International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878, Volume-8, Issue-2S11, September 2019

Segmentation of Blood cell Images Using Hybrid K-means with Cluster Center Estimation Technique

A.A. Mariena, J.G.R. Sathiaseelan

Abstract--- ***Image segmentation plays a predominant role in the field of image processing. k- Means clustering is one of the most powerful algorithms for medical image segmentation. However, the randomly selected cluster number and initial centroids cause inconsistency in the image segmentation results. To overcome this limitation we have proposed a combined approach namely Hybrid K-Means with Cluster Center Estimation (HKMCCE) for image segmentation. The proposed technique use histogram peaks of the image to find the cluster number and initial cluster centers automatically.Moreover, it requires lessuser interaction to determine k-means initialization parameters. The performance of the proposed technique is compared with traditional segmentationmethods and it yields better results with less execution time.***

Keywords---K-Means,HKMCCE, Histogram.

I. INTRODUCTION

The microscopic images of blood cells are observed to diagnose various diseases. Variations in the blood condition reveals the development of diseases in an individual. Leukemia is a group of haematological disease that affects blood, bone marrow, lymph nodes which characterized by overproduction of abnormal white cells. Leukaemia is the leading cause of cancer death more particularly in India.Based on the cancer statistics 2019 by American Cancer Society it is found that leukaemia is the fifth cause of death in men and sixth cause of death in women.Careful microscopic examination of stained blood smear is the only way for the effective diagnosis of Leukemia.

Early detection of the disease is essential for the proper treatment.Leukaemia is detected only by extracting and analysing the WBC from other blood cell components such as Red Blood Cells (erythrocytes), platelets and plasma and it is one of the prior method to analyse the image accurately in a decision oriented application. Segmentation of WBC from blood smear is a tedious task however there exists a plethora of techniques for image segmentation. Moreover, the existing techniquesconsumes more time for execution and also less performance in terms of accuracy. Clustering techniques plays a vital role in the medical image segmentation. Clustering is an unsupervised learning problem through which meaningful and useful objects were grouped together based on some similarity measure. K-Means clustering algorithm is widely used in image segmentation due to its computational simplicity. The main problem associated with the k-means algorithm is the selection of cluster number and the initial centroids. To overcome this limitation we have proposed a combined approach namely Hybrid K-Means with Cluster Center Estimation (HKMCCE).Here we have used three distance measures like Euclidean, Manhattan and Chi Square distance for comparison and this is the novelty of HKMCCE.

The rest of the paper has been organized in the following manner; Literature survey explained in the next section followed by methodology as third section. Fourth section includes performance analysis of the results and conclusions are described in the last section.

II. LITERATURE SURVEY

Pallavi et al.[10] introduced an enhanced k-means technique for generating the cluster center by reducing the mean square error of the final cluster. The algorithm works better for dense dataset yet, it results in less accuracy for sparse dataset. Soumen et al. [2] proposed a thresholding estimation based watershed transforming technique for blood cell detection. Initially the image is converted in to Conversion of RGB to grayscale image then apply the sobel filter in frequency domain to detect the boundary. Finally iterative threshold is applied on the watershed transforming image for segmentation.

Algorithm has used 30 numbers of blood microscopic images as test images and obtained higher accuracy results of around 93%.Still it lacks in accuracy for higher number of datasets. Madhu et al.[12] proposed an adaptive k-means clustering technique with improved initial cluster center. This technique acquires good accuracy with less computation time still, it initializes the cluster number randomly. Dhanachandra et al. [6] devised a subtractive clustering algorithm to generate initial centroids based on the potential value of the image. The resultant image having good quality however its computation complexity grows exponentially according to dimensionality of data.

Intedhar et al. [5] designed a hybrid approach by integrating region growing and k-means method to segment the image. The first approach is used to get the number of objects from the image and second approach is used to cluster the image according to the input given from first approach. This technique is not suited for complicated images moreover, it will not work efficiently for non-uniform texture images. Saranya et al. [13]developed an adaptive k-means clustering approach for clustering high dimensionality data and it gives better quality results.

Manuscript received September 16, 2019.

A.A. Mariena, Research Scholar, Dept. of Computer Science, Bishop Heber College, Trichy. T.N, India.

J.G.R. Sathiaseelan, Asso. Professor and Head, Dept. of Computer Science, Bishop Heber College, Trichy. T.N, India.

