Fuzzy C-Means Clustering Based Segmentation in CT and MRI Image

Final Project Report

Xunzhe Wen
Electrical and Computer Engineering
University of Ottawa
Ottawa, ON, Canada
xwen055@uottawa.ca

***Abstract*—This project aims to implement the functional task of segmentation, based on the a new learning algorithm, Fuzzy c-means clustering with spatial information. The traditional Fuzzy c-means algorithm is one of the soft clustering methods, which does not completely take the spatial information of an image into consideration. In this project, we implement a fuzzy c-means algorithm that counts the spatial information into the membership function for clustering according. The spatial function information can be generated by make the summation of the membership function in the neighbor region of every single pixel under consideration. The method was applied for the medical image, to fulfill one of the sub-tasks of the image processing, segmentation. Image segmentation plays a crucial role in many medical imaging applications. The advantages and the disadvantages are discussed in the report. The common used segmentation algorithms are introduced and compared with the spatial information related fuzzy c-means clustering algorithm at this project.**

Keywords—Medical image processing; Segmentation; Machine Learning; Fuzzy c-means clustering; Spatial information.

I. INTRODUCTION

Digital imaging is the creation of digital images. In all classes of digital imaging, the information is converted by image sensors into digital signals that are processed by a computer and outputted as a visible-light image. Such as given a set of possible gray levels or colors and a grid, a digital image attributes a gray value (i.e., brightness) or a color to each of the grid points or pixels. In a digital image, the gray levels are integers. Although brightness values are continuous in real life, in a digital image we have only a limited number of gray levels at our disposal. The conversion from analog samples to discrete-valued samples is called quantization. When insufficient gray values are used, contouring appears [1]. The field of digital image processing is highly interdisciplinary and draws upon a great variety of areas such as mathematics, computer graphics, computer vision, visual psychophysics, optics, and computer science. It is a subject that lends itself to a rigorous, analytical treatment and which, depending on how it is presented, is often perceived as being rather theoretical [2].

Digital imaging has demonstrated its worth in a variety of fields from education to medicine. Medical imaging, a branch of digital imaging, seeks to assist in the diagnosis and treatment of diseases. It is growing at a rapid rate because it works as a technique and process of creating visual representations of the interior of a body for clinical analysis and medical intervention. It can help people to reveal internal structures hidden by the skin and bones. In this case, once it appears different, the disease can be diagnosed and treated in time. Advances in medical imaging technology have made it possible routinely to acquire high-resolution, three dimensional images of human anatomy and function using a variety of imaging modalities. There exist many different strategies for image analysis, however, few of them are suitable for medical applications. Since that both the medical image data or even the description of features to be analyzed, are typically quite complex. The aim of medical image enhancement is to allow the clinician to perceive better all the relevant diagnostic information present in the image.

The segmentation is one of the difficult sub-tasks in the medical image processing. We can segment the image according to the anatomy structure, tissue intensity, or functional regions. That idea was similar with concept of clustering. Therefore, we figure out the idea of fuzzy clustering. Fuzzy c-means was first proposed in 1969 by Ruspin. People proposed a variety of clustering methods using this concept. It can be roughly divided into three categories according to the different clustering process: (1) based on the relationship of clustering (2) based on the objective functions (3) based on the neural network. One of the most widely used fuzzy clustering algorithms is the fuzzy c-means Algorithm [3]. Standard fuzzy c-means algorithm seldom incorporates spatial information in image segmentation.

II. LITERATURE REVIEW

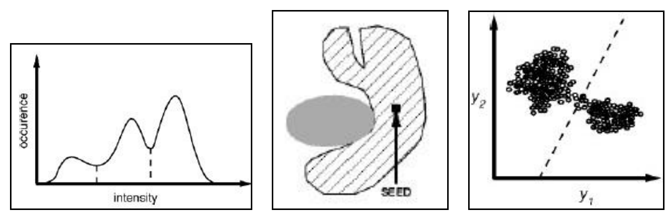
A. *Segmentations in medical image*

Segmentation of the interest region for the objects is a difficult task in the analysis of digital images. Some of the existing automatic methods can fail, producing the false results or even requires the intervention of a human operator. This is quite common in medical applications. Medical images contain many structures including normal structures such as organs, bones, muscles, fat, and abnormal structures such as tumors and fractures. Segmentation is the process of identifying structures, both normal and abnormal. It is fundamental to the interpretation of medical images. Segmenting structures from medical images is not trivial due to the complexity and variability of the region of interest. Therefore, image segmentation is particularly difficult due to restrictions imposed by image acquisition, pathology and biological variation.

Several common approaches have appeared in the recent literature on medical image segmentation. They can be divided into eight categories: threshold approaches, region growing approaches, classifiers, clustering approaches, Markov random field models, artificial neural networks, deformable models, atlas-guided approaches.

1). Thresholding:

Thresholding approaches segment scalar segment scalar images by creating a binary partitioning of the image intensities. The figure 1 shows the histogram of a scalar image that possesses three apparent classes, corresponding to the three modes. A thresholding procedure attempts to determine an intensity values, the threshold, which separates the desired classes. The segmentation is then achieved by grouping all pixels with intensities greater than the threshold into one class and all other pixels into another class. Thresholding is a simple yet often effective means for obtaining a segmentation of images in which different structures have contrasting intensities or other quantifiable features. The partition is usually generated interactively, although automated methods do exist.



(a) (b) (c)

Figure 1. Feature space methods and region growing. (a) Histogram showing three apparent classes; (b) 2-D feature space; (c) Example of region growing.

Variations on classical thresholding methods have been proposed for medical image segmentation that incorporate information based on local intensities and connectivity.

2). Region Growing:

Region growing is a technique for extracting an image region that is connected based on some predefined criteria. These criteria can be based on intensity information and edges in the image [4]. Region growing requires a seed point that is manually selected by an operator and extracts all pixels connected to the initial seed based on some predefined criteria such as growing the region until an edge in the image is met. Region growing is seldom used alone but usually within a set of image-processing operations, particularly for the delineation of small, simple structures such as tumors and lesions [5]. The primary disadvantage of region growing is that it requires manual interaction to obtain the seed point. Thus, for each region that needs to be extracted, a seed must be planted. Region growing can also be sensitive to noise, causing extracted regions to have holes or even become disconnected. Conversely, partial-volume effects can cause separate regions to become connected.

