Unsupervised Segmentation Method for Brain MRI Based on Fuzzy Techniques

Dr. Nasser N. Khamiss
Information Engineering Collage
Nahrain University, Iraq
nasser_alani@yahoo.com

Abstract

In the present research a novel spatially weighted Fuzzy C-Means (FCM) clustering algorithm for image thresholding is presented. The segmentation technique is for magnetic resonance (MR) images of the brain based on fuzzy algorithms for learning vector quantization (FALVQ) by creating of a combined method in utilizing both LVQ (learning vector quantization) and the fuzzy technique. Such a technique is obtaining more efficient method for the process of diagnosis of the human brain tumor without the need for sophisticated steps or human manner. To speed up the FCM algorithm, the iteration is carried out with the statistical gray level histogram of image instead of the conventional whole data of image. Some comparisons with classical thresholding algorithm and fuzzy thresholding algorithm are also considered in this research. Experimental results on real images are given to demonstrate the effectiveness of the proposed algorithm. In addition, due to the neighborhood model, the proposed method is more tolerant to noise.

1. Introduction

Vector quantization is carried out by relating each training vector to a single codebook vector. Thus, the application of the fuzzy c-means to vector quantization should be based on assigning each training vector to the codebook vector, which appears the maximum membership degree with respect to that training vector. However, such a crisp interpretation of the fuzzy c-means during the codebook design may have serious effects on the quality of the final codebook, since this approach hides the existence of outliers and replaces them by their closest codebook vectors. This behavior of the fuzzy c-means justifies the need for the transition from fuzzy (soft) to crisp (hard) decisions, as described in [1]. In the same time, learning vector quantization (LVQ) as introduced by Kohonen constitutes a particularly intuitive and simple though powerful classification scheme [2] which is very appealing for several reasons: the method is easy to implement; the complexity of the resulting classifier can be controlled by the user; the classifier can naturally deal with multiclass problems; and, unlike many alternative neural classification schemes such as feed forward networks and support vector machines, the resulting classifier is human understandable because of the intuitive classification of data points to the class of their closest prototypes. For these reasons, LVQ has been used in a variety of academic and commercial applications such as image analysis, telecommunication, robotics, etc. [3], as shown in **Figure (1)**.

The simplest idea for combining classification and unsupervised learning methods consist of partitioning the feature space using just the feature vectors and labeling each partition using the labels. There are, nevertheless, algorithms that combine properties of clustering and classification algorithms. This work discussed the Learning Vector Quantization methods proposed by Kohonen [4], which are sometimes described as being special cases of Neural Networks. A much simpler interpretation is the following: a Learning Vector Quantizer is just a vector quantizer where centroids have labels, and are iteratively selected to best represent the corresponding classes.

This is done through iterative algorithms that push the centroids towards regions containing numerous samples of the associated class, and away from regions that contain numerous samples of other classes**.**

An interpretation for a family of competitive learning algorithms and investigates their relationship to fuzzy c-means and fuzzy vector quantization are presented in this work. These algorithms map a set of feature vectors into a set of prototypes associated with competitive network that performs unsupervised learning. All algorithms formulas are accomplished by minimizing an average generalized distance between the feature vectors and prototypes using gradient descent method. A closed relationship between the resulting algorithms and fuzzy c-mean is involved. It is also shown that the fuzzy c-mean and fuzzy learning vector quantization algorithms are related to the proposed algorithms if the learning rate for each iteration is selected to satisfy a certain condition. The work also describes the three variants: LVQ1, LVQ2, and LVQ3.