Retrieval Number: B10260982S1119/2019©BEIESP
DOI: 10.35940/ijrte.B1026.0982S1119

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Moreover the initial cluster number calculation yet to be implemented adaptively. Elindra et al. [14] proposed a hybrid approach of k-means and PSO for image segmentation. PSO is used to generate the best initial cluster centroids. Proposed technique used three distance measures to compare the performance like execution time and RMSE. Euclidean based k-means-PSO consumes more processing time compared to other distance measures. The above mentioned techniques have high computational complexity and it consumes more time for processing.

III. METHODOLOGY

K-means clustering is an unsupervised image segmentation technique that has no prior knowledge about the type of image to be segmented. So the selection of these initialization parameters affects the accuracy of the final results. To deal with this problem, a combined approach namely Hybrid K-Means with Cluster Center Estimation (HKMCCE) technique is proposed in this paper.

The proposed algorithm use histogram peaks of an image to find the cluster number and initial cluster centers automatically.

The cluster number and the cluster centers obtained from the above technique will be given to the k-means algorithm, which assigns each pixel to the nearest cluster center and all the intensity levels of the pixels within the cluster will be replaced by the intensity level of the respective cluster center. The work flow of theHybrid K-Means with Cluster Center Estimation (HKMCCE)is shown in figure1.

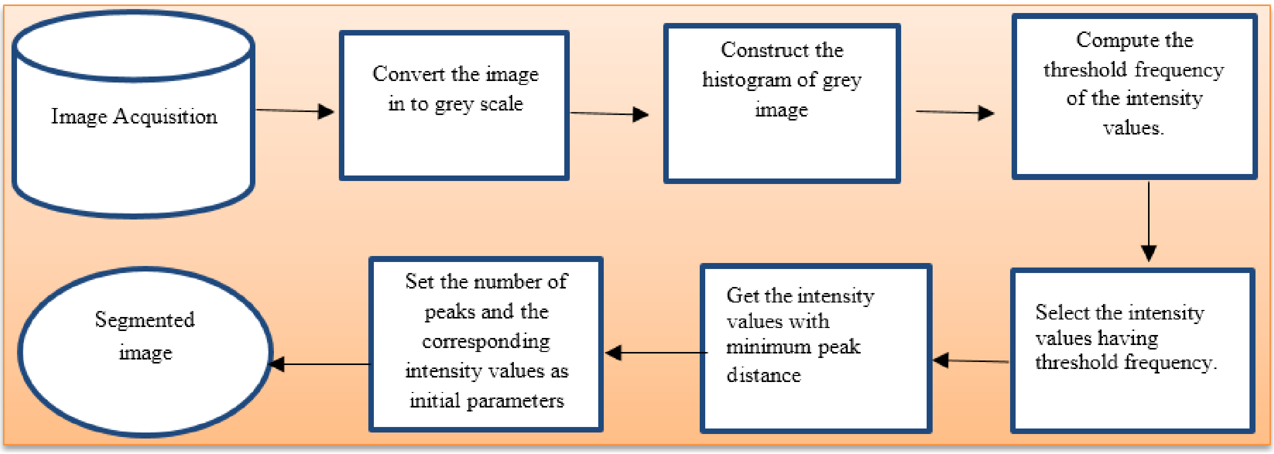

Fig. 1: HKMCCE Framework

Hybrid K-Means with Cluster Center Estimation(HKMCCE)

In the grey image histogram the x-axis represents grey levels and y-axis represents the frequency of grey levels in the image. The grey level intensities lies in between 1 and 256. Initially compute the threshold frequency by dividing the total frequency by number of intensity values.

The grey levels having a frequency greater than the threshold frequencyand their immediate neighbouring grey levels, are processed to obtain possible peak grey levels. The selected peak gray levels from the above steps are again scanned to get the real peaks.

The peak grey levels having minimum peak distance are selected as real peaks since the noisy peaks are having higher frequency, and they are close to each other in the x-direction. The minimum peak distance is computed by dividing the total number of grey levels having threshold frequencyby the constant value β whose value can be selected within the range [1,N]where N represents the total number of grey levels in the histogram having threshold frequency. In the proposed approach β is set to a value 2 for all the grey images.

