Segmentation of Histological Images using Artificial Bee Colony Algorithm

Khalif Muse WARSAME[1], Rifat KURBAN[2], Ali DURMUŞ[3]

[1]*Erciyes University, Graduate School of Natural and Applied Sciences, Dept. of Computer Engineering, Kayseri, Turkey, khalifmuse515@gmail.com*
[2]*Kayseri University, School of Technical Sciences, Dept. of Computer Technologies, Kayseri, Turkey, rifatkurban@kayseri.edu.tr*
[3] *Kayseri University, School of Technical Sciences, Dept. of Electricity and Energy, Kayseri, Turkey, alidurmus@kayseri.edu.tr*

***Abstract*— Image segmentation is an important topic in image processing. Various algorithms are proposed for segmentation of medical images. In this paper, we study 3 methods for color image segmentation, particularly histological images stained with Hematoxylin and Eosin. Color image segmentation is a process of extracting information from the images more connected regions satisfying uniformity criterion which is based on features derived from spectral component. Color images provide more information than gray version of it. In this paper, images obtained from the literature are segmented by K-means, fuzzy C-means and artificial bee colony (ABC) algorithm based iterative Euclidean distance minimization. Experiments are conducted both on RGB color spaces and Lab color spaces. Experimental analyses show that methods achieve successful segmentation results.**

***Keywords*— Histological image segmentation, K-means, Fuzzy c-means, ABC (artificial bee colony).**

I. INTRODUCTION

Pathology analyzes the change in cell structure and tissue distribution by using microscope and then determines if a biopsy is malignant [1]. Image segmentation is a difficult and complex task in image processing, computer vision and pattern recognition applications [2,3]. Image segmentation is an essential technique in medical diagnosis through the analysis of the microscopic images. Image segmentation separates objects of interest from background through various methods in image processing such as thresholding where the intensity value is used to separate the regions [4]. Well known image segmentation methods can be grouped as: thresholding algorithms [5, 6], clustering techniques [7], compression-based methods [8], histogram-based methods [9], edge detection [10] and region-based techniques [11].

In image segmentation, the gray level of pixels of the object are very different from the gray level of the pixels of the background. Thresholding is an effective technique in separating the object from the background. The main aim of thresholding is to choose the gray-level threshold value for discriminating the target object in an image from the background in term of their gray-level distribution. In threshold technique, binary image is computed from converted gray level image. The information of the binary image should include the positions and shapes of the objects of interest [5]. The main problem of global thresholding techniques is that they cannot distinguish those pixels which share the same gray level but do not belong to the same group [6].

Clustering method is to divide data set at hand into groups are more like each other than elements in different clusters. Clustering is a method of grouping a set of patterns into several clusters such that similar patterns are assigned to one cluster. Each cluster has a well-defined centroid. The objective of clustering technique is to measure the similarity or to compute the distance between the respective patterns. Clustering methods can be categorized into two such as supervised and unsupervised. Supervised learning-based image segmentation algorithms separate the image based on the sample of object colors using prior knowledge. Unsupervised methods are automated methods, and they segment regions in feature space that have higher density [7].

In image segmentation techniques based on similarity criteria and region splitting and merging technique are very similar methods. The main aim of region splitting technique is to define similarity of the image. Region splitting technique checks the quadrants for the same defined criteria and divides it further into four quadrants if the test result is negative and the process continues till the criteria is satisfied [12].

Medical imaging becomes an essential part of health care, recently. There are many medical imaging methods are widely used such as X-rays, ultrasound, magnetic resonance imaging (MRI) and computed tomography (CT). When images are segmented perfectly, one can easily extract information from them, and this makes them simple and easy to analysis. That makes medical decision more reliable and increases patient safety.

Developing an automatic image segmentation algorithm, which aims at automated extraction of object boundary feature, plays an important role in understanding medical image contents for diagnosis and analysis. Medical image processing plays an important role in assisting health care providers for proper diagnosis and suitable treatment. In general, most of existing methods do not use adaptive optimization techniques and have a complex structure. RGB color model is also preferred instead of using other color spaces.