3). Classifiers:

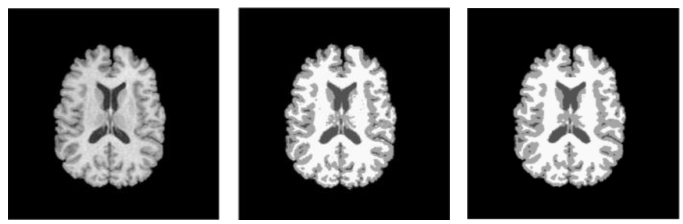
Classifier methods are pattern recognition techniques that seek to partition a feature space derived from the image by using data with known labels [6]. A feature space is the range space of any function of the image, with the most common feature space being the image intensities themselves. Classifiers are known as supervised methods because they require training data that are manually segmented and then used as references for automatically segmenting new data. There are a number of ways in which training data can be applied in classifier methods. A simple classifier is the nearest-neighbor classifier, in which each pixel is classified in the same class as the training datum with the closest intensity. Standard classifiers require that the structures to be segmented possess distinct quantifiable features. Because training data can be labeled, classifiers can transfer these labels to new data as long as the feature space sufficiently distinguishes each label as well. A disadvantage of classifiers is that they generally do not perform any spatial modeling. Another disadvantage is the requirement of manual interaction to obtain training data. Training sets can be acquired for each image that requires segmenting, but this can be time consuming and laborious. On the other hand, use of the same training set for a large number of scans can lead to biased results that do not take into account anatomical and physiological variability between different applications.

4). Clustering:

Clustering algorithms essentially perform the same function as classifier methods without the use of training data. Thus, they are termed unsupervised methods. To compensate for the lack of training data, clustering methods iteratively alternate between segmenting the image and characterizing the properties of each class. In a sense, clustering methods train themselves, using the available data. Three commonly used clustering algorithms are the K-means [7], the fuzzy c-means algorithm [6], and the expectation-maximization algorithm [8]. The K-means clustering algorithm clusters data by iteratively computing a mean intensity for each class and segmenting the image by classifying each pixel in the class with the closest mean [9]. The fuzzy c-means algorithm generalizes the K-means algorithm, allowing for soft segmentations based on fuzzy set theory. The EM algorithm applies the same clustering principles with the underlying assumption that the data follow a Gaussian mixture model, Equation 1. It iterates between computing the posterior probabilities and computing maximum likelihood estimates of the means, covariance, and mixing coefficients of the mixture model.

$$f(y_j;\theta,\pi)=\sum_{k=1}^{K}\pi_k f_k(y_j;\theta_k) \tag{1}$$

Although clustering algorithms do not require training data, they do require an initial segmentation. Like classifier methods, clustering algorithms do not directly incorporate spatial modeling and can therefore be sensitive to noise and intensity inhomogeneity. This lack of spatial modeling, however, can provide significant advantages for last computation [10].



(a) (b) (c)

Figure 2. Segmentation of a magnetic resonance brain image. (a) Original image. (b) Segmentation using the K-means algorithm. (c) Segmentation using the K-means algorithm with a Markov random field prior.

5). Markov Random Field Models

Markov random field modeling itself is not a segmentation method but a statistical model that can be used within segmentation methods. Markov random field model spatial interactions between neighboring or nearby pixels. These local correlations provide a mechanism for modeling a variety of image properties [11]. In medical imaging, they are typically used because most pixels belong to the same class as their neighboring pixels. Markov random field models are often incorporated into clustering segmentation algorithms such as the K-means algorithm under a Bayesian prior model [12]. Markov random field methods usually require computationally intensive algorithms. Despite the disadvantages, Markov random field are widely used not only to model segmentation classes, but also to model intensity inhomogeneity that can occur in MR images and textural properties, which is useful in the segmentation of digital mammograms [13].

6). Artificial Neural Networks:

Artificial neural networks are parallel networks of processing elements or nodes that simulate biological learning. Each node in an artificial neural network is capable of performing elementary computations. Learning is achieved through the adaptation of weights assigned to the connections between nodes. Artificial neural networks represent a paradigm for machine learning and can be used in a variety of ways for image segmentation. The most widely applied use in medical imaging is as a classifier [14], in which the weights are determined by using training data and the artificial neural networks is then used to segment new data. Artificial neural networks can also be used in an unsupervised learning as a clustering method [15]. Because of the many interconnections used in a neural network, spatial information can be easily incorporated into its classification procedures. Although artificial neural networks are inherently parallel, their processing is usually simulated on a standard serial computer, thus reducing this potential computational advantage.

B. Clustering as a machine learning concept

By definition, clustering algorithm is a method that divide the whole data set into different categories, these different categories are formed by the data set elements, the data set elements in the same category have similar features with each other, and are different from the elements in other categories. In the machine learning concept, the clustering is an unsupervised machine learning algorithm. To gain a better intuition for the clustering, it is important to know the concept of unsupervised learning and the difference of it with respect to supervised learning.

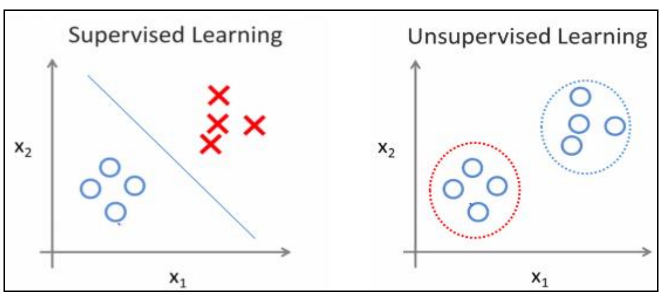


Figure 3. Supervised Learning and Unsupervised learning.

As shown in the figure above, the input data has exactly the same distribution, the axis here denote some certain feature of the input data, the only difference between supervised learning and unsupervised learning is that in supervised learning, each data will be assigned with an additional value called the "label". And the ultimate goal in supervised learning is to generate a "decision boundary" so that when receive new input data, one would know which label should be tagged to it.

As for the clustering algorithm, the situation become a little bit complicated. There are no additional features here, no "label". The only way to separate the data elements is by the distribution or say, the structure within the data set. And because we don't need to label the data manually and such algorithm can adapt itself with input data that has different distribution pattern, so it entitled with the name "Unsupervised learning". There are many clustering algorithms. And the principle of these algorithm can be totally different. This difference lies in the data structure, the purpose of clustering and the practical application. So there are many intersections between different algorithms, in other word, they are hard to be classified in a strict way. If to divide them by development, there are (1) Traditional clustering algorithms. (2) Newly developed clustering algorithms. And Fuzzy c-means clustering belongs to the newly developed algorithms for it uses the fuzzy logic, which is distinct with traditional "hard" clustering.

Introduction intuition over the traditional clustering. The method can be sub-divided in to: (1) Hierarchical clustering, which make separation based on the hierarchy structure on the data elements. (2) Partition method (Distance based), this method uses the distance, not always the Euclidean distance for the separation metric. One representative and most widely-used algorithm is K-means clustering. (3) Density-based method: this method can detect patterns from any given set theoretically. In comparison, the partition method could only detect sphere-like shaped cluster. And many other methods had been invited but are relatively less used.

