach, clustering approaches, artificial neural network, fuzzy technique, watershed algorithm etc for the segmentation of leukemia images and attempted to find better result.

S.Rubhala et al [7] presented a work to classify a lymphocyte as a normal or a lymphoblast. Here acute lymphocytic leukemia (ALL) is only considered. This paper presents a fast and effective segmentation procedure for blast images which is very helpful for improving the hematological procedure and accelerating diagnosis of leukemia diseases. A kernel-induced new metric was used to replace the Euclidean norm in fuzzy c-means algorithm in the original space and then derived the alternative kernel-based fuzzy c-means algorithm. Features are extracted from the segmented output and best features are analysed. SVM are employed for classification.

Ms. Minali et al [8] presented a method in which WBC nucleus segmentation of stained blood smear images followed by relevant feature extraction, selection and cell classification to the recognition and differentiation of normal cell from the blast cell for leukemia detection. In this paper Otsu's method of segmentation is utilized along with image arithmetic. The image histogram equalization and linear contrast stretching methods are also used. After this step simple arithmetic operations are used along with Global thresholding method to detect white blood cell nucleus. After image arithmetic operation, minimum filter is applied to remove noise. The paper mostly concentrates on measuring area, circularity, perimeter etc. features for better detection accuracy. Leukemia detection with the proposed features was classified with kNN classifier.

Stanislaw Osowski et al [9] presented the preprocessing methods of the leukemic blast cells image in order to generate the features well characterizing different types of cells. The solved problems include: the segmentation of the bone marrow aspirate by applying the watershed transformation, selection of individual cells, and feature generation on the basis of texture, statistical and geometrical analysis of the cells. These features are used as the input signals applied to the support vector machine used as the classifier for final recognition and classification of cells.

ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization Volume 3, Special Issue 5, July 2014

International Conference On Innovations & Advances In Science, Engineering And Technology [IC - IASET 2014]

Organized by

Toc H Institute of Science & Technology, Arakunnam, Kerala, India during 16th - 18th July -2014

## III. METHODOLOGY

In this paper cost effective and robust automated system for leukemia screening is developed. The method has been evaluated using a set of 50 images (with 25 abnormal samples and 25 normal samples). In the proposed system, the microscopic blood images containing multiple nuclei are exposed to series of preprocessing steps which include color correlation, and contrast enhancement. By performing FLICM clustering algorithm on the resultant images, the nuclei of the cells invested with cancer are obtained.

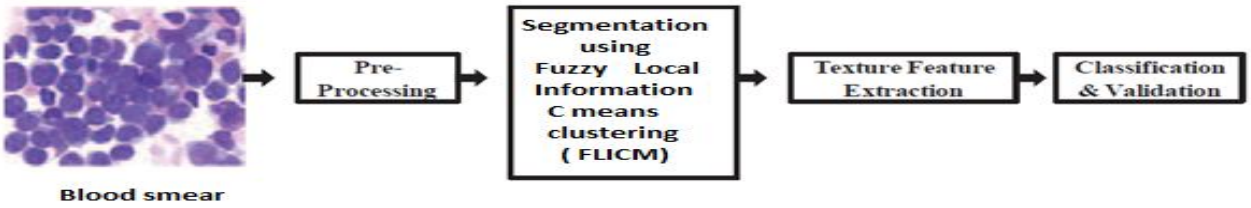


Fig 3.1. Overview of system using FLICM clustering

The proposed approach aims to present a more robust system with an efficient segmentation of blood images for high performance. To achieve this goal, the system we propose follows four main processing steps.

1) To preprocess the image in order to reduce background non-uniformities and perform color correlation.
2) To employ segmentation on whole images by combining different methods in order to exploit all the available a-priori information and thereby achieving a robust identification of the nuclei of the white cells.
3) To extract different sets of features for a database of images.
4) To run the classifier system and validate the output based on the results obtained.

### *A. Image Acquisition*

Blood image from slides will be obtained from nearby hospital with effective magnification and also from the web site http://www.hematologyatlas.com/leukemias.htm. Microscopic peripheral blood smear images containing multiple nuclei are considered. These all are the camera acquired images of blood samples. All images are in JPG format.