Recently, automatic image segmentation techniques for histological image segmentation become an essential for medical diagnosis and analysis to understand if cell is affected by the cancer or not. Motivation of this study is to develop an automatic image segmentation algorithm that can be used to segment histological images stained with Hematoxylin and Eosin (H&E) to produce an optimum segmented image that makes possible to analyze and detect easily for the cancer.

Aim of this study is to simplify representation of pathological images to make the image reliable, more useful and simply to analysis. In this paper, a simple segmentation technique based on Euclidean distance based iterative clustering and Artificial Bee Colony algorithm on histological images using both RGB and Lab color space is developed.

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

II. COLOR SPACES AND SEGMENTATION ALGORITHMS

A. RGB and Lab Color Spaces

A color space is a systematic way that can be representation of a set of colors. Color space is a technique that can be specified and visualize existing colors. Humans can describe a color by its hue (color), saturation and brightness (luminosity). However, computers define a color by using its amount of RGB phosphor emission required to match colors. Clustering method scan used to define similarity clusters of elements in the feature space and label cluster as a different region. Color space allows for reproducible representations of color, in analog and digital representations.

The Lab space is useful for sharpening images and the removing artifacts in JPEG images or in images from digital cameras and scanners. The Commission Internationale del'Eclairage (CIE) has been developed this color space. Lab color space is a color opponent space has three components L (lightness), a (involved red and green colors) and b (controls blue and yellow colors). It has non-linear coordinates and is designed to approximate human vision. Also, it is used as an interchange format between different devices. Lab model is three dimensional and can only presented in a three-dimensional space [13].

B. Segmentation Methods

Partitioning an image into a set of regions that can cover most of the image is called as image segmentation. It is the task of partitioning an image into similar areas or groups. Changing representation and simplifying an image into something useful and easier to diagnose that is the goal of an image segmentation technique. An unsupervised method is an automatic clustering that can be grouping unleveled data set into clusters. Segmentation techniques can be classified into the following basic concepts: pixels-based, edge-based, cluster-based, region-based, color-based and hybrid methods.

III. IMAGE SEGMENTATION USING ABC ALGORITHM

In this study, a method based on ABC optimization algorithm is compared with well-known methods.

A. Artificial Bee Colony (ABC) Algorithm

Karaboga proposed an efficient optimization technique using less parameter compared with others [14]. Artificial Bee Colony (ABC) algorithm compared to other population techniques, ABC demonstrates simplicity and flexibility. In ABC, food position stands a solution to the problem and nectar represents the objective function of the solution. Artificial bees are grouped into three: employed bees, onlooker bees, and scout bees. Each employed bee carries with her information about the associated food source and shares the information to onlookers. Onlooker bees wait in the hive on the dance area for the information about the possible food sources, and then make decision to choose a food source in order to exploit it. Onlooker bees select a food source related with the quality. Food sources with high quantity of nectar attract more onlooker bees compared to ones with a lower quantity of nectar. The employed bee whose food source has been exhausted becomes a scout bee which explores a possible promising food source in the neighborhood of the hive. In ABC algorithm, the number of employed and onlooker bees are equal.

B. Proposed Segmentation Method

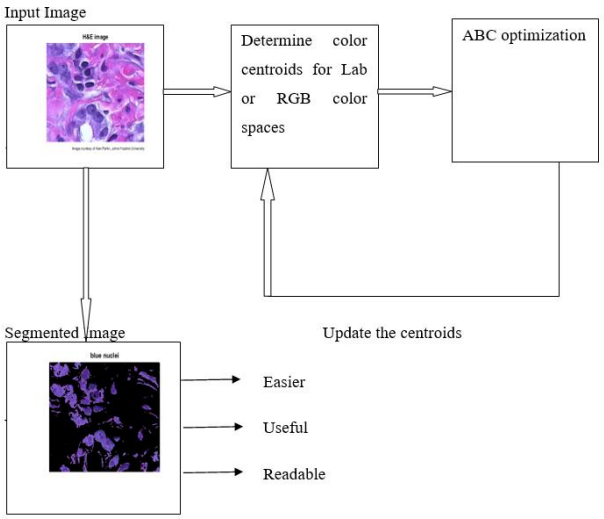


Fig. 1. Flowchart of the proposed segmentation method.