Total number of significant peaks and the corresponding intensity values serve as the number of clusters and the initial cluster centersrespectively for the k-means method. To get the segmented image k-means technique is applied on the image using the initial centroids and cluster number. In the subsequent steps new cluster centers are formed by computing the mean of all the pixels within a cluster. This procedure continues until convergence occurs.

Proposed HKMCCE Algorithm

Input: Blood Smear Image
Output: Segmented Image

Step 1: Read the Blood smear image.
Step 2: Convert the color image into its corresponding grey image.
Step 3: Resize the image into 255x 255.
Step 4: Construct the histogram of the image.
Step 5: Compute the threshold frequency by finding the sum of all frequencies divided by total number of intensity values
Step 6: **for** each intensity value having threshold frequency
Step 7: Check whether the frequency of the intensity value is greater than the previous one and less than the next one.
Step 8: **If** yes increment the number of peaks
Step 9: **end if**
Step 10:**end for**
Step 11:Compute the minimum peak distance by dividing the number of peaks by beta where beta=2.
Step 12: **for** each peak
Step 13:Check whether the distance between the current peak and the next peak is greater than or equal to the minimum peak distance.
Step 14: **If** yes increment the number of significant peaks
Step 15: **end if**
Step 16: **end for**
Step 17: Set the number of significant peaks as the cluster number and its corresponding intensity values as initial centroids.
Step 18: Input the cluster number and initial centroids to K-Means clustering technique.
Step 19: Final output is the segmented Image.

Fig. 2: Proposed Work

IV. PERFORMANCE EVALUATION AND EXPERIMENTAL RESULTS

To evaluate the proposed approach we have taken the benchmark blood cell dataset namely ALL-IDB [15]. It is a public dataset developed for the evaluation of blood cell segmentation and classification that contains 108 blood cell images in JPG format with 24 bit color depth and resolution 2592 x 1944. The quantitative metrics such as Rand Index (RI), Variation of Information (VOI), Global Consistency Error (GCE) and Execution time were used for evaluating the performance of proposed segmentation. Comparison of the blood smear images with standard segmentation algorithms are shown in Fig. 3, 4, and 5 where Figure (3a) is the Original blood smear image, (3b) is the segmented image obtained by using k-means (3c) is the result of Region Growing, (3d) is output of Watershed and (3e) is the result of HKMCCE Technique. Similarly figure 4, and figure 5 are the results for other blood smear images respectively. From the results it is observed that the proposed HKMCCE algorithm shows better performance than other standard segmentation algorithms.