### *B. Pre-Processing: Color-Correlation*

Images generated by digital microscopes are usually in RGB color space which is difficult to segment. Thus RGB input image is converted into the CIELAB color space. The L*a*b* space consists of a luminosity layer L*, chromaticity layer a* and b*. Here the color information is represented in two components i.e. a* and b*. Due to less color dimension L*a*b* color space is mostly employed in color based clustering.

### *C. Segmentation: FLICM Clustering*

The segmentation technique is used to extract the nuclei from the complete blood smear images. It plays a key role since the efficiency of subsequent feature extraction and classification relies greatly on the correct segmentation of the blasts. In acute leukemia cell images, cytoplasm is scanty. So nucleus of WBC is focussed. To overcome the drawbacks of FCM such as for noisy images it does not take into account spatial information, which makes it sensitive to noise & other image artifacts, fuzzy local information C-means clustering algorithm (FLICM) is used for image segmentation.FLICM incorporates local spatial information and grey level information in a novel fuzzy way. FLICM is completely free of any

ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization Volume 3, Special Issue 5, July 2014

International Conference On Innovations & Advances In Science, Engineering And Technology [IC - IASET 2014]

Organized by

Toc H Institute of Science & Technology, Arakunnam, Kerala, India during 16th - 18th July -2014

parameter determination, while balance between the noise and image details is automatically achieved by the fuzzy local constraints, enhancing concurrently the clustering performance. Its characteristics include it provide noise immunity, it preserves image details and it is free of any parameter selection. We consider only the cluster which contains the blue nucleus, which is required for the feature extraction.

*D. Feature Extraction*

Transforming the input data into the set of features is called feature extraction. Feature extraction plays a crucial role in obtaining relevant information from the input data. A set of features are extracted inorder to allow a classifier to distinguish between normal and abnormal pattern. The abnormality can be identified on the basis of textural appearance.

*(1) GLCM Features*

The paper proposes efficient classification of cancer cells based on famous Haralick features. These texture features based on Grey Level Co-occurrence Matrix (GLCM) is one of the most widely used techniques for texture analysis. GLCM offers a tabulation of how frequently different combinations of gray levels co-occur in an image section. GLCM texture considers the relation between two pixels at a time (reference pixel and neighboring pixel). Haralick defined several coefficients, which can be calculated from the normalized GLCM, Pi,j. Some of the main Haralick's coefficients on GLCM for texture analysis considered are: Contrast, homogeneity, energy and correlation.

*(2) Fractal Dimension*

The fractal dimension, *D*, is a statistical quantity that gives an indication of how completely a fractal appears to fill space. Box-counting mechanism is the simplest methods to calculate fractal dimension. The algorithm offers two main advantages over other methods: one, it can be applied to any type of image irrespective of the complexity and two, it is very easy to implement in case of using a computer. The following approximation depicts 'box-counting' fractal dimension derived from Hausdorff coverage dimension.

$$HD= \log(R)/\log(R(s))$$

Where, *R* is the number of squares in the superimposed grid and R*(s)* is the number of occupied squares or boxes (box count).

Perimeter roughness of the nucleus is an important measure that decides whether a particular nucleus represents a myeloblast . Fractal geometry is a more convenient way to parameterize the cell boundary surface. HD is an essential feature for fractal geometry and will be an essential quantitative measure for cell boundary roughness measurements. Higher HD signifies higher degree of roughness. Number of nuclei under field of view was much higher for a cancerous case as opposed to non cancerous case. This resulted in steep difference in box count between two cases and thereby proved to be an effective feature in whole image.