By definition, clustering algorithm is a method that divide the whole data set into different categories, these different categories are formed by the data set elements, the data set elements in t

All the traditional clustering algorithm shared a common feature: Every data elements could only be assigned into one specific cluster. This is called "Crisp Clustering", but many real-life problems cannot be defined by a clear threshold. And the traditional methods struggle with this kind of problems. The fuzzy clustering provide a powerful way in solving this kind of problems.

In fuzzy clustering (soft clustering), data elements can belong to different clusters at the same time [16], and a set of fuzzy membership levels is associated with each element. The fuzzy membership indicates the strength of the association between that data element and a particular cluster, or say, the probability can one element belong to one particular cluster. The Fuzzy c-means algorithm is one of the most widely used fuzzy clustering algorithms. The algorithm tries to partition a finite number of data elements into a matrix that contains fuzzy clusters by some criterion. The description of the fuzzy clusters is a list of cluster centers and another matrix to store the fuzzy membership which indicates the strength of the association between elements and clusters. Apart of fuzzy c-means clustering algorithm, there are many other fuzzy algorithms, and they use different functions to depict the similarity between different elements. The advantage of fuzzy clustering is that it considers not only the belonging status to the clusters, but also to consider to what degree do the objects belong to the clusters [17]. But apart from this, experimental results prove that fuzzy clustering seems also to be more robust in terms of local minimum of the objective function. Different kinds of fuzzy clustering have different disadvantages and applications. A conclusive table is listed below for reference.

For more details about fuzzy c-means clustering, it has such advantages: (1) Allows a data point to be in multiple clusters, (2) It is a more natural representation of the complex structure. (3) Fuzzy c-means clustering will always converge. It also has the disadvantages: Need to define c, the number of clusters, (1) Need to determine membership cutoff value, (2) Clusters are sensitive to initial assignment of centroids, (3) Fuzzy c-means is not a deterministic algorithm, in other words, the output is only a separation based on the distribution, may not reveal the intrinsic property of data.

III. METHODS

Clustering analysis is a branch of unsupervised pattern recognition. The fuzzy cluster analysis attracts more and more attention recently with introducing the fuzzy set theory. Clustering algorithms are used to find groups in unlabeled data, based on a similarity measure between the data elements. This means that similar patterns are placed together in the same cluster. The main difference between fuzzy clustering and other clustering techniques is that it generates fuzzy partitions of the data instead of hard partitions. Therefore, data patterns may belong to several clusters, having in each cluster different membership values.

The fuzzy c-means algorithm is one of the most widely used method for data clustering, and has been successfully applied to feature analysis and classifier designs in fields such as astronomy, geology, medical imaging, target recognition, and image segmentation. But, in the fuzzy c-means based image segmentation algorithm, feature vectors are assumed to be independent of each other and independent of their spatial coordinates. However, real-world images usually have strong correlation between neighboring pixels. Adjacent pixels in an object are generally not independent of each other. Thus, the incorporation of local spatial interaction between adjacent pixels in the fuzzy clustering process can produce more meaningful classification, as well as help to resolve classification ambiguities due to overlap in intensity value between clusters or noise corruption. In this paper, we implemented a fuzzy c-means image segmentation by utilizing local spatial information.

	Advantages	**Disadvantages**	**Suitable for…**
Fuzzy C Varieties (FCV)	Each cluster represents an r dimensional variety in the dimension of the data space	o The areas of high membership exceed beyond the line segments o A higher number of clusters increases the number of local minima	Lines, planes and hyper planes
Adaptive Fuzzy Clustering(AFC)	Able to recognize elliptic or circular clusters	The eigenvalues have to be computed to update the prototypes, any changes are hardly visible	Line segments
Fuzzy C Means (FCM) Algorithm	Few iterations steps already provide good approximation to the final solution	FCM tends to locate centroid in the neighborhood of the larger cluster and misses the small, well-separated cluster.	Spherical shape
Gustafson Kessel (GK) Algorithm	Faster than AFC. In order to adapt to different structures in data, GK used the covariance matrix to capture ellipsoidal properties of clusters.	The clusters are narrower and the areas with higher membership are thinner	Line segments
Gath-Geva (GG) Algorithm	Unlike FCM and GK algorithm, it is not based on objective function. It is a fuzzification of statistical estimators	Because the occurrence of the exponential function within the distance, the distance divided into two range, close and remote	Line segments

Table.1 The features of different fuzzy cluster algorithms. [18]

1). Fuzzy C-Means Clustering Algorithm.

The main idea of fuzzy c-means clustering is to optimize an "objective function" (cost function), to constrain the error and lifting the performance in each iteration. The cost function are shown below [19]:

$$\min_{(U,V)}\{J_m(U,V)\} = \sum_{i=1}^{c}\sum_{k=1}^{n} u_{ik}^m D_{ik}^2 \tag{2}$$

where the D_{ik} term here is the distance metric, and it satisfies this equation: $D_{ik}^2 = \|x_k - v_i\|_A^2$ (the double slash denotes the norm metric, the norm we use is a plain distance in Euclidean space). There is the fuzzy term which represents the probability (as introduced in literature review). u_{ik} is under the constrain: $\sum_{i=0}^{c} u_{ik} = 1, \ \forall k$, and c is the number of the clusters.

The fuzzy term u_{ik} is computed with the equation:

$$u_{ij}^{m} = \sum_{k=1}^{c} \left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{-\frac{2}{m-1}} \quad (3)$$

the parameter m here controls how fuzzy the algorithm will be, typically, m can be any value that are greater than 1. When m=1, the fuzzy clustering will degenerate to a crispy clustering. A typical value of m=2 is widely chosen. After the fuzzy membership be computed, the vector of clusters centers can also be carried out by this equation:

$$v_i = \left(\sum_{k=1}^{n} u_{ik}^{m} \cdot x_k \right) / \left(\sum_{k=1}^{n} u_{ik}^{m} \right) \quad (4)$$

The first set of clusters centers are randomly generated, of course they have little chance to be a global minimum for the cost function. Then, the fuzzy membership term will be computed based on the cluster center vector computed previously. And whether the error of clusters are small enough and iteration should stop, is determined by this termination measure:

$$E_t = \|V_t - V_{t-1}\| \quad (5)$$

The norm is still the Euclidean distance. From the equation we know the termination condition is that the distance between two consecutive clusters are acceptably small. And this threshold is adjustable base on the specific application and different distribution of input data. A typical threshold for a wide range of data is $\varepsilon = 0.01$ [20].

