International Journal of Soft Computing 1(3): 232-238, 2006

A Neuro-Fuzzy Image Segmentation Alogrithm

Abdelouahab Moussaoui
Department of Computer Science, Faculty of Engineering Sciences,
Ferhat Abbas University, Setif, 19000 Algeria

Abstract: To segment automatically a cerebral RMN image in a reliable and robust manner is a delicate problem. However, it is a task that the expert could do with a certain precision, if he had one unlimited time. The expert has a very precise priori knowledge of what he wishes to segment. He knows the general shape and the disposition of the objects that he must segment. He can use principles logical of "common sense". This type of knowledge is integrated, to different degrees and often of very implicit manner, in all methods of segmentation. These methods are generally founded on very different principles as for example the processes of classification, the use of models of deformable contours or models of knowledge. But these last don't use in general that only one volume of image what causes in most cases a considerable loss of information. Besides these methods tend to be dependent of some parameters as the number of classes, the functions of similarities that are often imposed what implies that the application of these techniques becomes non suitable to the multimodal medical images. The goal of our work is to show that it is possible to define a common setting of work to permit the implementation of cooperation between heterogeneous approaches. The interest of such an approach is to be able to exploit the complementarity of information that results from the application of several methods in order to propose a complete system of segmentation. For it we developed one automatic classification algorithm baptized ONFPCM (Optimal Neuro-Fuzzy Possibilist C-Means) using a supplementary stage of Bootstrap by the slant of an auto-organizing neuron network algorithm named CENN (Capture Effect Neural Network). In this study a number of classic automatic classification methods as the HCM algorithm (Hard C-Means) and FCM (Fuzzy C-Means) are elaborated as well as the NFMEP algorithm (Neuro-Fuzzy Maximum Entropy Principle). These last are applied on multimodal medical images and the results are compared to those gotten by the ONFPCM algorithm.

Key words: Fuzzy clustering, PCM algorithm, MEP algorithm, automatic classification, image segmentation

INTRODUCTION

In the clinical domain, the concept of the fuzzy diagnosis is applied extensively. It can stand between the supervised classification of several clinical cases with different classes of pathologies, through a set of linguistic rules and numerical data and the non supervised segmentation of medical images, through methods based on data analysis concerning numerical variables spatially distributed[1,2]. In all these cases, the uncertainty is present in the data as well as in the rules and it suggests that the fuzzy methods can be applied with success thanks to their intrinsic flexibility. In this study we concentrate our attention on the methods for the segmentation of multimodal medical images[3,4].

The segmentation can be described as the definition of groupings, in the multimodal parametric space, where the points are associated to the different sets of values of similar intensities in the different images. As consequence, in this approach, the process of grouping is the main step in the segmentation procedure[5-9] and the techniques based on the automatic grouping are reputed to be more robust for the separation of the different tissues, in presence of noise and imprecise data, in relation to the techniques of contours detections[10,11,5].

Besides, the uncertainty is largely present in the medical images because of the noise (during the acquisition process) and of the effects of partial volumes. It means that the values of the voxels, especially to the borders between volumes of interest, correspond to the miscellanies of different anatomical tissues, because of the low resolution of the sensors. As consequence, the borders between tissues are not defined correctly and the adherences in the limits of the regions are intrinsically fuzzy[12,13].

According to all these considerations, our choice for the segmentation of the anatomical tissues carried itself on the methods of data analysis whose principle is founded mainly on the fuzzy automatic grouping by progressive and iterative refinement.

IMAGE CLASSIFICATION

We consider a multimodal volume resulting of a set of S different volumes of images. Their voxels are associated to an array of S values; each represents the intensity of only one image volume. In other words, the S different values of intensities are joined to all voxels in which every multimodal volume can be seen as the coordinates of the voxels in a parametric space of S dimension where a multimodal analysis can be operate.

Two differential spaces are considered however for the continuation in the description of the segmentation problem:

- an image space (generally 3D) definite from the spatial coordinates of the set of data and;
- a multimodal parametric space, that is described subsequently.