2. Fuzzy algorithms for learning vector quantization

The learning vector quantization is frequently based on the minimization of the functional error [5], where;

$$L_x = (v_1, v_2, ..., v_c)$$

$$L_x = \int\int...\int_{R^n} \sum_{j=1}^{c} u_j(x) \left\| x - v_j \right\|^2 f(x) dx \quad (1)$$

Where *n* is the input vector length and *u* is the member function of clusters *c*. Eq. (1) represents the expectation of the loss function as given below:

$$L_x = L_x(v_1, v_2, ..., v_c)$$

$$L_x = (v_1, v_2, ..., v_c) = \sum_{j=1}^{c} u_j(x) \left\| x - v_j \right\|^2 \quad (2)$$

In the above definitions, $u_j(x), 1 \le j \le c$ represent membership functions of clusters *c*, that regulate the competition between the prototypes $v_j, 1 \le j \le c$ for the input vector ***x***. The specific form of the membership functions determines the strength of attraction between each input and the prototypes during the learning process. The loss function is often defined with respect to the winning prototype. In such a case, the loss Eq. (2) measures the locally weighted error of each input vector with respect to the winning prototype**.**

Pal et al. suggested that the loss function shown in Eq. (1) can be minimized by using the gradient of the instantaneous loss function, Eq. (2), when the probability distribution function is not known. This approach implies the sequential update of the prototypes x with respect to the input vector χ, where $x \in \chi$.

If $u_{ij} = 0, \forall_j \neq i$, then minimization of the loss function, Eq. (2), using gradient descent leads to Kohonen's (unlabeled data) LVQ, which can be used to generate crisp *c* partitions of unlabeled data vectors. According to this learning scheme, only the winning prototype is updated during learning to match the input vector. Because of the inherent bias toward the winning prototype, Kohonen's (unlabeled data) LVQ depends strongly on the initial set of prototypes and is susceptible to local minima.

The development of FALVQ algorithms requires the selection of the membership functions assigned to the prototypes**.** A fair competition among the prototypes is guaranteed if the membership function assigned to each prototype:

1) is invariant under uniform scaling of the entire data set;
2) is equal to one if the prototype is the winner;
3) takes values between one and zero if the prototype is not a winner;
4) approaches zero if the prototype is not a winner and its distance from the input vector approaches infinity.

A variety of FALVQ algorithms can be derived by minimizing the loss function using gradient descent. If ***x*** is the input vector, the winning prototype $\mathbf{v}_j$ can be updated by:

$$\Delta \mathbf{v}_i = \eta(\mathrm{x} - \mathrm{v}_i)\left(1 + \sum_{j \neq i}^{c} \omega_{ij}\right) \quad (3)$$

Where η is the learning rate and

$$\omega_{ij} = \omega\left(\left\|\mathrm{x} - \mathrm{v}_i\right\|^2 \Big/ \left\|\mathrm{x} - \mathrm{v}_j\right\|^2\right) \quad (4)$$

With;

$$\omega(z) = u'(z)$$

Each non-winning prototype $\mathbf{v}_j \neq \mathbf{v}_i$ can be updated as follows (the formula below based on Karayiannis and Pai [6]):

$$\Delta v_j = \eta(x - v_j)\eta_{ij} \quad (5)$$

Where

$$\eta_{ij} = n\left(\left\|x - v_i\right\|^2 \Big/ \left\|x - v_j\right\|^2\right) \quad (6)$$

with;

$$n(z) = u(z) - zu'(z) \quad (7)$$

The update of the prototypes depends on the learning rate $\eta \in [0,1]$, which is a monotonically decreasing function of the number of iterations v. The learning rate can be a linear function of *v* defined as:

$$\eta = \eta(v) = \eta_0(1 - v/N) \quad (8)$$

where η_0 is its initial value of the learning rate, and N the total number of iterations predetermined for the learning process. The above formulation provided the basis for the development of the FALVQ 1, FALVQ 2, and FALVQ 3 families of algorithms [7].

Table (1) shows the membership functions $u(\cdot)$ that generated these families of algorithms and the

corresponding interference functions ω(·) and *n*(·). If ***x*** is the input vector, then the winning prototype v_j is updated by Eq. (3), with ω_{ir} evaluated in terms of the interference function $\omega(\cdot)$ shown in Table (1) as:

$$\omega_{ij} = \omega\left(\left\|\mathrm{x} - \mathrm{v}_i\right\|^2 \Big/ \left\|\mathrm{x} - \mathrm{v}_j\right\|^2\right) \quad (9)$$

The nonwinning prototypes $\mathrm{v}_j \neq \mathrm{v}_i$ can be updated by Eq. (6), with n_{ij} evaluated in terms of the interference function $n(\cdot)$ shown in Table (1) as $n_{ij} = n\left(\left\|x - v_i\right\|^2 \Big/ \left\|x - v_i\right\|^2\right)$.