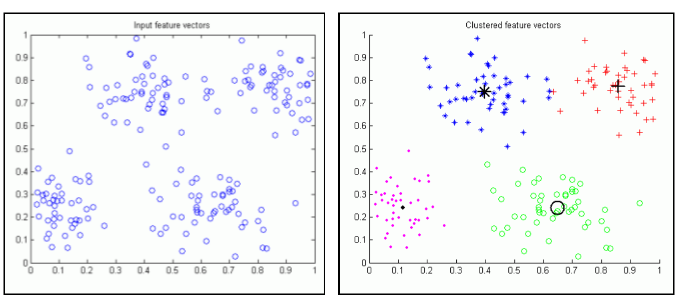


Figure 4. Data after Fuzzy c-means clustering

At last, the algorithm of fuzzy c-means clustering can be concluded as follows:

Step 1: Initialize membership for data point   of cluster by random;

Step 2: At the k-th step, compute the fuzzy centroid V, for i= 1, ...n, where n is the number of clusters, using:

$$v_i = \left(\sum_{k=1}^{n} u_{ik}^{m} \cdot x_k \right) / \left(\sum_{k=1}^{n} u_{ik}^{m} \right) \quad (6)$$

Step 3: Update the fuzzy membership, using:

$$u_{ij}^{m} = \sum_{k=1}^{c} \left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{-\frac{2}{m-1}} \quad (7)$$

Step 4: If $\|V_t - V_{t-1}\| \le \varepsilon$, then stop the iteration, else return to step 2;

Step 5: Determine membership cutoff: for each data point x_i, assign x_i to cluster v_i.

2). Spatial Information Related.

One of the significant features of a medical image is that neighborhood pixels are highly correlated. In other words, these neighborhood pixels can have a similar feature values, and the probability that they belong to the same cluster is great. This spatial related information is important in the clustering algorithms, but it is not utilized in the conventional fuzzy c-means clustering algorithm. To make the use of the spatial information, a function can be defined as [19]:

$$h_{ij} = \sum_{k \in NB(x_j)} u_{ik} \quad (8)$$

where $NB(x_j)$ represents a square window centered on pixel x_j in the spatial domain. A 5×5 window was used throughout this work. Just like the membership function, the spatial function h_{ij} represents the probability that pixel x_j belongs to *i*-th clusters.

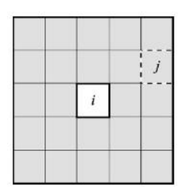


Figure 5. Local image window.

The spatial function is incorporated into membership function as following:

$$u_{ij}^{i} = \frac{u_{ij}^{p} \cdot h_{ij}^{q}}{\sum_{k=1}^{c} u_{kj}^{p} \cdot h_{kj}^{q}} \tag{9}$$

where p and q are parameters to control the relative importance of both functions. In a homogeneous region, the spatial functions simply fortify the original membership, and the clustering result remains unchanged. However, for a noisy pixel, this formula reduces the weighting of a noisy cluster by the labels of its neighboring pixels. As a result, misclassified pixels from noisy regions or spurious blobs can easily be corrected.

In the traditional fuzzy c-means clustering image segmentation algorithm, clusters are based on the distribution of pixel attributes in the feature domain, and do not take the spatial distribution of pixels in an image into consideration. We present a novel fuzzy c-means clustering image segmentation algorithm by utilizing local feature information and the high level inter-pixel correlation inherent. Firstly, a local spatial similarity measure model is established, and the initial clustering center and initial membership are determined adaptively based on local spatial similarity measurement. Secondly, the fuzzy membership function is modified according to the high inter-pixel correlation inherent. Finally, the image is segmented by using the spatial information related fuzzy c-means clustering algorithm.

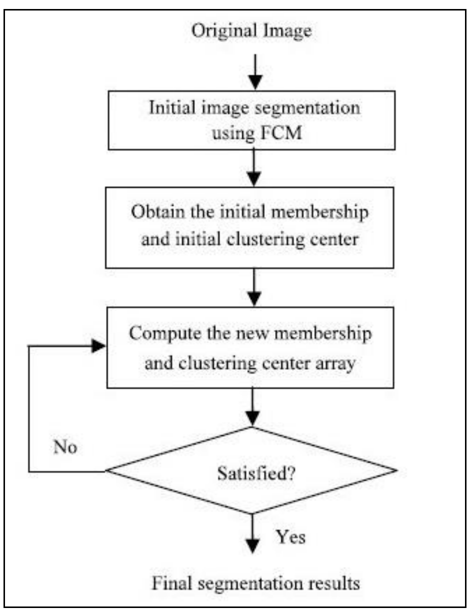


Figure 6. Image segmentation using FCM with spatial information.

Step 1: Set the number of c of the cluster prototypes, initialize randomly those prototypes and set $\varepsilon > 0$ to a very small value. (In this paper, we set $c = 2$, $m = 2$, $\varepsilon = 0.0001$);

Step 2: Compute the local image feature using for all neighbor windows over the image;

Step 3: Go into the iteration to evolve the membership function, which related to the current clustering center results and initial membership;

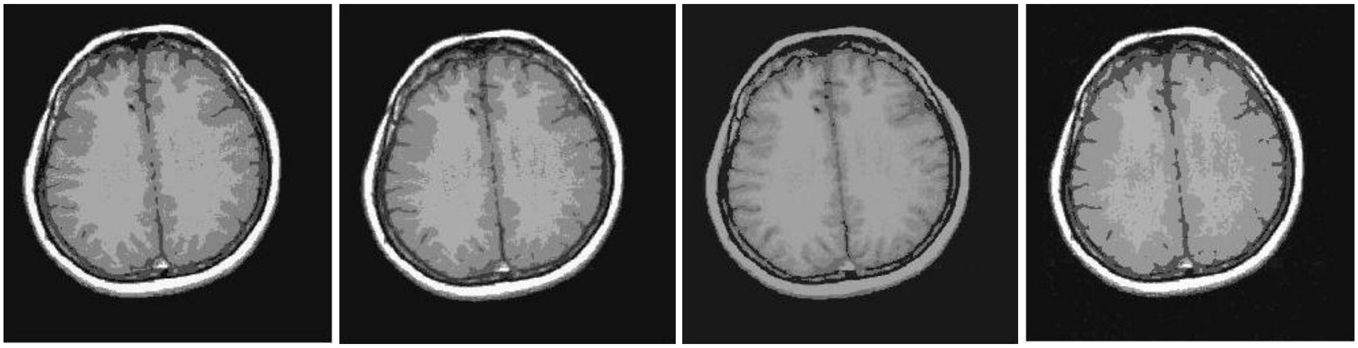
Step 4: Repeat the iteration, until the final segmentation results are obtained.

