I J C T A, 9(26) 2016, pp. 335-342

A Proposed Hybrid Fuzzy *C*-means Algorithm With Cluster Center Estimation For Leukemia Image Segmentation

A.R. Jasmine Begum[1] and T. Abdul Razak[2]

ABSTRACT

Fuzzy *C*-means clustering (FCM) is an important technique used in cluster analysis. The standard FCM algorithm calls the centroids to be randomly initialized resulting in the requirement of making estimations from expert users to determine the number of clusters. To overcome these observed limitations of applying the FCM algorithm, an efficient image segmentation model, Hybrid Fuzzy *C*-means Algorithm with Cluster Center Estimation (HFCMCCE) using subtractive clustering for Leukemia infected blood sample Image Segmentation is presented in this paper. In this algorithm, the image is initially subjected to Partial Contrast Stretching (PCS) method to modify the force level of the dark scale image, then Subtractive Clustering Method (SCM) is applied to determine the thickness measure of the pixel and the pixel with the highest thickness measure is marked as the first cluster center. These cluster centers are used in FCM to perform the image segmentation to obtain the required segmented image. The performance of this algorithm is then compared with other existing algorithms, on the basis of image quality measures such as PSNR (Peak-Signal-to-Noise ratio), MSE (Mean Square Error).

Keywords: Fuzzy *C*-means, Classification, Segmentation, Subtractive Clustering, Peak-Signal-to-Noise ratio, Mean Square Error.

1. INTRODUCTION

Leukemia is a type of blood cancer that occurs in white blood cells (WBCs) created from the bone marrow. It interrupts the balance of the blood system which is analyzed through qualified specialists using exclusive laboratories. In all such assessment cases, it is reported that 50% of patients are not properly diagnosed of its subtypes. The issue of improper diagnosis emerges because of the impersonation of comparable signs in different issues [1] and the complex nature of blood smear images. Thus the selection in slide arrangement methods needs to be improved to meet genuine clinical requests. The intense leukemia segmentation and classification method depends on four primary categories, such as threshold, boundary, region and hybrid. Most of the procedures combines boundary and region criteria [1] [2]. The Otsu and histogram , commonly used threshold based strategy, segments the WBCs basically from the blood smear image utilizing the intensity level[3][4].

Form based techniques recognizes the anomalies of the nucleus boundary consolidated with particular filtering segment leukocytes from other blood components [5] [6]. The area and hybrid segmentation depends on multi decision analysis extract the predictable maximization of hue, infiltration, and value color space to recognize the cytoplasm and nucleus of the WBCs [7]. Generally color image segmentation is more desired, while compared with gray level images. Watershed segmentation algorithm and unsupervised color segmentation provides good level segmentation of nuclei to recognize acute leukemia [8].

[1] Asst.Professor, Department of Information Technology, Cauvery College for Women, Trichy-18. *E-mail: jasmine_akram2003@yahoo.com*

[2] Associate Professor, Department of computer Science, Jamal Mohamed College, Trichy-20. *E-mail:* abdul64@yahoo.com

2. RELATED WORKS

A number of researchers utilized various segmentation strategies like thresholding strategy, region based methodologies, edge detection approach, clustering approaches, artificial neural network, Fuzzy procedure, watershed algorithms and others for the segmentation of leukemia images in an attempt to achieve better results. S. Rubhala *et. al.* [9] presented a work to classify a lymphocyte as a typical or a lymphoblast, where only acute lymphocytic leukemia is considered. The paper showed that a quick and successful segmentation methodology for blow up images is extremely useful for enhancing the hematological technique and quickening in finding of leukemia illnesses. A kernel-induced new metric was utilized to supplant the Euclidean standard in Fuzzy *C*-means algorithm in the original space and then determined the option of kernel-based Fuzzy *C*-means algorithm. Features are segmented in the output and the best features are analyzed. SVM (Support Vector Machine) is employed for classification. N. H. Harun *et. al.* [10] proposed three clustering algorithms such as *k*-means, fuzzy *c*-means and moving *k*-means algorithms have been applied on the saturation component image. Then the median filter and seeded region growing area extraction algorithms have been applied. The Comparisons among the three clustering algorithms are made in order to measure the performance of each clustering algorithm on segmenting the blast area and moving *k*-means clustering algorithm has successfully produced the fully segmented blast region in Acute Leukemia image.