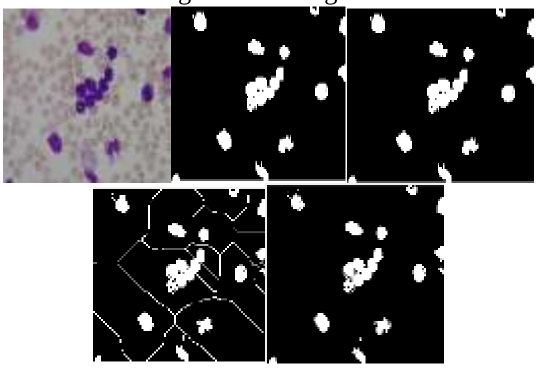

Fig. 3: (a) BSI 1) (b) K-means (c) Region Growing (d) Watershed (e) HKMCCE

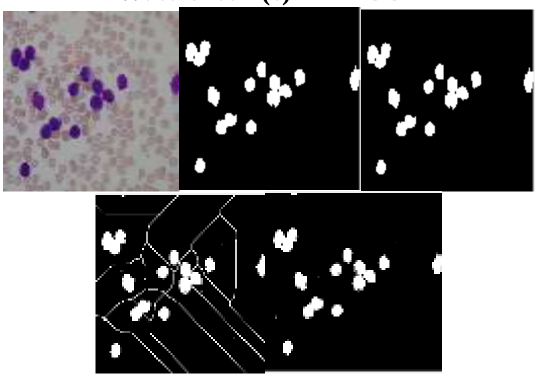

Fig. 4: (a) BSI 2 (b) K-means (c) Region Growing (d) Watershed (e) HKMCCE

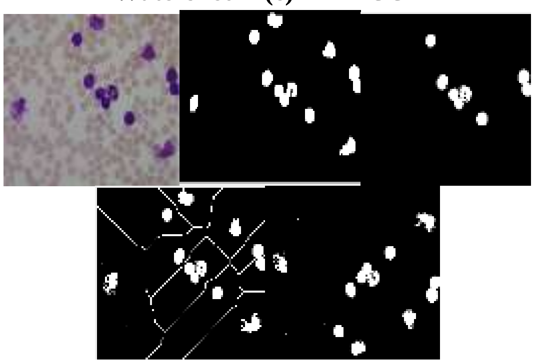

Fig. 5: (a)BSI 3 (b) K-means (c) Region Growing (d) Watershed (e) HKMCCE

Table 1: Rand Index

Techniques	BSI1	BSI2	BSI3
K-means	.9071	.8110	.7391
Region Growing	.8511	.8273	.7562
Watershed	.9122	.7968	.7465
HKMCCE	.9281	.8461	.7866

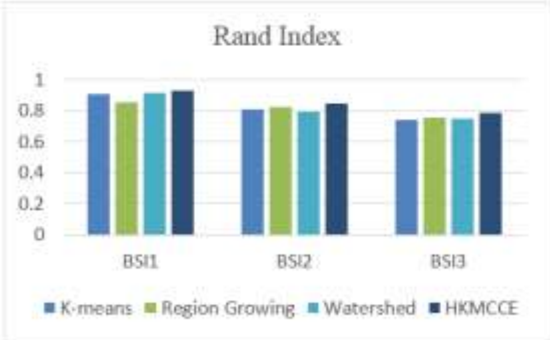

Fig. 6: Performance Analysis of RI

Table 2: Variation of Information

Techniques	BSI1	BSI2	BSI3
K-means	.3940	.6382	.7796
Region Growing	.3886	.6803	.4857
Watershed	.3489	.7967	.3889
HKMCCE	.3641	.6191	.3079

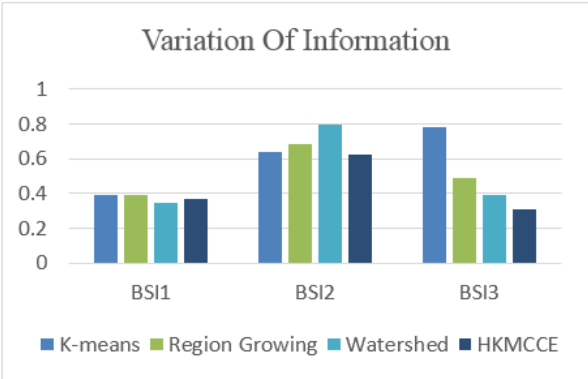

Fig. 7: Performance Analysis of VOI

Table 3: Global Consistency Error

Techniques	BSI1	BSI2	BSI3
K-means	.0584	.0785	.0791
Region Growing	.0672	.0819	.8246
Watershed	.0531	.0712	.0797
HKMCCE	.0468	.0603	.0743

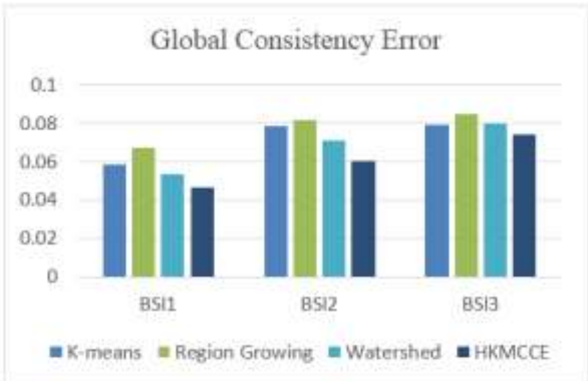

Fig.8: Performance Analysis of GCE

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Table 4: Execution Time

Techniques	BSI1	BSI2	BSI3
K-means	10.305	6.903	8.158
Region Growing	7.134	8.132	5.281
Watershed	6.721	5.980	4.870
HKMCCE	6.538	5.220	4.116

Fig. 9: Performance Analysis of Execution Time

Table 5: Performance Analysis of various distance measures

Techniques	BSI1	BSI2	BSI3
Euclidean HKMCCE	5.8292	5.919	5.9479
Manhattan HKMCCE	5.2411	5.8767	5.8114
Chi-square HKMCCE	9.2958	10.9396	10.1618