An iterative segmentation technique based on optimizing the distance of sub-neighbors of each centroid by using ABC is proposed. The method can use both RGB and Lab color values of the pixels. The block diagram of the method is given in Figure 1. The main steps of the algorithm can be summarized as follows:

1) Read RGB image and decomposed R, G and B channels.
2) Convert image from RGB color space to Lab color space and decompose L, a and b channels.
3) Choose either to use RGB or Lab color space.
4) Segment the pixel values by using ABC:
   a) Generate initial random 3 centroids because of the images are composed of 3 parts: cell nuclei, cytoplasm and the remaining. Centroids are initialized in continuous space in the range of [0,255].
   b) Calculate the Euclidean distance of each pixel color value the generated centroids: $D_j = \sum_{i=1}^{N}(P_i - C_j)^2$ (1) where N is the total count of pixels, P is a pixel's color value, C is a generated centroid and j={1,2,3}.
   c) Find the $E_i = \min(D_j)$ for each pixel and assign C_j to L_i where D_j is minimum. E is the error vector for all pixels and L is the segmented image.
   d) Minimize the total error of each pixel by updating the centroids C using employed, onlooker and scout bees.
   e) Repeat the steps until a predefined cycle number is met.
5) Use ED as the error metric and L as the segmented image.

IV. EXPERIMENTS AND RESULTS

K-means, Fuzzy C-means and the ABC based method are used to segment histological images stained with H & E, and

the results are compared. Also, sum of Euclidean distance errors (ED), mean squared error (MSE) between the original image and segmented the image and similarity between original images and segmented image (SSIM) are also calculated. The test images in the dataset are summarized in Table I.

TABLE I. DESCRIPTION OF IMAGES DATASET USED IN THE EXPERIMENTS

Image Number	Description
Image #1	Size 200 x 132 Granular cell variant of clear cell renal cell carcinoma on hematoxylin and e stain. Institution Medscape [15]
Image #2	size 586 x 345 Cross-section of multi-cellular bio printed human liver tissue, stained with hematoxylin and Eosin (H&E) [16]

Parameter settings for the methods used in the experiments are given in Table II. The parameters are the default values of the methods.

TABLE II. PARAMETER SETTINGS FOR DIFFERENT CLUSTERING METHODS.

K-means		Fuzzy C-means		ABC	
Parameter	**Value**	**Parameter**	**Value**	**Parameter**	**Value**
Distance	Euclidean	Exponent for U	2	Pop. size	32
EmptyAction	Singleton	Max Iter	100	Max. cycle number	540
Max Iter Start	100 Plus	Min. Amout of Improvement	1e-5		

Experiments have been conducted on a desktop computer which have the following specifications: AMDX6 1055 six core CPU, 8GB of RAM and Windows 2012 Server operating system. Methods are implemented by using MATLAB programming language. Experiments are repeated 30 times due to the usage of stochastic methods.

In Figure 2, two test images stained with H&E are visualized. These images are used in the experiments to compare the algorithms. The images used are encoded at 24-bit true color RGB. In the experiments, both RGB and Lab channels are used for all methods.

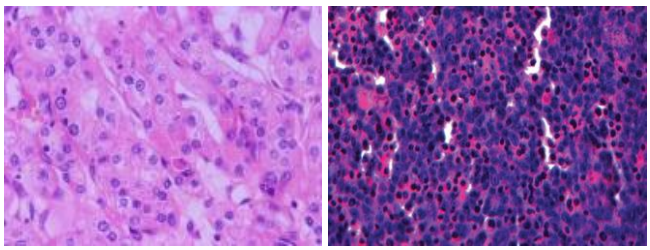


Image #1 Image #2

Fig. 2. Test images used in the experiments.

Figure 3 shows segmented images results. K-means, Fuzzy C-Means and ABC based clustering methods are used for both RGB and Lab channels. Methods clusters the images into three colors.

Table 3 illustrates the results of the two test images. The bold results show which method has better result than the other methods. In Table 3, ED, MSE and SSIM results of all algorithms are given. For ED and MSE, small values are better. For SSIM higher values are better. Quantitative results show that ABC based method is slightly better than the others. The number in the parenthesis indicates the standard deviation of 30 runs.