The update of the winning prototype is affected by the term $\left[1 + \sum_{j \neq i}^{c} \omega_{ij}\right]$ which depends on the number of prototypes. The effect of the number of prototypes on the performance of FALVQ algorithms can be moderated by replacing in the update Eq. (5) the learning rate η by $\eta/\left(1 + \hat{\omega}(c-1)\right)$ where $\hat{\omega} = \hat{\omega}(v)$ increases linearly with the iteration number v from its minimum value $\hat{\omega}_{\min}$ to its maximum value $\hat{\omega}_{\max}$ as:

$$\hat{\omega} = \hat{\omega}_{\min} + v\left(\hat{\omega}_{\max} - \hat{\omega}_{\min}\right)/N \quad (10)$$

The minimum value $\hat{\omega}_{\min}$ of ω can be determined by observing that the interference function ω_{ij} attains its minimum value if the following condition is satisfied:

$$\left\|x - v_j\right\|^2 \approx \left\|x - v_i\right\|^2, \forall j \neq i \quad (11)$$

which is more likely to occur in the beginning of the learning process. It is also reasonable to expect that near the end of learning process, where;

$$\left\|x - v_j\right\|^2 >> \left\|x - v_i\right\|^2, \forall j \neq i \quad (12)$$

That implies that the interference function ω_{ir} attains its maximum value. Typical values for $\hat{\omega}_{\min}$ and $\hat{\omega}_{\max}$ can be obtained from the extreme values of the interference function corresponding to the FALVQ 1 algorithm with α=1. In this case, the following condition:

$$\left\|x - v_j\right\|^2 \approx \left\|x - v_i\right\|^2$$

implies that $\omega_{ij} \approx 1/4$, while the condition

$$\left\|x - v_j\right\|^2 >> \left\|x - v_i\right\|^2$$

implies that $\omega_{ij} \approx 1$.

The resulting FALVQ algorithms can be summarized as follows:

1) Select *c*; fix η_0, *N* ; set *v*=0 ; randomly generate an initial codebook: $\overline{v}_0 = \{v_{1,0}, v_{2,0}, ..., v_{c,0}\}$.
2) Calculate
$\eta = \eta_0\left(1 - v/N\right)/\left(\hat{\omega}_{\min} + v\left(\hat{\omega}_{\min} - \hat{\omega}_{\max}\right)/N\right)$
3) Set $v = v + 1$
4) For each input vector ***x***:
Find i such that:
$$\left\|\mathrm{x} - v_{i,v-1}\right\| < \left\|\mathrm{x} - v_{j,v-1}\right\|^2, \forall j \neq i;$$
Then calculate
for $\forall j \neq i$; $u_{ij,v} = u\left(\left\|x - v_{i,v-1}\right\|^2 \Big/ \left\|x - v_{j,v-1}\right\|^2\right)$;
$\omega_{ij,v} = u'\left(\left\|x - v_{i,v-1}\right\|^2 \Big/ \left\|x - v_{j,v-1}\right\|^2\right)$;
$n_{ij,v} = u_{ij,v} - \left(\left\|x - v_{j,v-1}\right\|^2 \Big/ \left\|x - v_{j,v-1}\right\|^2\right)\omega_{ij,v}$
Update v_i by:
$v_{i,v} = v_{i,v-1} + \eta\left(x - v_{i,v-1}\right)\left(1 + \sum_{r \neq i}^{c} \omega_{ij,v}\right)$
update $\mathrm{v}_j \neq \mathrm{v}_i$ by:
$v_{j,v} = v_{j,v-1} + \eta\left(x - v_{j,v-1}\right)n_{ij,v}$.

5) If $v<N$, then go to step 2.