IV. EXPERIMENTAL RESULTS

For the implementation, we used the cross-section of brain MRI image to test the performance of fuzzy c-means clustering. The image labeled (a)-(b) respectively are: brain MRI cross-section image after applying: (a) conventional fuzzy c-means clustering without spatial information, (b) spatial information included with p=1, q=1, (c) p=0, q=2, (d) p=2, q=2.

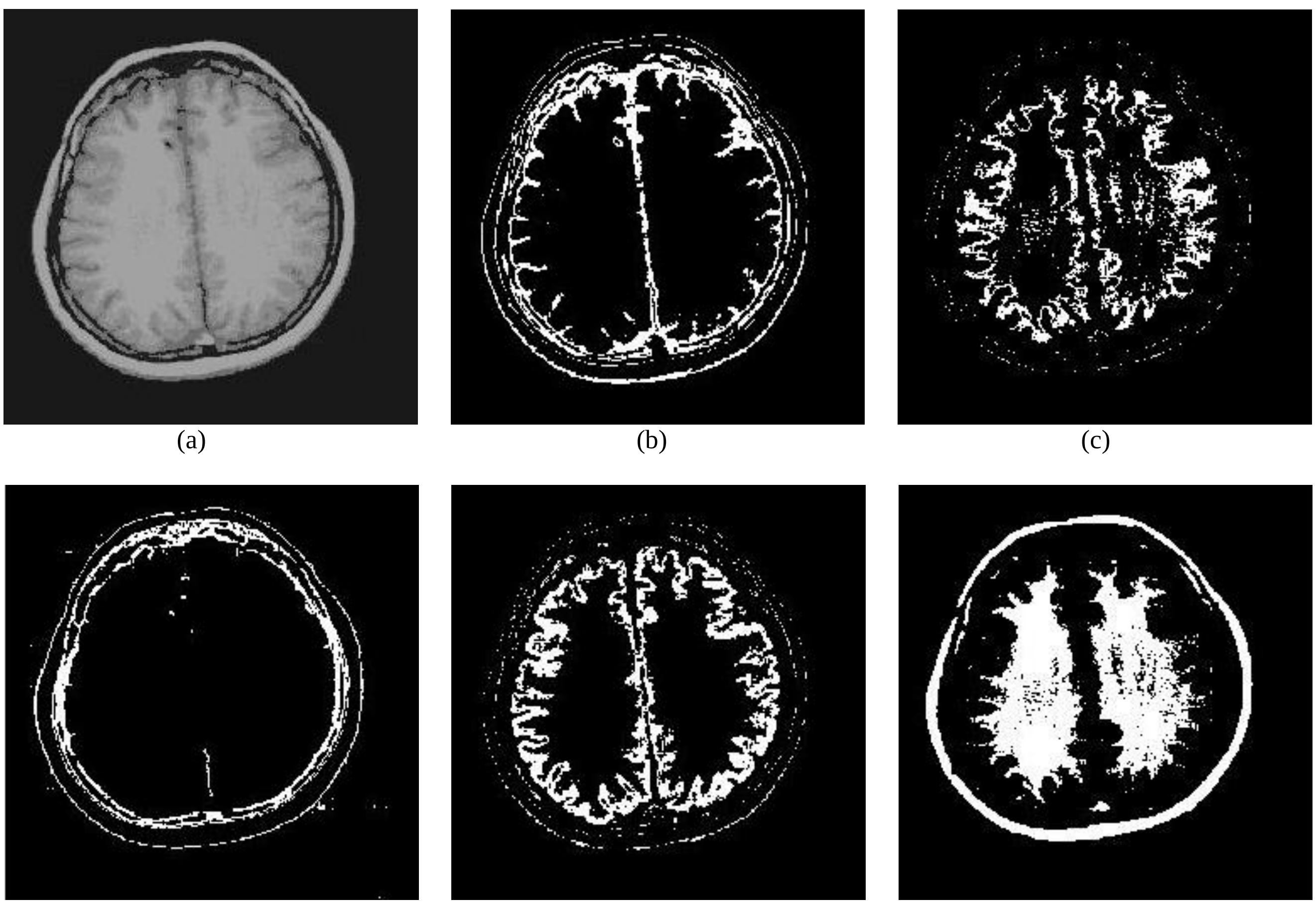
In implementation we observed that the conventional fuzzy c-means clustering successfully classifies the MRI images into 6 categories. In fig.7, these different categories of brain matter are shown. However, we observed that there are spurious blobs of gray matter appear inside the white matter cluster. The spatial function do the job to modify the fuzzy membership function of a pixel according to the neighborhood situation. Such neighboring effect will distort the clustering output. And after apply spatial information into conventional fuzzy c-means clustering, and with proper chosen p and q value, this spurious blob decreased, and the segmented images become more homogeneous. And the higher q parameter showed a better smoothing effect. But the disadvantage is that if the q value become too high, it may cover some of the fine details and blur the whole image (Although with a q value of 2, this distortion effect is not so obvious).

These two sets of images showed us the advantage of fuzzy c-means clustering with spatial information. The algorithm has exactly the same parameter as previous (Figure .7). As we can see, even if we input the images that has totally different distribution of matters (Figure 7 compared to Figure .8), or the imaging principle is completely different (Figure .7, the MRI compared to Figure .9, the CT image). For the advantage of the fuzzy c-means we discussed before, with very easy adjustment, the algorithm could work with different input image. And the performance are also acceptable. The reason is that the fuzzy c-means clustering with spatial information utilizes the distance between pixels and cluster centroids in the feature domain (which is intensity and spatial information



(a) (b) (c) (d)

Figure 7. (a). Conventional fuzzy c-means clustering; (b). Spatial fuzzy c-means clustering, with p=1, q=1; (c). Spatial fuzzy c-means clustering, with p=0, q=2; (d). Spatial fuzzy c-means clustering, with p=2, q=2.



(a) (b) (c)

(d) (e) (f)

Figure 8. The images from left to right, top to down are: (a). spatial information related fuzzy c-means clustering image; (b). Fat tissue; (c). Fluid; (d). Bone and skin; (e). Gray matter; (f). White matter.