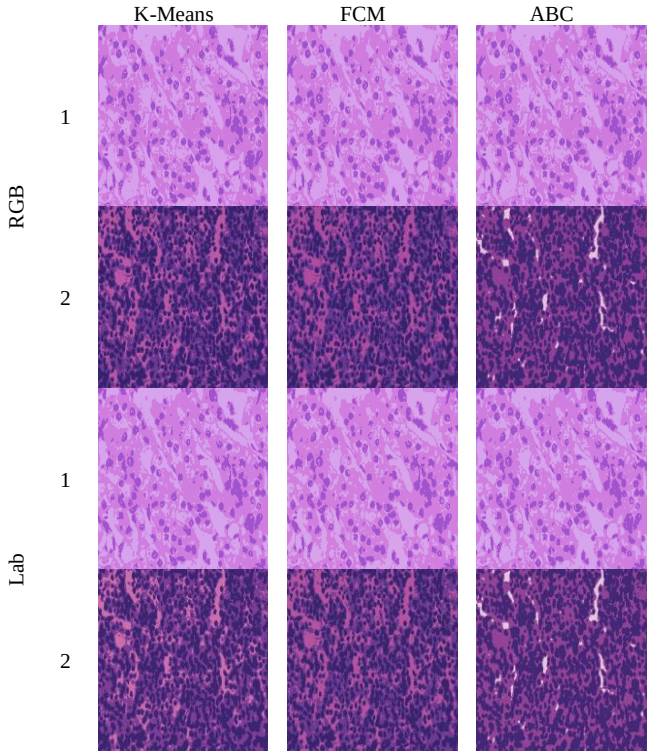


Fig. 3. Comparison of segmented images obtained with K-means, Fuzzy C-means and ABC methods

TABLE III. QUANTITATIVE RESULTS OF THE SEGMENTATION EXPERIMENTS.

			K-means	Fuzzy C-means	ABC-ED
Image #1	**RGB**	ED	141.309(0.000)	141.593(0.000)	**141.308(0.001)**
		MSE	226.850(0.071)	228.047(0.000)	**226.809(0.068)**
		SSIM	1.364(0.000)	1.361(0.000)	**1.365(0.000)**
	LAB	ED	353.887(0.000)	354.344(0.000)	**353.886(0.000)**
		MSE	226.340(0.000)	**225.953(0.000)**	226.340(0.009)
		SSIM	1.389(0.000)	1.387(0.000)	**1.390(0.001)**
Image #2	**RGB**	ED	505.498(12.684)	530.581(0.000)	**479.843(0.001)**
		MSE	941.319(16.260)	**923.256(0.000)**	974.617(0.597)
		SSIM	0.989(0.026)	**1.050(0.000)**	0.936(0.001)
	LAB	ED	1558.065(0.000)	1582.742(0.000)	**1544.946(0.040)**
		MSE	**907.836(0.006)**	927.132(0.000)	940.597(0.299)
		SSIM	1.044(0.000)	**1.065(0.000)**	0.934(0.001)

Segmentation results are visualized in the Figures 4-7. Each row of Figure 4 indicates the different parts (segments) of the images such as cell nuclei, cytoplasm or the remaining. Three segments of each image using K-means, fuzzy c-means and ABC methods for both Lab and RGB color spaces are given. In Figure 4 and 5, segmentation results of Image #1 is given for Lab and RGB color spaces, respectively.

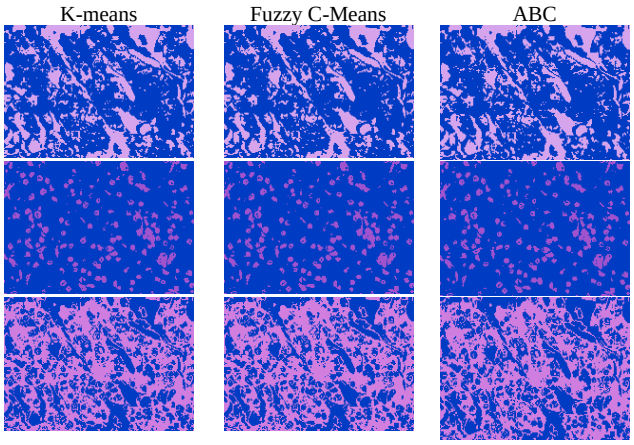


Fig.4. Visual segmentation results of Image #1 using Lab color space.