3. CLUSTERING ALGORITHM FOR IMAGE SEGMENTATION

A. Partial Contrast Stretching Technique

Contrast Stretching is one of the image enhancement techniques that attempts to improve the contrast in an image by stretching the range of the intensity values it contains to span a desired range of values. Partial contrast is a linear mapping function that is used to increase the contrast level and brightness level of the image.

B. Fuzzy *C*-means Clustering

Fuzzy clustering algorithms treat clusters as flexible gatherings to which each data object has a membership degree [12]. These degrees are esteemed somewhere around 0 and 1, with a high esteem representing to a high similitude between the object and the group (Bezdek, *et. al.*, 1984). The most well-known of these techniques is Fuzzy *c*-means.

Fuzzy *C*-means Clustering is separate from *K*-means that utilizes hard AP segmenting. FCM permits pixels to have a place with various cluster with alterable degrees of participation. In light of its extra adaptability, FCM is also called as Soft clustering strategy. FCM segments an accumulation of n vector x_i, $i = 1, n$ into c Fuzzy gatherings, and finds a cluster focus in every gathering such that a cost capacity of divergence measure is minimized. FCM utilizes Fuzzy AP segmenting such that given information point can have a place with a few gatherings with the level of belongingness determined by enrollment grades somewhere around 0 and 1. To oblige the presentation of Fuzzy AP segmenting, the enrolment network U is permitted to have components with qualities somewhere around 0 and 1. Forcing standardization stipulates that the summation of degrees of affiliation for information set dependably be equivalent to solidarity.

$$\sum_{j=1}^{c} u_{ij} = 1, \ \forall j = 1 \ldots n \tag{1}$$

The cost capacity (or target capacity) for FCM

$$J(U, c_1, \ldots c_c) = \sum_{i=1}^{c} J_i = \sum_{i=1}^{c} J_i \sum_{j}^{n} u_{ij}^{m} d_{ij}^{2} \tag{2}$$

Where U_{ij} is somewhere around 0 and 1; *ci* is the cluster focal point of Fuzzy gathering *i*;

$$dij = \|ci - xj\| \tag{3}$$

Is the Euclidean separation between i[th] cluster focus and j[th] point; and mC[1, ∞] is a weighting type.

$$J(U, c_1, c_c, \lambda_1 \ldots \lambda_n) = J(U, c_1 \ldots c_c) = \sum_{j=1}^{n} \lambda_j \left(\sum_{i=1}^{c} u_{ij} - 1 \right) \tag{4}$$

$$= \sum_{i=1}^{c} \sum_{j}^{n} u_{ij}^m u_{ij}^2 + \sum_{i=1}^{n} \lambda_j \left(\sum_{i=1}^{c} u_{ij} - 1 \right) \tag{5}$$

where λ_j, $j = 1$ to n, are the Lagrange multipliers for the n requirements in Equation (2). By separating $J(U, c_1\, c_c, \lambda_1 \ldots \lambda_n)$ to all its inputs contentions, the essential conditions for Equation (2) to achieve its base are

$$c_i = \frac{\sum_{j=1}^{n} u_{ij}^m x_j}{\sum_{j=1}^{n} u_{ij}^m} \tag{6}$$

$$u_{ij} = \frac{1}{\sum_{k=1}^{c} \left[\dfrac{d_{ij}}{d_{kj}} \right]^{2/(m-1)}} \tag{7}$$

Step 1: Initialize the enrollment lattice U with irregular qualities somewhere around 0 and 1 such that the imperatives in Equation (1) are fulfilled.