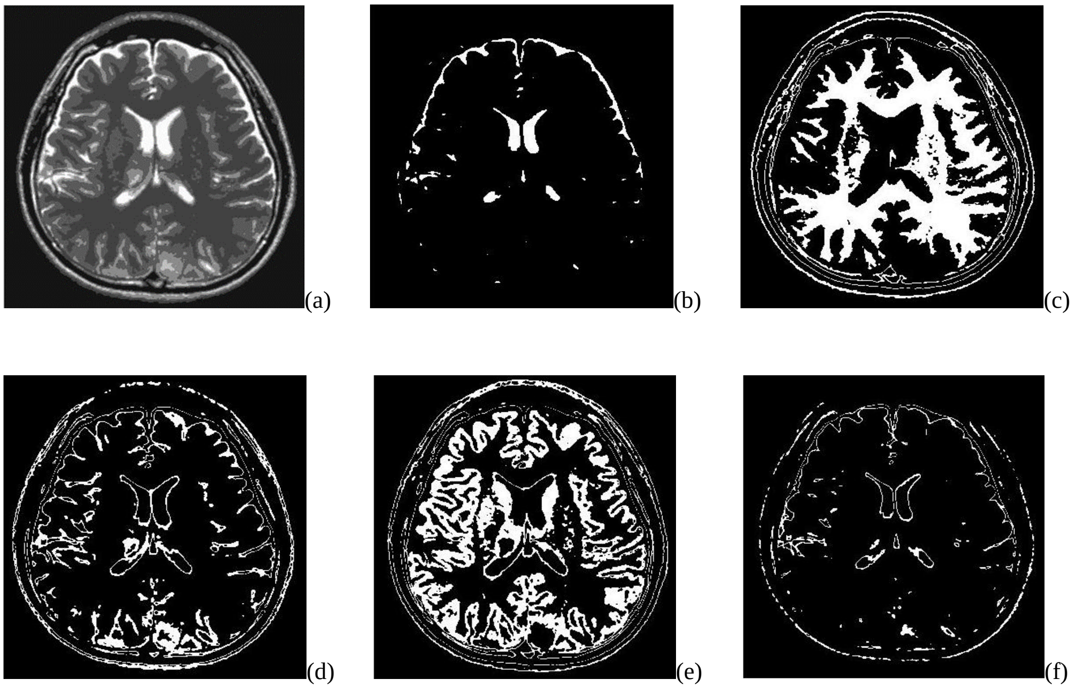


Figure 9. The segmentation results of brain CT image.

Iteration times	*Cluster 1*	*Cluster 2*	*Cluster 3*	*Cluster 4*	*Cluster 5*	*Cluster 6*
1	20.2793	56.4470	154.0304	82.9628	253.6363	201.3368
2	20.1445	61.6897	149.2909	86.1771	253.5038	201.0525
3	20.2373	64.0797	146.1387	89.3637	253.4747	200.3314
4	20.2955	65.2866	144.1567	91.7896	253.4435	199.4652
5	20.3280	65.9646	142.9451	93.5784	253.4088	198.6169
6	20.3473	66.3790	142.2259	94.8930	253.3754	197.8620
7	20.3596	66.6848	141.8168	95.8524	253.3457	197.2258
8	20.3678	66.8320	141.5990	96.5477	253.3280	196.7085
9	20.3735	66.9594	141.4958	97.0493	253.3006	196.2985
10	20.3755	67.0492	141.4587	97.4108	253.2847	195.9798

Table 2. The trend of the clustering centroids for each cluster, within 10 iterations (CT image).

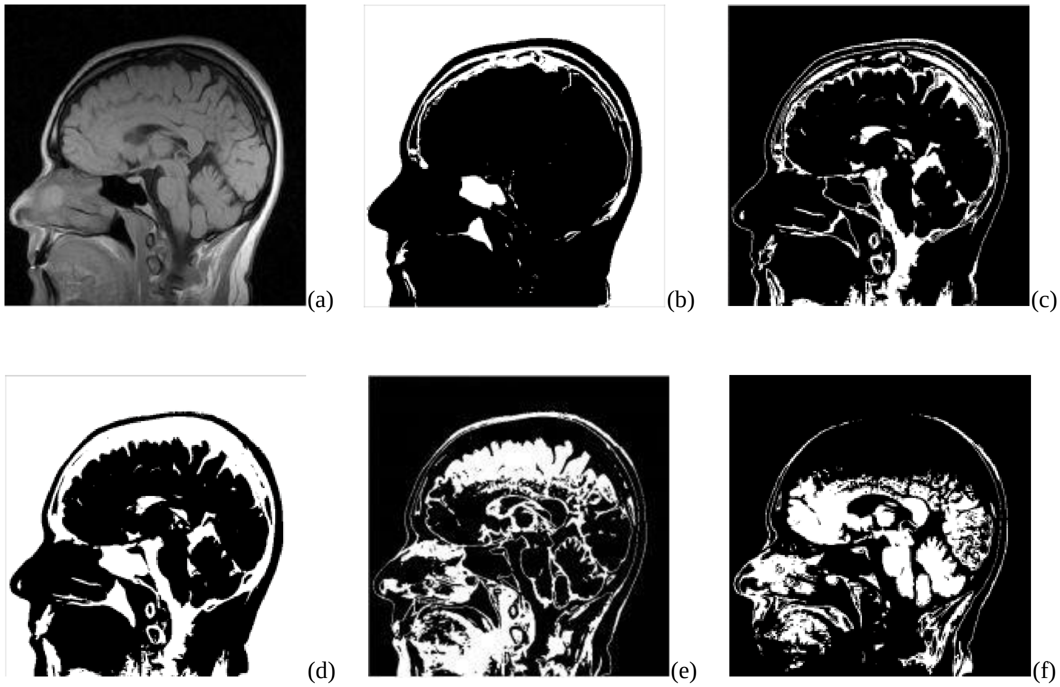


Figure 10. The head longitudinal section MRI image.

Iteration times	*Cluster 1*	*Cluster 2*	*Cluster 3*	*Cluster 4*	*Cluster 5*	*Cluster 6*
1	5.2405	50.7173	145.5991	88.4227	209.9773	206.3350
2	2.5953	49.3159	133.2666	93.1432	212.0931	202.7370
3	2.4677	48.4662	129.1241	94.9373	213.4388	198.3945
4	2.4554	47.9961	127.7342	95.6029	213.8887	194.2850
5	2.4508	47.7128	127.1816	95.8590	213.8437	190.3646
6	2.4481	47.5302	126.9306	95.9489	213.6499	186.6571
7	2.4462	47.4060	126.7951	95.9674	213.4409	183.3027
8	2.4449	47.3172	126.7044	95.9539	213.2465	180.4702
9	2.4439	47.2511	126.6320	95.9257	213.0741	178.2047
10	2.4432	47.2000	126.5687	95.8908	212.9277	176.4170

Table 3. The trend of the clustering centroids for each cluster, within 10 iterations (MRI image).

here). The pixels on the image (no matter their imaging principle or intensity distribution) are highly correlated, and this spatial information shows a vital characteristic that could be utilized to help their tagging. Nonetheless, the spatial relationship in between pixels is seldom used in fuzzy c-means clustering, which limited the performance of FCM, with spatial information included, the algorithm will be more robust against the noise, and the performance will be relatively better when compared to traditional fuzzy c-means clustering (e.g. the spurious blob in white matter reduced drastically).