*E. Classification*

The features are extracted from the segmented images and classified using the Support Vector Machine. SVM is a classifier which constructs an *N*-dimensional hyper plane that optimally separates the data into two sets. Here linear SVM is used inorder to reduce the complexity. SVM is now considered to be one of the powerful kernel based classifier that can be adopted for resolving classification problems. An SVM kernel based algorithm builds a model for transforming a low dimension feature space in to high dimension feature space to find maximum margin between the classes. The vectors near the hyper plane are the support vectors. Extracted features are considered as input to classifier. Desired output was specified as 1 for normal and 0 for abnormal. Classification process is divided in to training and testing phase. In training phase,

ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization Volume 3, Special Issue 5, July 2014

International Conference On Innovations & Advances In Science, Engineering And Technology [IC - IASET 2014]

Organized by

Toc H Institute of Science & Technology, Arakunnam, Kerala, India during 16th - 18th July -2014

known data are given. In testing phase, unknown data are given and classification is performed using classifier after training. Accuracy of classification depends on efficiency of training. Classification of images is followed by validation.

## IV. FLICM CLUSTERING ALGORITHM

Fuzzy c-means (FCM) algorithm is one of the most widely used fuzzy clustering algorithms in image segmentation. Although the conventional FCM algorithm works well on most noise-free images, it fails to segment images corrupted by noise, outliers and other imaging artifacts. Its non-robust results are mainly because of ignoring spatial contextual information in image and the use of non-robust Euclidean distance.

To overcome the above mentioned problems, fuzzy local information c-means clustering algorithm (FLICM) is used for image clustering, which is free of any parameter selection, as well as promoting the image segmentation performance.

The major characteristic of FLICM is the use of a fuzzy local (both spatial and gray level) similarity measure, aiming to guarantee noise insensitiveness and image detail preservation. Furthermore, FLICM algorithm is fully free of the empirically adjusted parameters incorporated into all other fuzzy c-means algorithms. Also it is effective and efficient, providing robustness to noisy images**.** It can detect the clusters of an image.

The algorithm is relatively independent of the type of the added noise, and as a consequence, in the absence of prior knowledge of the noise, FLICM is the best choice for clustering. Spatial and gray level image information are combined in the algorithm; the factor $G_{ki}$ combines in a fuzzy manner the spatial and gray level information, rendering the algorithm more robust to all kind of noises, as well as to outliers. Furthermore, all the other fuzzy c-means algorithms for image clustering exploit, in their objective functions, a crucial parameter, which is used to balance the robustness and effectiveness of ignoring the added noise. This parameter is mainly determined empirically or using the trial-and-error method. The FLICM is completely free of any parameter determination, while the balance between the noise and image details is automatically achieved by the fuzzy local constraints, enhancing concurrently the clustering performance. This is also enhanced, by the fact, that almost all the other methods perform the clustering on a precomputed image, while FLICM is applied on the original image.

Thus in order to overcome the disadvantages a new factor in FCM objective function is needed. The new factor should have some special characteristics:

• to incorporate local spatial and local gray level information in a fuzzy way in order to preserve robustness and noise insensitiveness;
• to control the influence of the neighbourhood pixels depending on their distance from the central pixel;
• to use the original image avoiding preprocessing steps that could cause detail missing;
• to be free of any parameter selection. So, we introduce the novel fuzzy factor $G_{ki}$ defined as

$$G_{ki} = \sum_{\substack{j \in Ni \\ i \neq j}} \frac{1}{d_{ij} + 1} (1 - u_{kj})^m \, ||x_j - v_k||^2$$

where the ith pixel is the center of the local window (for example, 3x 3), k is the reference cluster and the jth pixel belongs in the set of the neighbours falling into a window around the ith pixel ($N_i$).$d_{i,j}$ is the spatial Euclidean distance between pixels i and j, $u_{kj}$ , is the degree of membership of the jth pixel in the kth cluster, m is the weighting exponent on each fuzzy membership, and $v_k$ is the prototype of the center of cluster .

ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization Volume 3, Special Issue 5, July 2014

International Conference On Innovations & Advances In Science, Engineering And Technology [IC - IASET 2014]

Organized by

Toc H Institute of Science & Technology, Arakunnam, Kerala, India during 16th - 18th July -2014

*General Framework of FLICM*

By using the definition of $G_{ki}$, FLICM clustering algorithm can be proposed as a robust FCM framework for image clustering. It incorporates local spatial and gray level information into its objective function, defined in terms of

$$J_m = \sum_{i=1}^{N}\sum_{k=1}^{c}[u_{ki}^m\|x_i - v_k\|^2 + G_{ki}]$$

The two necessary conditions for $J_m$ to be at its local minimal extreme, with respect to $u_{ki}$ and $v_k$ is obtained as follows:

$$u_{ki} = \frac{1}{\sum_{j=1}^{c}\left(\frac{\|x_i - v_k\|^2 + G_{ki}}{\|x_i - v_j\|^2 + G_{ji}}\right)^{1/m-1}} \qquad (1)$$

$$v_k = \frac{\sum_{i=1}^{N} u_{ki}^m x_i}{\sum_{i=1}^{N} u_{ki}^m} \qquad (2)$$

Thus, the FLICM algorithm is given as follows.

Step 1. Set the number of the cluster prototypes, fuzzification parameter m and the stopping condition €.

Step 2. Initialize randomly the fuzzy partition matrix.

Step 3. Set the loop counter b=0.

Step 4. Calculate the cluster prototypes using (2).

Step 5. Compute membership values using (1).

Step 6. Max { $U^{(b)}$- $U^{(b+1)}$ }< € then stop, otherwise, set b=b+1 and go to step 4.

When the algorithm has converged, a defuzzification process takes place in order to convert the fuzzy partition matrix U to a crisp partition. The maximum membership procedure is the most important method that has been developed to defuzzify the partition matrix U. This procedure assigns the pixel i to the class with the highest membership.

$$C_i = arg_k\{\max\{u_{ki}\}\}, k = 1,2,\ldots\ldots c$$

It is used to convert the fuzzy image achieved by the proposed algorithm to the crisp segmented image. The measure used in the FLICM objective function is still the Euclidean metric as in FCM, which is computationally simple. Moreover, differently from FCM, FLICM is robust because of the introduction of the factor $G_{ki.}$ It can preserve more image details than the other methods.

The major characteristics of the FLICM are summarized below:

- it provides noise-immunity;
- it preserves image details;
- it is free of any parameter selection;
- it is applied on the original image.

ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization Volume 3, Special Issue 5, July 2014

International Conference On Innovations & Advances In Science, Engineering And Technology [IC - IASET 2014]

Organized by

Toc H Institute of Science & Technology, Arakunnam, Kerala, India during 16th - 18th July -2014