Except the C-Means algorithm (CM)[14,15] and the algorithm based on the neuron network called Capture Effect Neuron Network (CENN)[16], we made several experiences with a number of fuzzy approaches of automatic classification for the multimodal medical images segmentation. These methods seem to give good results of automatic classification and they are not affected by the dimension of the parametric space.

However, the fuzzy methods of automatic classification produce a classification of voxels, in relation with the adherence function of the different classes and can bring an effect of smoothing to the voxels classification. It can drive to a cleaner definition of the surfaces of the anatomical objects by the segmentation.

The first studied method is based on the neuron network of where we applied it to multimodal medical images of different dimensions.

The second is the classification method baptised ONFPCM (Optimal Neuro-Fuzzy Possibilistic C-Means) derived from Masulli's method[7,12] and in which we used a metrics that takes in consideration the particularity of the multimodal images[17,2,4]. A cooperation between the previous method and the one of PCM is put in evidence.

The third is the fuzzy automatic classification method based on the maximum entropy principle (MEP)[23,24,8]. It doesn't require any prior assumptions about the number of classes. The methodology and the results of this implementation algorithm, concerning the multimodal medical images segmentation, are described in[4,2].

In the following section, the CENN algorithm is described and its results are used to launch the developed algorithm.

CENN ALGORITHM

The CENN[16,12] is an auto organizing neural network capable to take in consideration the local features of the point's distributions (grouping by adaptive resolution). CENN combines the standard competitive auto-organization of the weight vectors[18] with the non linear mechanism of the adaptive local modulation of the receptive fields (RFs) of the neurons (Capture Effect). The training phase of the CENN consists in a training step by calculating data vector quantification and the labeling step where the prototypes, gotten while executing the previous stage, are grouped in order to get robust classes.

In the training phase, a set of neurons $n^i = \{w_i, r_i\}$ is initialized with random weights vector w_i (representing the centers of the sub-classes) and with a large ray $r_i(r_i = R_0)$ of the receptive fields Rf_i (modeled by a Gaussian function γ). Next, the set of the data is presented to the CENN and the following training formula is applied:

$$\Delta w_i = \eta_w (x_k - w_j) \frac{\gamma(d_j(x_k))}{\sum_\ell \gamma(d_\ell(x_k))} \tag{1}$$

$$\Delta r_i = \begin{cases} \eta_r \left(d_i(x_k) - r_i\right) \exp\left(-d_i(x_k)/p\right) \\ 0 \quad \text{if } d_i(x_k) \geq R_0 \end{cases} \tag{2}$$

where η_w and η_r are the training rates, $d_i(x_k) = \|x_k - w_i\|$ is the Euclidian distance of the points to the weights vector and the parameter p is defined with:

$$p \equiv \frac{\langle d_i(x_k) \rangle}{D \ln 10}, \tag{3}$$

knowing that D is the dimension of the attributes space.

The labeling step abandons all neuron n_q with $n_q = R^0$ (i.e., neuron not being part of the elements of the training set) and their pairs of neurons, n_p and n_q must receive the same label (i.e., their associated classes are merged) if:

$$\left\| w_p - w_q \right\| < \left(r_p + r_q\right)\sigma, \quad \sigma \in (0,1), \tag{4}$$

ONFPCM ALGORITHM

In[19,20], Krishnapuram and Kellers present two versions of the Probabilistic C-Means algorithm (PCM) who relax the probabilistic constraint, to permit to assimilate the possibilistic interpretation of the adherence function to a typicality degree.

Let $X = \{x_k / k = 1,..., n\}$ the set of data to classify;

$Y = \{y_j / j = 1,..., c\}$ the set of the centers of classes (or prototype) and $U = [u_{jk}]$ the matrix of the membership degrees. In the PCM algorithm, the elements of U fill the following conditions:

$$u_{jk} \in [0,1] \quad \forall j,k;$$

$$0 < \sum_{k=1}^{n} u_{jk} < n \quad \forall j; \tag{5}$$

$$0 < \sum_{k=1}^{n} u_{jk} < n \quad \forall j;$$

PCM-II[19] is based on the modification of the objective function of CM[15] in order to avoid the determination of the fuzzy parameter. The objective function of PCM-II contains two terms; the first is the objective function of CM[15], whereas the second is the regulating term. The points that possess elevated typicality degrees have an important value u_{jk} and the non representative points have a value u_{jk} less important in all classes:

$$J_m(U,Y) = \sum_{j=1}^{c}\sum_{k=1}^{n} u_{jk} E_j(x_k) + \sum_{j=1}^{c}\xi_j \sum_{k=1}^{n}\left(u_{jk}\log u_{jk} - u_{jk}\right), \tag{6}$$

where $E_j(x_k) = \|x_k - y_j\|^2$ is the square of the Euclidian distance and the parameter ξ_j depends on the distribution of the points in the j^{th} class and must be assigned before the classification procedure.

As demonstrated already by Krishnapuram and Keller[20], (U, V) minimize J_m under the constraints (1-3) if and only if:

$$u_{jk} = \exp\left\{\frac{E_j(x_k)}{\xi_j}\right\} \quad \forall j,k, \tag{7}$$

$$\text{et } y_j = \frac{\sum_{k=1}^{n} x_k u_{jk}}{\sum_{k=1}^{n} u_{jk}} \quad \forall j. \tag{8}$$

A classification algorithm of bootstrap can be executed before, throwing the PCM, in order to get an initial distribution of the prototypes in the parametric space and thus to estimate some parameters used in this algorithm. While considering the CENN like bootsrap for the PCM[1,19,20,12], the following definition of ξ_j is initially used:

$$\xi_j = \frac{\sum_{k=1}^{n} (u_{jk})^m E_j(x_k)}{\sum_{k=1}^{n} (u_{jk})^m}, \tag{9}$$

where $m \in (1, +\infty)$ is the fuzzy parameter used by FCM and K is a normalization parameter. This definition makes that ξ_j is proportional to the average value of the deviation between the classes and depends strongly on the choice of k (in [12] they suggest that k = 1).

In the next optional refinement step, a second definition of ξ_j is used:

$$\xi_j = \frac{\sum_{xk \in (\pi_j)\alpha} E_j(x_k)}{\left|(\pi_j)\alpha\right|}, \tag{10}$$

where $(\pi_j)_\alpha$ is the set of the points of the j^{th} class of which the membership function is superior to a given threshold α (α- cut).

In the version defined by Masulli and al[7,9,21] the considered distance is the Euclidian distance since we suppose that the densities are not variable and that the distributions of the different objects have an equitable probability. What is not always true, especially for multimodal data and in presence of noise, like in the case of the medical images. It is why we developed a modified version that takes in consideration all these constraints. For that, we leaned on the works of Gath and Gevas[22].

The algorithm of Gath and Geva[22] is distinguishable of the PCM algorithm by use of an exponential distance measure based on the maximum likelihood. This distance measure permits to generate the hyper-ellipsoids classes that can vary at a time by their shapes and their densities.

The first stage consists to partition the data by the Possibilistic fuzzy C-Means algorithm (PCM).

The second stage consists on the one hand in initializing the centers of the classes by the prototypes generated by PCM algorithm during the first stage to improve the partition and on the other hand to affect a second partition while applying a method named Fuzzy Maximum Likelihood Estimation (FMLE).

The third phase consists in iterating the process that means partitioning the data by FCM and then by the FMLE method while incrementing the number of classes and while calculating, at the time of every iteration, the values of the performance measures proposed by Gath and Geva[22]. The used metrics is a given exponential distance by:

$$d_{ik}^2 = d_e^2(x_k - v_i) = \frac{|F_i|^{1/2} \exp\left(\frac{(x_k - v_i)^T F_i^{-1} (x_k - v_i)}{2}\right)}{p_i} \quad (11)$$

where $|F_i|$ represents the determinant of the fuzzy covariance matrix F_i. P_i is the priori probability to select the class c_i and $h\ (i / x_k)$ is the posteriori probability (i.e., probability to select the class c_i knowing the data x_k).