3. Present work test

This section presented the evaluation of a segmentation technique for MR images of the brain based on FALVQ algorithms. This segmentation was applying a fuzzy clustering technique and used of unsupervised LVQ algorithms. A method presented to produce segmentation system for multiform tumor disease from trans-axial MR image. The method using MATLAB 7.0 and applying the algorithm of FCM of *c* = 14, that is, the segmented images contained fourteen different segments. The segmented image produced by Kohonen's (unlabeled data) LVQ algorithm applied with initial value of the learning rate η= 0.001 and the total number of iterations was N = 20. The tumor and the surrounding edema were clearly

identified by all three algorithms from the FALVQ 1, FALVQ 2, and FALVQ 3 families tested in this work.

The competition between the prototypes during learning and its impact on the performance of FALVQ algorithms were further explored by an additional set of experiments, which evaluated the effect of the free parameter α on the performance of various algorithms from the FALVQ 1 family.

Combined by applying morphology operation as filter threshold that is, the low threshold was *lo*=100 and height threshold *hi*=255, and hole and filling by area open instruction of pixels intensity *bw*=150.

Finally, it is discriminate between normal tissues and abnormalities obtaining image that contain tumor only, see **Figure (2)**.

4. Segmentation System Result

Figure (3) shows a real case of the T1-weighted MR image of multiform intracranial tumors from RADIATION AND NUCLEAR MEDICINE CENTER which is captured as a (jpg) formula by means of especial interfacing to the proposed system.

In this work The MR image was segmented, algorithms from the FALVQ 1, FALVQ 2, and FALVQ 3 families were used. The feature vectors were formed using the pixel values of the T1-weighted.

In this study, c = 14, that is the segmented image contained fourteen different segments. **Figure (4)** shows the segmented image of the T1-weighted MR image produced by the algorithm of the FALVQ1 family corresponding at α = 1 (competition measures: A_u = 0.306; C_u = 0.629).

Figure (5), shows the segmented image of the T1-weighted MR image produced by the algorithm of the FALVQ 2 family corresponding to β=1 (competition measures: A_u = 0.264; C_u = 0.608).

Figure (6) shows the segmented image of the T1-weighted MR image produced by the algorithm of the FALVQ 3 family corresponding to γ_1=1 (competition measures: A_u = 0.167; C_u = 0.5).

In all of these experiments the total number of iterations was N = 20 and the initial value of the learning rate was η_0 = 0.001.

The tumor and the surrounding edema were clearly identified by all three algorithms from the FALVQ 1, FALVQ2, and FALVQ 3 families tested in this work.

The tumor boundaries were not properly captured therefore the T- threshold filter was applied to the tumor of segmentated image (keeping only those pixels whose T pixels intensity was greater than the T-threshold). The resultant image was considered the final tumor segmentation. Threshold filter was a relatively coarse manner because the boundary of enhancing tumor was obscured by pixels belonging to non-tumor tissues. With the removal of most of these non-tumor tissues, a greater level of focus can be placed and a more precise threshold can be applied. The threshold filter was determined using the principle that the spatial boundary between enhancing tumor and surrounding tissues contain pixels whose their intensities correspond to the tumor/non-tumor boundary. The threshold filter used was determined from training slice including in LVQ1 and LVQ3 low threshold =100, and high threshold =255 using only pixels contained in the initial tumor segmentation. While in FALVQ2 low threshold =130, and high threshold =255. **Figure (7)** shows the results of applying the T threshold filter. Segmentation image after threshold filter had detected tumor need not to be perfect, merely sufficient to indicate the appropriate intensity. Edge detected must still be reasonable, however, for the method to work. Then applying morphology operation helps to find the final image with tumor by using hole and filling method, as shown in **Figure (8).**

5. Conclusions

The presented work of image segmentation is built on previous results, suggesting a more systematic way of classifying the brain tissues. A clear emphasis is put on the detection of tumors or lesions related to demyelination of white matter. This segmentation approach is simple and easily implementable, while the use of unsupervised LVQ algorithms does not rely on prior information provided by human experts. It is remarkable that the performance of FALVQ algorithms degraded considerably in the limit where their behavior approaches that of Kohonen's (unlabeled data) LVQ, which allows only the winning prototype to be updated. The experiments revealed the ability of all Generalized FALVQ1, LVQ2 and LVQ3 algorithms tested to discriminate between normal tissues and abnormalities. LVQ1 and LVQ3 are more "robust" than LVQ2, which is probably why they are included with the available package used for the class. The suggestion is to use LVQ2 only for little iteration and with small learning rate.