Step 2: Calculates c Fuzzy cluster focus c_i, $i = 1$, c, utilizing Equation (5).

Step 3: Compute the cost capacity (or complaint capacity) by Equation (2). Stop if it is possible that it is underneath a specific resistance quality or its change over past emphasis is beneath a specific limit.

Step 4: Compute another U by Equation (6). Go to step 2.

C. Subtractive Clustering

In subtractive clustering, information focuses are considered as the possibility for cluster focuses. By utilizing this technique, the calculation is essentially corresponding to the quantity of information focuses and free of the measurement issue as specified. Consider an accumulation of n information focuses $\{x_1, \ldots, a_n\}$ in a M-dimensional space. The information focuses are accepted to have been standardized inside a hypercube. Since every information point is a possibility for cluster focuses, a thickness measure at information point X_i is characterized as

$$D_i = \sum_{j=1}^{n} \exp\left[-\frac{\|x_i - x_j\|^2}{(r_a/2)^2} \right] \tag{8}$$

Where the image $\|\cdot\|$ means the Euclidean separation, r_a is a positive steady. In this manner, an information point will have a high thickness esteem in the event that it has numerous neighboring information focuses. The sweep characterizes an area; information focuses outside this span contribute just somewhat to the thickness measure.

$$D_i = D_i - D_{c_1} \exp\left[-\frac{\|x_i - x_j\|^2}{(r_a/2)^2} \right] \tag{9}$$

Where is a positive consistent. In this way, the information focuses close to the principal cluster focus X_{c_1} will have fundamentally diminished thickness measure, consequently making the focuses unrealistic to be chosen as the following cluster focus. The consistent r_b characterizes an area that has quantifiable decreases in thickness measure. The steady r_a is regularly bigger than to forestall firmly separated cluster focuses; Where ($r_a = \beta * r_a$). β is a parameter called as squash component, for the most part r_b is equivalent to $1.5r_a$, which is duplicate by span qualities to decide the neighboring clusters inside which the presence of other cluster focuses are disheartened. r_b parameter utilized for determination of the quantity of clusters. After the thickness measure for every information point is changed, the following cluster focus X_{c_2} is chosen and the greater part of the thickness measures for information focuses are updated once more. This procedure is rehashed until an adequate number of cluster focuses are produced.

4. PROPOSED HYBRID FUZZY *C*-MEANS ALGORITHM WITH CLUSTER CENTER ESTIMATION (HFCMCCE)

A. Proposed Model for Image Segmentation

At first the fractional complexity extension is connected for extending the pixel power values and accordingly a high differentiation image can be obtained. In order to introduce the cluster focus, the point with the most elevated thickness measure is picked as the principal cluster focus by applying SCM on the grounds that the nature of the clustering depends firmly on the decision of its beginning cluster focus and their areas in FCM. Presently, utilizing the FCM calculation of the image is divided into a predetermined number of clusters. The following stride is to improve the image by evacuating the commotion by utilizing middle channel. The Proposed Algorithm is given below:

Input: Gray Scale Image

Step 1: Input the dark scale image to the framework

Step 2: Partial Contrast Stretching to alter the force level of the dark scale image.

Step 3: Initialize cluster *i.e.* k = number of clusters

Step 4: The thickness measure esteem for the each pixel in the image is assessed by the condition.

$$D_i = \sum_{j=1}^{n} \exp\left[-\frac{\|X_i - x_j\|^2}{(r_a/2)^2} \right]$$

Step 5: Select the pixel with the most astounding thickness measure as the main cluster focus and its relating thickness measure esteem as the greatest thickness measure.

Step 6: The thickness measure for every pixel worth is reconsidered by utilizing the condition

$$D = D_i - D_{c_1} \exp\left[-\frac{\|X_i - x_j\|^2}{(r_a/2)^2} \right]$$

Step 7: After the thickness measure for every pixel worth is overhauled, the following cluster focus with the most noteworthy thickness measure is chosen as a second cluster focus.