In conclusion, the fuzzy c-means clustering showed a promising result in segmenting different types of images while human labor needed to adjust the algorithm is limited, which made it a compatible algorithm to solve many kinds of medical image segmentation problem.

V. COMPARISON & DISCUSSION

The complexity and diversity of medical image made the segmentation of medical image extremely hard to apply, even today, this problem is not completely solved. However, years of work done by researchers have carried out several practical algorithms. Some of the most widely used segmentation method will be listed below:

1). Threshold Based Method:

Segmentation based on threshold, this method detect candidate region in a direct and parallel way. If choose only one threshold, it is single threshold segmentation (e.g. Otsu), the other way is multiple threshold segmentation [21]. The image will be segmented into multiple targets and background. In a gray-scale image, the intensity values of adjacent pixels are similar, but different targets and background has different intensity value. In histogram, different targets and background will have different peaks. The threshold should be located between different peaks in order for separation.

The advantage of threshold-based segmentation is its simplicity, when different targets and background have significant difference in intensity, the separation based on threshold will be not only effective, but also computationally inexpensive. The disadvantage is that it cannot be applied to multi-channel images, and when the intensity difference is insignificant, or then intensity range has overlapping, the performance will be poor. What's more, the threshold method only consider the intensity, the spatial information is lost, thus, it would be sensitive to noise and intensity asymmetry.

2). Classifier-based segmentation

The classification is a fundamental statistic method. The purpose of classification is with the training data that are already known, to find a curved face in 2D or 3D, or even a high dimension to partition the image. Use classifier for partition is a supervised machine learning method [22]. It needs manual cropped images as the training sample to segment the new input image. The most widely used classifiers are SVM, Artificial Neural Network, bayesian Classifier.

There are two advantages of classifier: (1) It doesn't need to run iterative computation. Thus has relatively small computational cost. (2)It can be applied to multi-channel images. But the classifier also doesn't utilize the spatial information, so its performance are limited when deal with image of asymmetric intensity. Classifiers also require the human labor to generate the training data, the work load can be very high. In the same time, when the amount of training data is insufficient, the classification will end up having large error, because it will fail to generalize the individual difference.

3). Edge detection method

The edge detection method is one of the earliest segmentation method. Based on the drastic change in the contour of the image, it tries to segment the image by detecting the contour of the different areas. The simplest edge detection method is to use the differential operator, it uses the property that the intensity is not continuous, and uses the first order or second order differential to detect the edge point. The common used differential operators are: Prewitt, Sobel, Laplacian and other non-linear operators. But the disadvantage is that these operators will not only be sensitive to edge information, not also to the noise.

To reduce the interference of the noise to image, usually before differentiation, a filtering step will be applied first. A Gaussian function for filtering is commonly used, John Canny think the first order differentiation of Gaussian function is a good approximation of the optimal filter he found [23].

Here we picked some of the most representative algorithms out of the methods introduced above, and made implementation in order to gain a direct and intuition towards the pros and cons for the different segmentation method. For edge detection, we've chosen the canny method. The SVM (support vector machine) to represent the Classifier method. And the Ostu method to represent the threshold method.

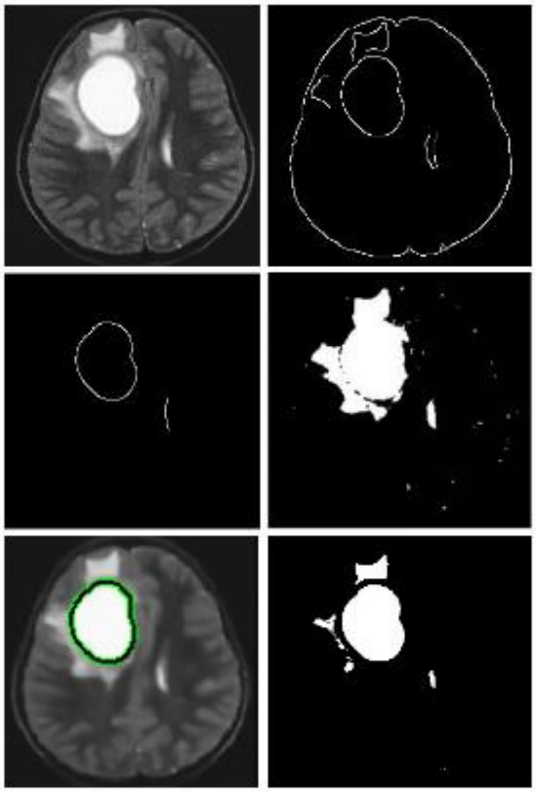


Figure 11. The comparison of different methods: The original image (top-left), Canny with 0.6 threshold (top-right), Canny with 0.9 threshold (middle-left), Otsu (middle-right), SVM (bottom-left), Spatial-FCM (bottom-right).

In the testing image shown above, what we observe is that for a complex medical image, the edge detection really struggles with segmentation task, although with proper settled threshold value, the significant part in the MRI image (e.g. Tumor) can be segmented, but the output is really unstable, what's more, the value has to be manually set, there are no measuring metric to tell whether the detected edge is good or not. The Otsu method has relatively better performance compared to edge detection in this task. But the drawback is also obvious, it cannot get rid of the points that have similar intensity, though they can be totally irrelevant spatially. The SVM as a representative algorithm for classifier-based method, has really promising performance in classifying the image it is trained to. But in medical image segmentation this advantage could be a double-edged sword. To gain a good performance, a fair amount of human labor are needed to collect the training data. But it is hard to generalize for different kinds of medical image, which means for different kinds of medical image, the large amount of human labor are needed again, this sometimes can be unacceptable(e.g. Brain MRI image at different cross-section). Finally, the FCM showed a good performance at different kinds of medical images, the only inconvenience is that the cluster number has to be manually set. But compared to SVM, this step is fairly simple and easy to adjust.