References

[1] George E. Tsekouras, "A fuzzy vector quantization approach to image compression", Applied Mathematics and Computation, pp 539–560, 2005.

[2] Teuvo Kohonen, "Self-organizing maps", Springer, Berlin, New extended edition, ISBN 3-540-67921-9, ISSN 0720-678X, 2001.

[3] Michael Biehl, Barbara Hammer, Petra Schneider, "Matrix Learning in Learning Vector Quantization", Technische Universit?t Clausthal, **ISSN:** 1860-8477, 2008.

[4] Teuvo Kohonen, J. Hynninen, J. Kangas, J. Laaksonen, and K. Torkkola. **"LVQ PAK: The learning vector quantization program package"**, Report A30, Helsinki University of Technology,1996.

[5] N. R. Pal, J. C. Bezdek, and E. C.K. Tsao, **"Generalized clustering networks and Kohonen's self-organizing scheme"** IEEE Trans. Neural Networks, vol. 4, pp. 549–557, Apr. 1993.

[6] R. J. Hathaway and J. C. bezek," Optimization of clustering criteria by reformulation, Fuzzy Systems, IEEE Transactions on signal processing, Volume 3, Issue: 2, ISSN: 1063-6706, May 1995.

[7] Nicolaos B Karayiannis, Pin I Pai, "A family of fuzzy algorithms for learning vector quantization", International journal of engineering intelligent systems for electrical engineering and communications, volume: 7, pp. 145-155, 1999

Table (1) Membership functions and interference functions for the FALVQ 1, FALVQ 2, and FALVQ 3 algorithms families

FALVQ Family	$u(x)$	$\omega(x)$	$n(x)$
FALVQ1 $(0<\alpha<\infty)$	$x(1+\alpha x)^{-1}$	$x(1+\alpha x)^{-2}$	$\alpha x^2(1+\alpha x)^{-2}$
FALVQ 2 $(0<\beta<\infty)$	$x\exp(-\beta x)$	$(1-\beta x)\exp(-\beta x)$	$\beta x^2\exp(-\beta x)$
FALVQ 3 $(0<\gamma_1<1)$	$x(1+\gamma_1 x)$	$1-2\gamma_1 x$	$\gamma_1 x^2$