Step 8: This procedure is rehashed until a k number of cluster focuses are created.

Step 9: Initialize k cluster focuses in the FCM with cluster focus created in the progression 8.

Step 10: Calculate the Euclidean separation between each pixel of the dark scale image and cluster focus by utilizing the equation

$$d_{ij} = \|ci - xj\|$$

Step 11: Create participation grid by taking the fragmentary separations from the point to the cluster focus by utilizing the equation

$$u_{ij} = \frac{1}{\sum_{k=1}^{c}\left[\frac{d_{ij}}{d_{kj}}\right]^{2/(m-1)}}$$

Step 12: Generate another Fuzzy cluster focus with the condition

$$c_i = \frac{\sum_{j=1}^{n} u_{ij}^{m} x_j}{\sum_{j=1}^{n} u_{ij}^{m}}$$

Step 13: Compute the cost capacity as indicated by Equation $J(U, c_1, ...c_c) = \sum_{i=1}^{c} J_i = \sum_{i=1}^{c} J_i \sum_{j}^{n} u_{ij}^{m} d_{ij}^{2}$ Stop if it is possible that it is beneath specific resistance esteem.

Step 14: Compute another *U* participation grid by Equation $u_{ij} = \frac{1}{\sum_{k=1}^{c}\left[\frac{d_{ij}}{d_{kj}}\right]^{2/(m-1)}}$ Go to step12.

Step 15: Transform the resultant cluster into image.

Output: The resultant image of Leukemia influenced cells.

The figure 1 represents to the framework of proposed hybrid Fuzzy *C*-means clustering with Cluster Center Estimation for Leukemia image segmentation. In this figure, the distinctive colours represent to the different operations of the Leukemia image. Violet colour represents to the information image. Yellow color is for applying filers and believer that information image to lab color. Red colour represents the proposed hybrid Fuzzy *C*-means calculation with Cluster Center. Blue colour indicates the segmentation of the clustered image after applying the proposed hybrid model.

5. SIMULATION RESULTS AND DISCUSSIONS

The proposed Hybrid Fuzzy *C*-means Algorithm with Cluster Center Estimation is actualized by using Image Processing Tools of MATLAB R 2013a. The corresponding set of figures represents the outcomes acquired from the proposed structure calculation. The sample images are collected from hospital labs and also from the dataset [13].

Initially, the Partial Contrast Stretching (PCS) is applied to adjust the force level of dark scale image then the Subtractive Clustering Method (SCM) is applied to acquire prior knowledge on the cluster centers. These cluster centers are used in FCM to perform the image segmentation which generates the resultant image of Leukemia influenced cells. From the Figure 3 and Table 1 it is observed that the performance of the proposed hybrid clustering method (HFCMCCE) for Leukemia blood sample is showing high PSNR (Peak Signal Noise Ratio) and less Mean Squared Error (MSE) than the input image and standard FCM applied image.

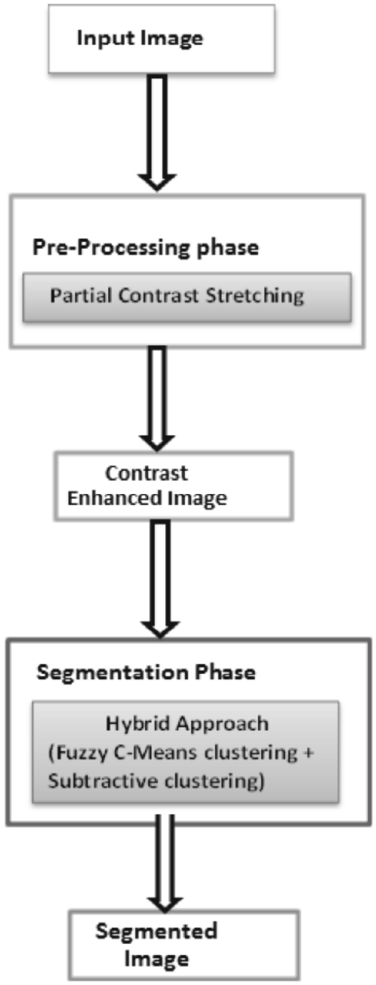
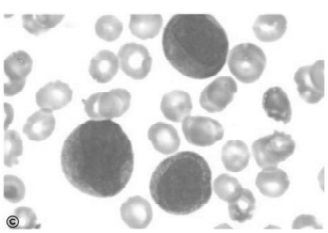
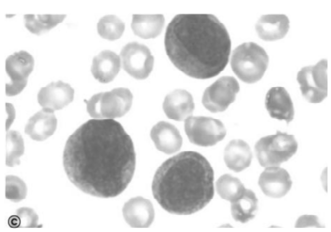