$$h(i/x_k) = \left[\sum_{j=1}^{C}\left(\frac{d_{ik}^2}{d_{jk}^2}\right)\right]^{-1} \quad (12)$$

$$F_i = \frac{\sum_{k=1}^{C} h(i/x_k)(x_k - v_i)(x_k - v_i)^T}{\sum_{k=1}^{n} h_i(i/x_k)} \quad (13)$$

$$p_i = \frac{\sum_{k=1}^{C} h_i(i/x_k)}{n} \quad (14)$$

DEVELOPED ALGORITHM (ONFPCM)

- Bootstrapping by CENN
  - Initialize the weights vectors $n_i = \{w_i, r_i\}$ randomly.
  - launch and label the CENN in order to obtain the parameters c, y^o_j and ξ (eq. 1 et 2);
  - Compute $U^{(0)}$ (eq. 5);
  - initialize the iteration meter ($\ell \leftarrow 0$) and the threshold ε for the stop test;
- OPCM Algorithm
  - Fix d^2_{ik} (Gath et Geva)
  - Compute h (I / x_k), p_i et $|F_i|$,
- Repeat
  - Update the prototypes $y_i^{(\ell+1)}{}_j$ (eq. 7 et 8)
  - Compute $U^{(\ell+1)}$ (eq. 5);
  - increment ℓ ($\ell \leftarrow \ell + 1$);
- Until

$$\nabla_k \left\| y_k^{\ell+1} - y_k^{\ell} \right\| \le \varepsilon$$

- Refinement by FMLE
- Compute the performance measures.
- Increment the class number and return to 2 until having the optimal value of the performance measure.

NFMEP ALGORITHM

The principle of the entropy maximization proposed by Jaynes[23] has been applied to the fuzzy clustering by Rose, Gurewitz and Fox[20,4].

Let p_{jk} the probability distribution of k^{the} attribute of the training set to the j^{th} class calculated in relation of the Shannon entropy set Ω.

While maximizing H under the constraints of the equations (6) and (2), the resulting membership functions of the classes' samples are Gibbs distributions of[23]:

$$u_{jk} = \frac{e^{-\beta E_j(x_k)}}{Z_k} \quad (15)$$

where

$$Z_k = \sum_{l=1}^{c} e^{-\beta E_l(x_k)} \quad (16)$$

is a normalization factor named Partition Function. Of the point of seen of the statistical mechanics the multiplier of Lagrange β is interpreted as the inverse of the temperature T.

$$\left(\beta = \frac{1}{T}\right) \quad (17)$$

Besides, its can be interpreted like a fuzzy parameter control. When β increase the associations of the samples with their respective classes become less fuzzy.

The limits cases are:

- for $\beta \rightarrow 0^+$ we have $\mu_{jk} = \frac{1}{c}$ for all j,k, i.e., every sample is associated in an equitable manner to every cluster;
- for $\beta \rightarrow \infty$ we have $\mu_{jk} = 0$ if the sample belongs to the cluster j and $\mu_{jk} = 0$ for all $I \neq j, i \in [1, c]$, i.e., every sample is associated to only one cluster (hard limit).

We define the efficient mistake (as named the free energy, by analogy to the statistical mechanics) by:

$$F = -\frac{1}{\beta} \ln Z \quad (18)$$

where

$$Z = \prod_k Z_k \quad (19)$$

is named Total Partition Function.we can easily demonstrated that

$$F = -\frac{1}{\beta} H_{max} + \langle E \rangle \quad (20)$$

and from there

$$\lim_{\beta \to \infty} F = \langle E \rangle \quad (21)$$

This limit allows us to find the minimization solution under constraints of $\langle E \rangle$ while performing the deterministic annealing algorithm on F proposed by Rose, Gure-witz and Fox[24,8]. This method that we name fuzzy clustering by the maximum entropy principle (MEP-FC), start while minimizing F in a high temperature T, for which its exists an unique solution $\mu_{jk} = \frac{1}{c}$ for all j,k and we reduce progressively T until the hard limit is reached.