## V. SIMULATION RESULTS

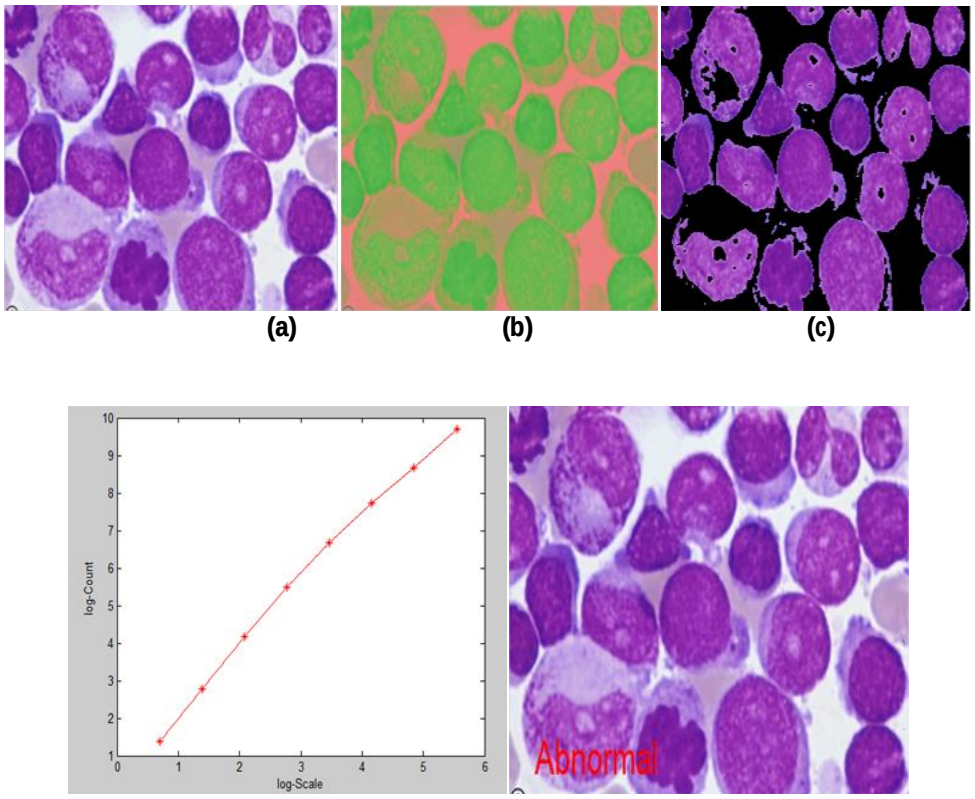


**Fig 5.1 (a) input image (b) lab image (c) segmented image (d) fractal dimension results (e) classified image**

## VI. CONCLUSION

Here a cost effective and robust automated system for leukemia segmentation and classification is presented. When preprocessing, microscopic images are converted from RGB color space to L*a*b* space so that the cell segmentation becomes more robust. Segmentation of images is done using FLICM clustering algorithm. Then different set of features such as GLCM features and fractal dimensions are extracted for a data base of images. Finally run the classifier system and validate the output based on the results obtained.

ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization Volume 3, Special Issue 5, July 2014

International Conference On Innovations & Advances In Science, Engineering And Technology [IC - IASET 2014]

Organized by

Toc H Institute of Science & Technology, Arakunnam, Kerala, India during 16th - 18th July -2014

## REFERENCES

[1] Monica Madhukar, Sos Agaian, AnthonyT.Chronopoulos, "Deterministic Model for Acute Myelogenous Leukemia Classification" 2012 IEEE International Conference on Systems, Man, and Cybernetics October 14-17, 2012.
[2] Fabio Scotti, "Automatic Morphological Analysis for Acute Leukemia Identification in Peripheral Blood Microscope Images", *CIMSA july 2005 IEEE International Conference.*
[3] Fabio Scotti, "Robust Segmentation and Measurements Techniques of White Cells in Blood Microscope Images", *IMTC 2006 - Instrumentation and Measurement Technology Conference Sorrento, Italy 24–27 April 2006.*
[4] Nor Hazlyna Harun, MY Mashor, NR.Mokhtar, Aimi Salihah, A.N, Rosline Hassan, R.A.A. Raof, MK Osman, "Comparison of Acute Leukemia Image Segmentation using HSI and RGB Color Space", *10th International Conference on Information Science, Signal Processing and their Applications (ISSPA 2010) IEEE.*
[5] Subrajeet Mohapatra, Dipti Patra, "Automated Leukemia Detection using Hausdorff Dimension in Blood Microscopic Images", *2010 IEEE.*
[6] Subrajeet Mohapatra, Dipti Patra, Kundan Kumar, "Blood Microscopic Image Segmentation using Rough Sets", *2011 International Conference on Image Information Processing (ICIIP 2011).*
[7] S.Rubhala, Mrs.J.Ramya , "Image segmentation of leukemia using kernel based fuzzy c means clustering" , *International Journal of Communications and Engineering Volume 03– No.3, Issue: 03 March2012.*
[8] Ms. Minal D. Joshi and Prof. A.H.Karode , "Detection of acute leukemia using white blood cells segmentation based on blood samples", *IJECET Volume 4, May June, 2013.*
[9] Stanislaw Osowski, Tomasz Markiewicz, Bozena Marianska, Leszek Moszczyński "Feature generation for the cell image recognition of myelogenous leukemia".
[10] Stelios Krinidis and Vassilios Chatzis, "A Robust Fuzzy Local Information C-Means Clustering Algorithm", *IEEE Transactions On Image Processing, Vol. 19, No. 5, May 2010.*