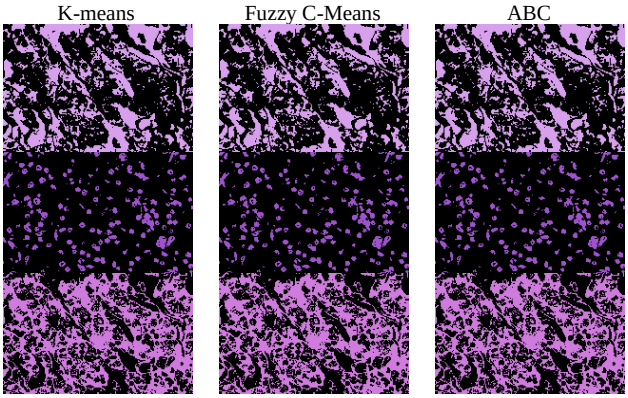


Fig.5. Visual segmentation results of Image #1 using RGB color space.

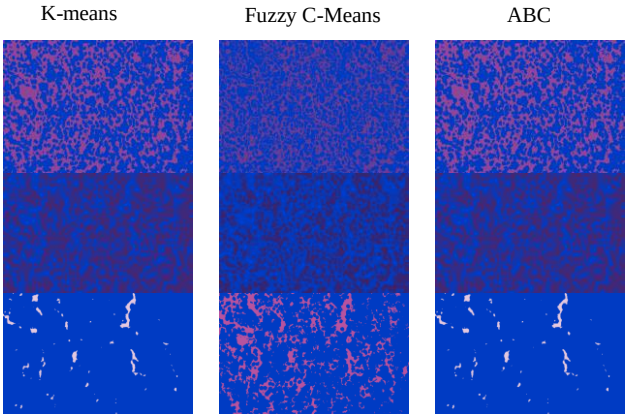


Fig.6. Visual segmentation results of Image #2 using Lab color space.

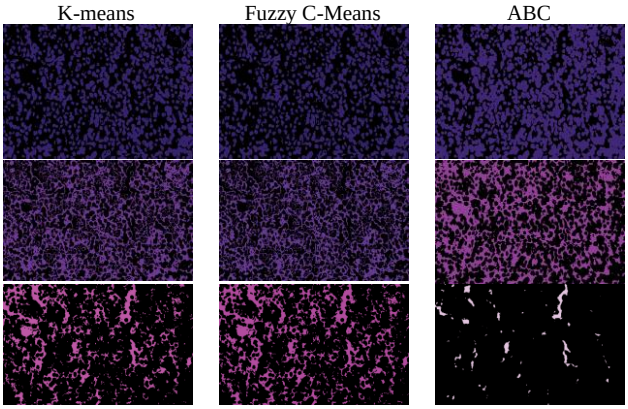


Fig.7. Visual segmentation results of Image #2 using RGB color space.

As shown in Figure 4 and 5, Lab and RGB spaces results are similar. K-means, fuzzy c-means and ABC results are similar.

In Figure 6 and 7, segmentation results of Image #2 are given for Lab and RGB color spaces, respectively.

As shown in Figure 6 and 7, Lab color space is better than RGB color space. Results of ABC method is better than K-means and Fuzzy C-Means methods.

V. CONCLUSIONS

Cancer has become major problem of the human health and the development of techniques for use in its diagnosis and treatment has increasingly become a focus of research. In this study, comparison of the three clustering techniques Fuzzy C-means, K-means and ABC optimization are compared. Fuzzy C-means and K-means are the most popular algorithms used clustering algorithms. Also, an iterative segmentation technique based on optimizing the distance of sub-neighbors of each centroid by using ABC is proposed. Image segmentation is the task of processing an image into its constituent objects. Recently, considerable work on histological segmentation has focused on fully automated approaches that rely on machine learning techniques. Color based features are important in histology applications since biologists stain tissues to highlight special structures. The color difference is used to classify specific objects to be investigated and to find relationship among them. The aim to use image segmentation technique is to find the objects of interest depending on the image and its characteristics.