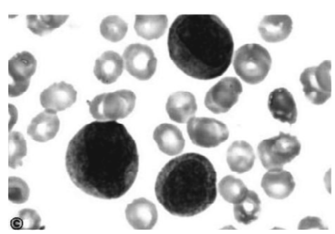
Figure 1: A Hybrid Fuzzy *C*-means Algorithm with Cluster Center Estimation Framework



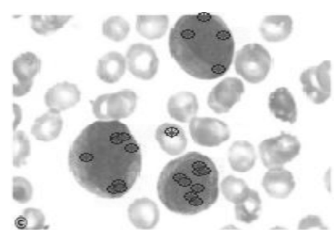
(a) Read and display an RGB image



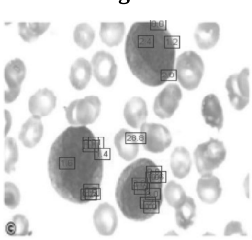
(b) Converting in to gray scale



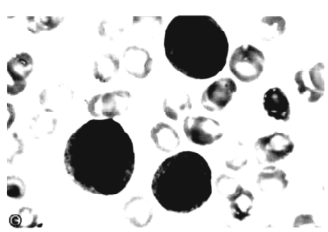
(c) Partial contrast stretching



(d) Thickness measure value of each pixel



(e) Thickness measure value of each pixel



(g) FCM

(h) Proposed HFCMCCE

Figure 2: Represents (an) Input image, (b) RGB to Gray Scale Conversion, (c) Result image from Partial Contrast Stretching (d) Thickness measure estimation of each pixel, (e) Thickness measure estimation of each pixel, (f) Density measure esteem, (g) FCM, (h) Proposed HFCMCCE

Table 1
Comparison of the error rate of the Input Image, FCM applied Image and Proposed HFCMCCE (Ideal values: PSNR-High, MSE-Low)

TABLE-I COMPARISON OF PSNR MSE VALUES FOR INPUT,FCM AND HFCMCCE IMAGES

File Edit View Insert Tools Desktop Window Help

	PSNR(dB)	MSE
Input Image	30.6300	21.23
FCM Image	26.4200	29.78
Proposed Hybrid Model Image	34.0800	20.73

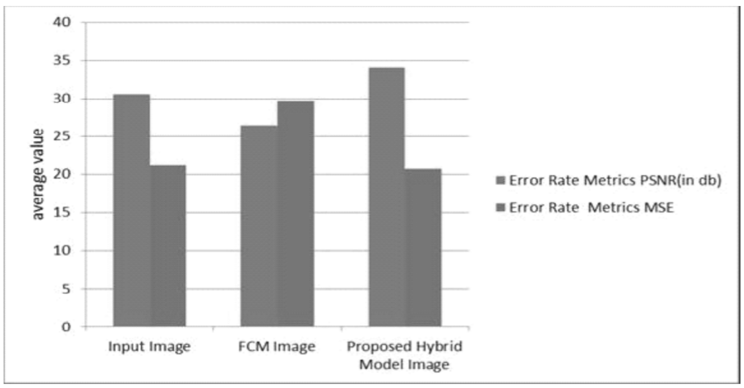


Figure 3: Graphical representation of comparison of the error rate of the Input Image, FCM applied Image and Proposed HFCMCCE