In the algorithm used here, we consider, as proposed in[19,20,21], $E_j(x_k) = \|x_k - x_j\|^2$ and for every value of β we calculate the minimization of F while considering Y by iteration of the following formula:

$$y_j = \frac{\sum_{k=1}^{n} u_{jk} x_k}{\sum_{k=1}^{n} u_{jk}} \quad (22)$$

RESULTS AND DISCUSSION

We have implement the two algorithms ONFPCM and NFMEP as well as the NFPCM algorithm (Neuro-Fuzzy Possibilist C-Merans) developed by Mussuli and al[7]. The algorithms are tested on resonance magnetic images (weighted in T1, T2 and in density of proton) in view of an automatic segmentation.

The first set of data represents a cerebral cut of a normal individual and the goal is to separate the white substance, the gray substance, the cerebro-spinal liquid, the encephalon and other structures in order to discover the possible morphological anomalies.

The second set of data represents the cut of a brain human of an individual reaches a meningioma and the goal is to delimit by contour the tumour localized in the right frontal lobe and it seen some separating correctly the tumour of the healthy tissues.

In any case the fusion of the different volumes produced a data set of dimension three (multimodal). Every triplet of voxel intensity in the multimodal data is represented by one point in an attribute space of dimension three where the different coordinates represent the respective intensities following every axis.

For the normal individual, the original images have a dimension of 220x220 pixels of 256 levels of gray. The unsupervised segmentation with NFPCM and ONFPCM is shown to the Fig. 1. The CENN algorithm finds automatically five classes with in general a good performance but with a clean overestimate of the white substance.

Table 1:Comparison of the results gotten by different algorithms

	Pcm	Nfpcm	Nfmep	Onfpcm
Tumor	0.31	0.59	0.74	0.83
WM	0.52	0.83	0.94	0.91
GM	0.40	0.76	0.63	0.88
CSL	0.55	0.79	0.95	0.96

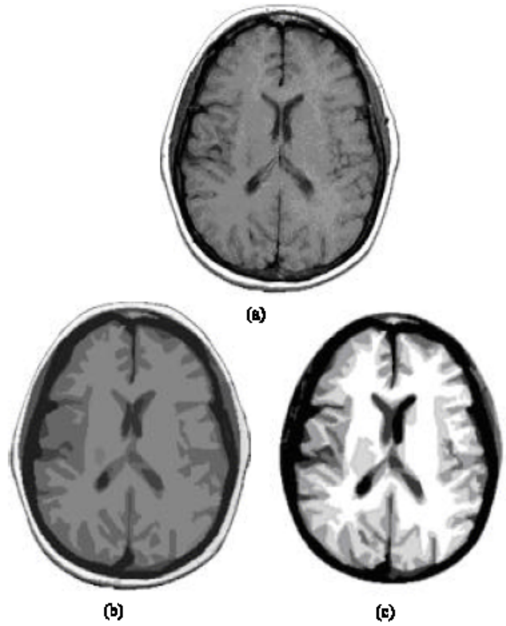

Fig. 1: Segmentation of RMN image (a) by (b) NFPCM and © ONFPCM of normal patient

In comparison with the PCM and NFMEP algorithm, on the same images (Table 1 and Fig. 1), a large and bad classification of the structures (close to the eyes) is well visible, on the other hand we notice well a very good separation of the gray and white matter.

The NFPCM and ONFPCM algorithms pull profit of the algorithm of CENN used like bootstrap what gives a very good classification notably of the cerebral structures in relation to the PCM and NFMEP algorithms (Table 1).

In the Fig. 2 we show the efficiency of the application of the NFPCM and ONFPCM in relation to the PCM and NFMEP in different conditions on anatomical images of an individual reaches a meningioma.

In end, we note that the ONFPCM algorithm gives a percentage of recognition of the a lot more important cerebral tissues than the one of Masulli (NFPCM) safe with regard to the white substance where the NFPCM algorithm remains the most effective.