ABC based segmentation aims to find nearest neighbors of each centroid to minimize the distance of each pixel to its corresponding centroid. This method is used to improve an efficient of Euclidean distance improvement by using ABC optimization algorithm. The comparative study of algorithms shows that the results are similar in most cases. However, in particular cases, ABC is slightly better than Fuzzy C-means and K-means.

Histology is the anatomical study of the microscopic structure of tissues. It is regarded as a gold standard for clinical diagnosis of diseased tissue and for the identification of therapy effects. The manual interpretation is however very time consuming and prone to subjective evaluation. The digitized tissues have become an important application of computer image analysis as well as with the use of machine learning techniques. The role of computer aided diagnosis (CAD) algorithms in medical imaging has escalated to a point where algorithms have been developed for disease detection, diagnosis, and prediction of prognosis that complements the opinion of the pathologist. Therefore, CAD systems are becoming crucial in histology analysis, as they could automatically identify regions of medical interest.

REFERENCES

[1] U. Rajyalakshmi, K. S. Prasad, and R. S. Koteswara, “Tissue processing, staining and image processing of Pathological Cancer Images A Review,” International Journal of Engineering and Advanced Technology (IJEAT), vol. 3(3), pp. 10-15, 2014.

[2] M. Xess and S. A. Agnes., “Survey on clustering based color image segmentation and novel approaches to FCM algorithm,” IJRET International Journal of Research in Engineering and Technology vol. 2(12), pp. 346-349, 2013.

[3] N. Senthilkumaran, R. Rajesh, “A study on rough set theory for medical image segmentation,”International Journal of Recent Trends in Engineering, vol. 2(2), pp. 236-238, 2009.

[4] D. Belsare, M. M. Mushrif, “Histopathological image analysis using image processing techniques an overview. Signal and Image Processing An International Journal (SIPIJ), vol. 3(4), pp. 23-36, 2012.

[5] S. S. Al-amri, N. V. Kalyankar, and ,S. D. Khamitkar, “İmage segmentation by using thershod techniques,” Journal of Computing, vol. 2(5), pp. 83-86, 2010.

[6] H. M. Som, J. M. Zain and A. J. Ghazali., “Application of threshold techniques for readibility improvement of jawi historical manuscript images.” Advanced Computing: An International Journal (ACIJ), vol. 2(2), pp. 60-69, 2011.

[7] M. C. Jobin, R. M. S. Christ Parvathi, “Magnetic resonance brain image.”, International Journal of VLSI design and Communication Systems (VLSICS), vol. 3(4), pp. 121-133, 2012.

[8] A. F. Jassim, E. H. Qassim, “Five modulus methods for image compression.” Signal and Image Processing An International Journal (SIPIJ), vol. 3(5), pp. 19-28, 2012.

[9] J. T. Orlando, and R. Seara., “Image Segmentation by Histogram Thresholding using Fuzzy Set.” IEEE Transection on Image Processing, vol. 11(12), pp. 1457-1465, 2002.

[10] N. Senthilkumaran, R. Rajesh, “Edge detection techniques for image segmentation servey of soft computing approaches” International Journal of Recent Trends in Engineering, vol. 1 (2), pp. 250-254, 2009.

[11] S. Lakshmi, V. Sankaranarayanan, “A study of edge detection techniques for segmentation computing approaches,” IJCA Special Issue on “Computer Aided Soft Computing Techniques for Imaging and Biomedical Applications” CASCT, pp. 35-41, 2010.

[12] A. M. Khan , S.Ravi, “Image Segmentation Methods: A Comparative Study,” International Journal of Soft Computing and Engineering (IJSCE), vol. 3(4), pp. 84-92, 2013.

[13] V. S. Rathore1, M. S. Kumar, A. Verma., “Colour Based Image Segmentation Using L*A*B* Colour Space Based On Genetic Algorithm.,” International Journal of Emerging Technology and Advanced Engineering, vol. 2(6), pp. 156-162, 2012

[14] D. Karaboga, B. Akay., “A comparative study of Artificial Bee Colony algorithm. Applied Mathematics and Computation,” vol. pp. 108–132, 2009.

[15] http://emedicine.medscape.com

[16] http://www.organovo.com/sites/default/files/assets/images/pic-liver2.jpg