CONCLUSION

Exploratory results demonstrates better execution of segmentation images utilizing the proposed Hybrid Fuzzy *C*-means with Cluster Center Estimation, which hybridizes the Fuzzy *C*-means clustering and Subtractive clustering. This proposed structure is utilized as a programmed framework for the segmentation of Leukemia infected blood cells when the leukemia influenced cell images are given as the info image. The proposed hybrid model is well suited for image with noise density up to 90%; in addition it reduces the Mean Squared Error in the proposed model. The PSNR value is increased for HFCMCCE applied image than the input image and FCM applied image. In future, HFCMCCE can be further improvised by combining it with morphological operations in order to improve the quality of the segmented output image; also the classification of the different types of Leukemia can be carried out using SVM.

REFERENCES

[1] M. Subrajeet, D. Patra, and S. Satpathy, “Automated Leukemia Detection in Blood Microscopic Images using Statistical Texture Analysis”, *International Conference on Communication ,Computing and Security (ICCCS’11)*, 184-187, 2011.

[2] P.Viswanathan, and P. Venkata Krishna, “Morlet Wavelet Fingerprint Invariant Automated Authentication System”, *International Journal on Recent Trends in Engineering and Technology*, **4(1)**, 1-5, 2010.

[3] S. Agaian, K. Panetta, S. Nercessian, and E. Danahy, “Boolean Derivatives with Application to Edge Detection for Imaging Systems”, *IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)*, **40(2)**, 371-382, 2010.

[4] A. Pentland, “Fractal Based Description of Natural Scenes”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-**6(6)**, 661-674, 1984.

[5] S. Mohapatra, D. Patra, and S. Satpathi, “Automated Cell Nucleus Segmentation and Acute Leukemia Detection in Blood Microscopic Images”, *International Conference on Systems in Medicine and Biology (ICSMB)*, 49-54, 2010.

[6] S. Mohapatra, D. Patra, and S. Satpathi, "Image Analysis of Blood Microscopic Images for Acute Leukemia Detection", *International Conference on Industrial Electronics, Control and Robotics (IECR)*, 215-219, 2010.

[7] P.Viswanathan, and P.Venkata Krishna, "Multimodal Biometric Invariant Moment Fusion Authentication System, Business Administration and Information processing", *Communications in Computer and Information Science, Springer verlag,***70(3)**, 136-144, 2010.

[8] C. E. Pedreira, L. Marcini, M. G. Land, and E. S. Costa, "New Decision Support Tool for Treatment Intensity Choice in Childhood Acute Lymphoblastic Leukemia", *IEEE Transactions On Information Technology and Biomedicine,***13(3)**, 284-290,2009.

[9] S.Rubhala, Mrs. J.Ramya, "Image Segmentation of Leukemia using Kernel Based Fuzzy *C*-Means Clustering" , *International Journal of Communications and Engineering*, 03,2012.

[10] N. H. Harun, A. S. Abdul Nasir, M. Y. Mashor, R. Hassan "Unsupervised Segmentation Technique for Acute Leukemia Cells Using Clustering Algorithms", *World Academy of Science, Engineering and Technology, International Journal of Computer, Control, Quantum and Information Engineering*, **9(1)**, 2015.

[11] Savita I Basanagoudar, Chidanandamurthy M V, M Z Kurian, "Medical Image Enhancement using Multi Scale Retinex Algorithm with Gaussian and Laplacian Surround Functions", *International Journal of Advance Electrical and Electronics Engineering (IJAEEE),* **4(1)**, 2015.

[12] Adnan Khashman, Esam Al-Zgoul,"Image Segmentation of Blood Cells in Leukemia Patients", *International Conference on Computer Engineering and Applications (ICCEA'10),* 104-109, 2010.

[13] Dataset Source: http://www.hematologyatlas.com/leukemias.htm.