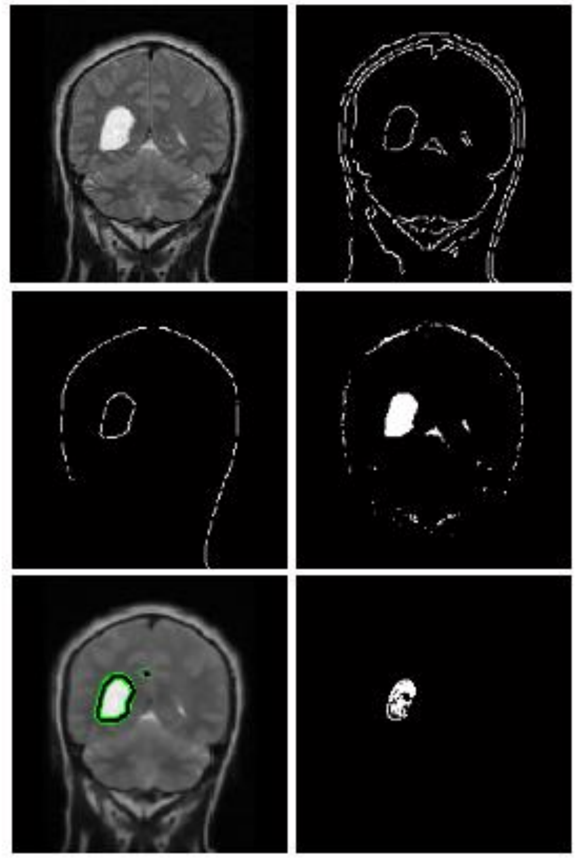


Figure 12. Another tumor image for comparison: The original image (top-left), canny with 0.6 threshold (top-right), canny with 0.9 threshold (middle-left), Otsu (middle-right), SVM (bottom-), Spatial-FCM (bottom-right).

Therefore, in conclusion, for brain MRI image analysis, the FCM is the most promising way, for the advantages described above.

VI. CONCLUSION & FUTURE WORK

Image segmentation is an important low-level preprocessing step for many computer vision problems. For this project, we implemented an efficient, good performance algorithm, the spatial information related fuzzy c-means algorithm for image segmentation. This clustering algorithm incorporates the spatial information into the membership function to improve the segmentation results. The membership functions for the neighborhoods of a centered pixel in the spatial domain are computed to obtain the cluster distribution statistics respectively. These statistics have been transformed into the evolved weighting function and incorporated into the previous membership function. According to the reference, this kind of neighboring effect can reduce the number of spurious blobs and biases the solution toward piecewise homogeneous labeling. Our project implemented the segmentation system, and apply this method on MRI images, and CT images. By make the comparison with other common used segmentation methods, the spatial information related fuzzy c-means algorithm performed excellent.

The advantages of the new method are the following: (1) it yields regions more homogeneous than those of other methods, (2) it reduces the spurious blobs, (3) it removes noisy spots, and (4) it is less sensitive to noise than other techniques. This technique is a powerful method for noisy image segmentation and works for both single and multiple-feature data with spatial information.

In the future work, we could apply this system in more accurate segmentation by choosing the right centroids for clustering, and the clusters of the natural objects can be improved by assign the precise number of clusters. Furthermore, the region of interest can be frequency in spectral domain, or other feature domain, which would generate more applications in all kinds of the fields.

REFERENCES

[1] Suetens, Paul. Fundamentals of medical imaging. Cambridge university press, 2009..

[2] Bovik, Alan C., ed. The essential guide to image processing. Academic Press, 2009.

[3] Bezdek, James C., et al. "Detection and characterization of cluster substructure i. linear structure: Fuzzy c-lines." SIAM Journal on Applied Mathematics 40.2 (1981): 339-357.

[4] Haralick, Robert M., and Linda G. Shapiro. "Image segmentation techniques." Computer vision, graphics, and image processing 29.1 (1985): 100-132.

[5] Gibbs, P., et al. "Tumour volume determination from MR images by morphological segmentation." Physics in medicine and biology 41.11 (1996): 2437.

[6] Bezdek, J. C., L. O. Hall, and L. P. Clark. "Review of MR segmentation technique in pattern recognition." Medical Physics 10.20 (1993): 33-48.

[7] Coleman, Guy B., and Harry C. Andrews. "Image segmentation by clustering." Proceedings of the IEEE 67.5 (1979): 773-785.

[8] Liang, Zhengrong, James R. MacFall, and Donald P. Harrington. "Parameter estimation and tissue segmentation from multispectral MR images." Medical Imaging, IEEE Transactions on 13.3 (1994): 441-449.

[9] Jain, Anil K., and Richard C. Dubes. Algorithms for clustering data. Prentice-Hall, Inc., 1988.

[10] Hebert, Thomas J. "Fast iterative segmentation of high resolution medical images." Nuclear Science, IEEE Transactions on 44.3 (1997): 1362-1367.

[11] Gonzalez, Rafael C. "RE woods, Digital Image Processing." Addison–Wesely Publishing Company (1992).

[12] Held, Karsten, et al. "Markov random field segmentation of brain MR images." Medical Imaging, IEEE Transactions on 16.6 (1997): 878-886.

[13] Chen, Chi Hau, and Gwo Giun Lee. "On digital mammogram segmentation and microcalcification detection using multiresolution wavelet analysis." Graphical Models and Image Processing 59.5 (1997): 349-364.

[14] Gelenbe, Erol, Yutao Feng, and K. Ranga R. Krishnan. "Neural network methods for volumetric magnetic resonance imaging of the human brain." Proceedings of the IEEE 84.10 (1996): 1488-1496.

[15] Reddick, Wilburn E., et al. "Automated segmentation and classification of multispectral magnetic resonance images of brain using artificial neural networks." Medical Imaging, IEEE Transactions on 16.6 (1997): 911-918.

[16] Ahmed, Mohamed N.;Yamany, Sameh M.; Mohamed,Nevin;Farag, Aly A.; Moriarty, Thomas(2002).”A Modified Fuzzy C-Means Algorithm for Bias Field Estimation and Segmentation of MRI Data”(PDF).IEEE Transactions on Medical Imaging 21(3): 193-199.

[17] Sato-Ilic, Mika,and Lakhmi C. Jain.Innovations in fuzzy clustering. Heidelberg: Springer, 2006.

[18] Klawonn, Frank. "Fuzzy clustering: insights and a new approach."?Mathware & soft computing. 2004 Vol. 11 Num.3(2004).

[19] Chuang, Keh-Shih, et al. "Fuzzy c-means clustering with spatial information for image segmentation." computerized medical imaging and graphics 30.1 (2006): 9-15.

[20] Xia, Yong, et al. "Image segmentation by clustering of spatial patterns." Pattern Recognition Letters 28.12 (2007): 1548-1555.

[21] Sahoo P K, Soltani S, Wang A K C, Chen Y C. A Survey of Thresholding Techniques. Computer Vision. Graphics and Image Processing, 1988, 41(2): 233-260

[22] Zhang Y J, Gerbrands J J. Transition Regon Determination Based Thresholding. Pattern Recognition Letter, 1991, 12(1):13-23.

[23] Canny J.A Computational Approach to Edge Detection. IEEE Trans on Pattern Analysis and Machine Intelligence,1986, 8(6);679-698