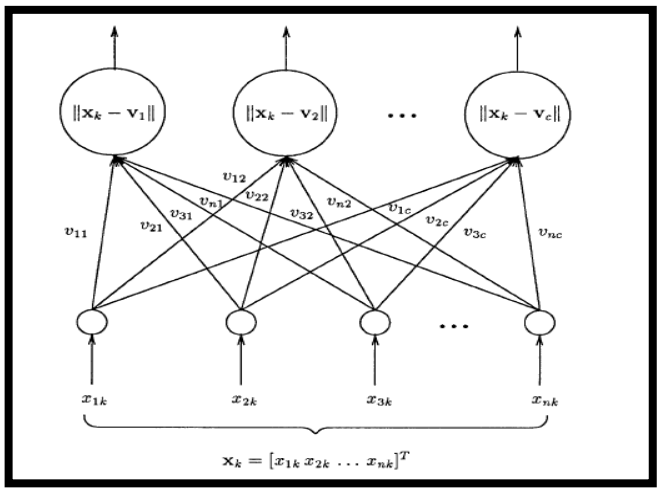


Fig. (1) The LVQ competitive learning network

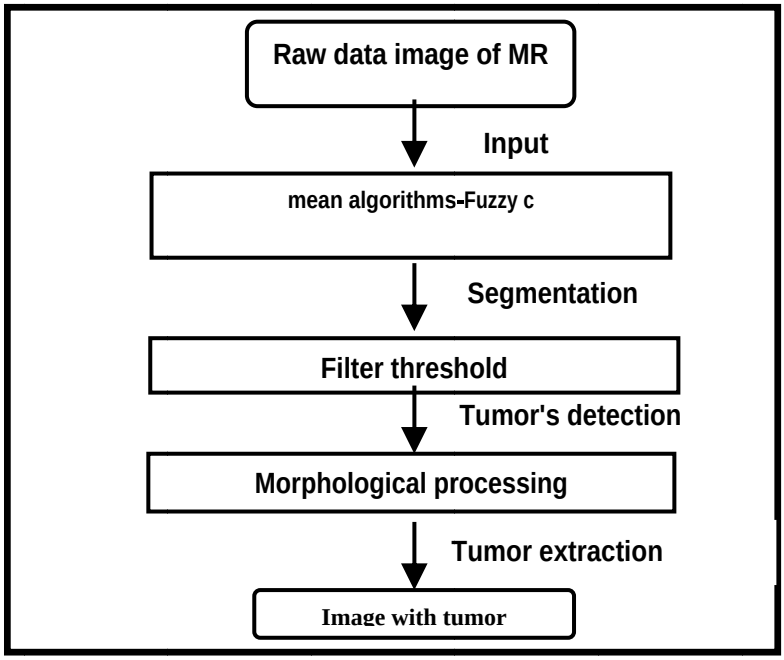


Fig. (2) Outline of the architecture of segmentation system

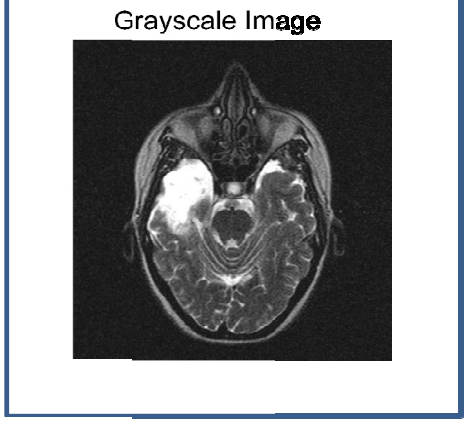


Figure (3) shown the T1-weighted MRI

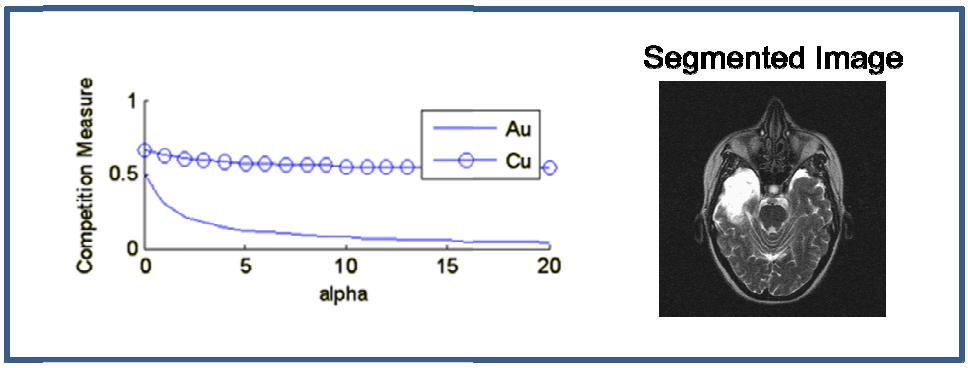


Figure (4) T1-weighted MRI shown the segmented image produced by the FALVQ1

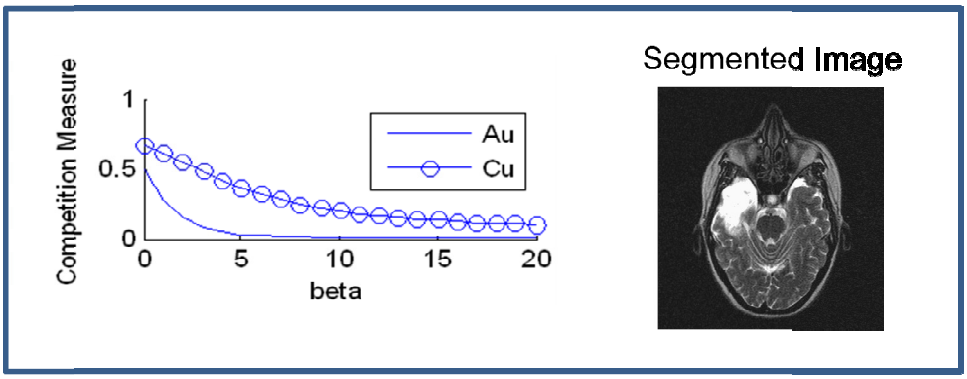


Figure (5) T1-weighted MRI shown the segmented image produced by the FALVQ2

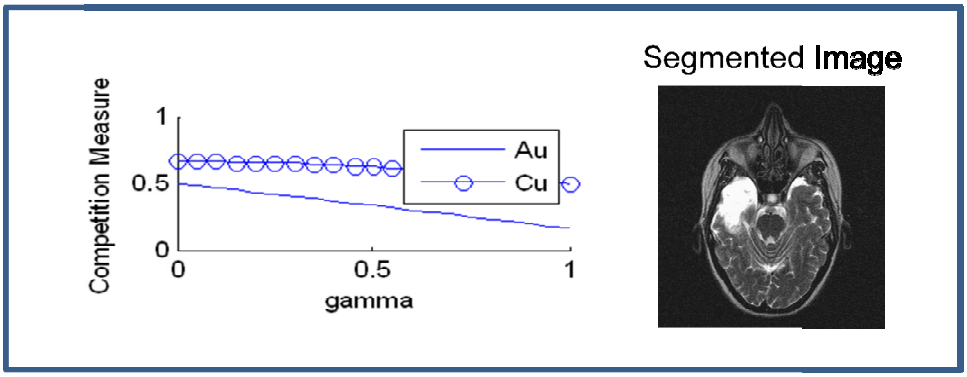
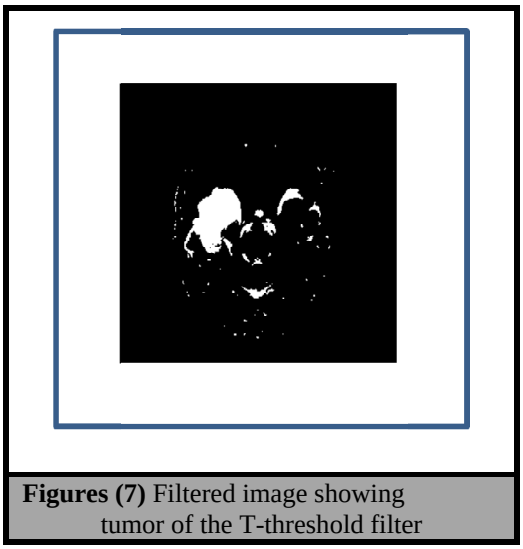


Figure (6) T1-weighted MRI shown the segmented image produced by the FALVQ3



Figures (7) Filtered image showing tumor of the T-threshold filter

Figure (8) shows the image with brain tumor

طريقة الانقسام الآلي لصور رنين مغناطيسي للدماغ باستخدام التقنيات الضبابية

د. نصر نافع خميس
جامعة النهرين/ كلية هندسة المعلومات
بغداد / العراق

الخلاصة

في هذا البحث تم تقديم خوارزمية جديدة لايجاد التجميعات في صور الرنين المغناطيسي للدماغ بأستخدام الشبكات العصبية العشوائية . إن تقنيات أيجاد التجمعات في صور الرنين المغناطيس للدماغ المبني على خوارزمية Fuzzy لتدريب المتجة الكمي من خلال الاستفادة من طريقة المتجهة الكمي وتقنية الـ Fuzzy معاً. هذه الطريقة أعطت نتائج كفؤة لتشخيص الحالات المرضية لسرطانات للدماغ البشري بدون الرجوع الى معالجات إضافية أو تدخل بشري . ولتقليل الوقت المطلوب بتهيئة المعلومات ، فأن جزء من المعالجات أنجزت خارج الخوارزمية .

أعطى البحث بعض المقارنات مع الطرق العادية ، حيث النتائج على صور حقيقية أوضحت كفاءة الخوارزمية المقترحة مع الملاحظة بأمكانية التعامل مع الضوضاء في الصور