Fig. 10: Performance Analysis of various distance measures

V. CONCLUSION

In this paper, a combined approach namely Hybrid K-Means with Cluster Center Estimation (HKMCCE) technique has been proposed, to overcome the drawbacks of k-means algorithm in image segmentation scenario. First of all, computation of initial parameters has been done by taking the real peaks of the grey level histogram and the respective intensity values. Those values were fed in to the k-means technique to get the segmented image. From the experimental results it is observed that the Rand index improved from 81% to 85% with the gain of 4%.The global consistency error decreased from 7% to 6% with the gain of 1%. Variation of information has been reduced from 54% to 43% with the gain of 11% and the execution time decreased from 7.05 seconds to 5.29 seconds with the gain of 1.76 seconds. The results endorsed the performance of the proposed HKMCCE method. The proposed method implemented and compared with various distance measures. Manhattan HKMCCE consumes less execution time compared to other two distance measures and Euclidean HKMCCE acquired high accuracy. For the future research, segmentation can be extended for the color images too.

REFERENCES

1. Soudani, A., &Zagrouba, E. (2017). Adaptive Region Based Active Contour Model for Image Segmentation. 2017 *IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA)*.
2. Soumen Biswas and Dibyendu Ghoshal, "Blood Cell Detection using Thresholding Estimation Based Watershed Transformation with Sobel Filter in Frequency Domain" Twelfth *International Multi-Conference on Information Processing*-2016 (*IMCIP*-2016, *Procedia Computer Science* 89 (2016) 651 – 657.
3. Rejula, M. A., &Dr.Jeya Kumar, M. K. (2016). An Efficient Approach of Segmentation on White Blood Cells Using Morphological Operators and 2D-Discrete Wavelet Transform with Anfis Classification. *International Journal of Emerging Research in Management &Technology*, 5*(8)*, 105-115.
4. Chen, Z., Qi, Z., Meng, F., Cui, L., & Shi, Y. (2015). Image Segmentation via Improving Clustering Algorithms with Density and Distance. *Procedia Computer Science, 55*, 1015-1022.
5. Intedharshakirnasir, "the proposed image segmentation method Based on adaptive k-means algorithm", *Journal of Theoretical and Applied Information Technology*, 15th November 2018. Vol.96. No 21.
6. Nameirakpam Dhanachandra, Khumanthem Manglem &Yambemjina Chanu (2015) "Image segmentation using K-means clustering algorithm and subtractive algorithm" *Procedia computer Science* 54,764-771.
7. Viswanathan, P. (2015). Fuzzy C Means Detection of Leukemia Based on Morphological Contour Segmentation. *Procedia Computer Science*, *58*, 84-90.
8. Lalaoui, L., Mohamadi, T., &Djaalab, A. (2015). New Method for Image Segmentation. *Procedia - Social and Behavioral Sciences*, *195*, 1971-1980.
9. Ruberto, C. D., &Putzu, L. (2014). Accurate Blood Cells Segmentation through Intuitionistic Fuzzy Set Threshold. *Tenth International Conference on Signal-Image Technology and Internet-Based Systems*.
10. Pallavi Purohit & Ritesh Joshi (2013)"A New Efficient Approach towards k-means Clustering Algorithm" *International Journal of Computer Applications* (0975 – 8887) Volume 65– No.11, March 2013.
11. Mohammed, E. A., Mohamed, M. M., Naugler, C., & Far, B. H. (2013). Chronic lymphocytic Leukemia cell segmentation from microscopic blood images using watershed algorithm and optimal thresholding. 26th *IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*.
12. Madhu Yedla, Srinivasa Rao Pathakota, T M Srinivasa (2010),"Enhancing K-means Clustering Algorithm with Improved Initial Center" *International Journal of Computer Science and Information Technologies*, Vol. 1 (2), 121-125.
13. M.Saranya& P. Krishnakumari (2011)" An Efficient Algorithm to fix Initial Centroid for Clustering High Dimensional data" *International Journal of Advanced Research in Computer Science*, Volume 2, No. 4, July-August.
14. Elindra Ambar Pambudi, Pulung Nurtantio Andono, Ricardus Anggi Pramunendar," Image Segmentation Analysis Based On K-Means PSO by Using Three Distance Measures", *ICTACT Journal On Image And Video Processing*, August 2018, Volume: 09, Issue: 01.
15. ALL-IDB "Acute Lymphoblastic Leukemia Image Database for Image Processing," Department of Information Technology -UniversitàdegliStudi di Milano, 2005. Available: http://crema.di.unimi.it/~fscotti/all/