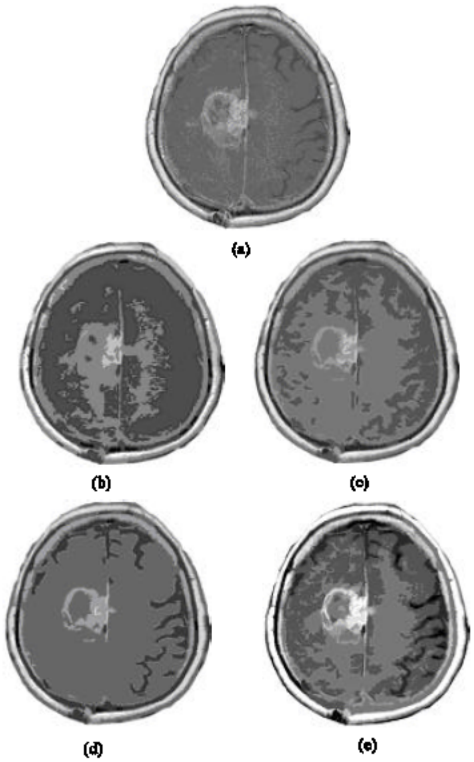

Fig. 2: Pathological image (a) and its segmentation gotten by (a) PCM; (b) MEP; © NFPCM et (d) ONFPCM (patient with meningioma)

CONCLUSION

For the diagnosis of pathologies through the treatment of medical images, the approaches based on the techniques of automatic classification (clustering) [5,11,17,24,25] show a hardiness as for the discrimination of the regions in relation to the techniques based on the detection of the contours[5,15,23], because of the weak ratio signal/bruit that characterizes the majority of the data of medical images. Besides, the fuzzy approaches of automatic classification seem to be especially interesting because, due to the effect of the partial volume during the acquirement, the value of the voxel to the borders between volumes of interest corresponds to the miscellanies of different anatomical tissues.

The probabilistic approach for the automatic classification[13,14], processed the degree of adherence of a voxel to a particular tissue as degree of typicality is especially useful. In the implementation described in this paper by the NFPCM, NFMEP and ONFPCM algorithms, we combine a bootstrap based on the CENN algorithm[7] with the second version of the PCM-II[14], with a simple heuristic of fusion of redundant classes and the MEP algorithm with a very big initialization of T[21].

The algorithms have been applied on different types of medical images (IRM) and resulted prove their efficiency as segmentation of the cerebral tissues and/or of the timorous tissues are shown respectively in Fig. 1 and Fig. 2.

REFERENCES

1. Barni, M., V. Capellini and A. Mecocci, 1966. Comments on a possibilistic approach to clustering, IEEE Trans Fuzzy Syst, pp: 85-93.
2. Liew, A.W.C and Hong Yan, 2003. An adaptive spatial fuzzy clustering algorithm for 3-D MR image segmentation, IEEE Transactions on Medical Imaging, pp: 1063-1075.
3. Gerig, G., J. Martin, R. Kikinis, O. Kubler, F.A. Shenton and M. Jolesz, 1992. Unsupervised tissue type segmentation of 3D dual-echo MR head data, Im Vis Comput, pp: 49-60.
4. Maisey, M.N. and al, 1992. Synergistic imaging, Eur J. Nucl. Med., pp: 2-5.,
5. Bezdek, J.C., L.O. Hall and L.P. Clarke, 1993. Review of MR image segmentation techniques using pattern recognition, Med Phys, pp: 33-48.
6. Karayiannis, N.B., 1997. A methodology for construction of fuzzy algorithms for learning vector quantization, IEEE Trans Neural Netw, pp: 5-18.
7. Masulli, F., P. Bogus, A. Schenone and M. Artuso, 1997. Fuzzy clustering methods for the segmentation of multivariate images, Proceedings of the Seventh International Fuzzy Systems Association Word Congress IFSA '97, Prague: Academia, pp: 123-128.
8. Rose, K., E. Gurewitz and G. Fox, 1993. Constrained clustering as an optimization method, IEEE Trans Pattern Anal Mach Intl., pp: 85-94.
9. Schenone, A., F. Firenze, F. Acquarone, M. Gambaro, F. Masulli and L. Andreucci, 1996. Segmentation of multivariate medical images via unsupervised clustering with adaptive resolution, Comput Med Imaging Graph, pp: 1929.
10. Beni, G and X. Liu, 1994. A least biased fuzzy clustering method, IEEE Trans Pattern Anal Mach Intel, pp: 54-60.
11. Bensaid, A.M., L.O. Hall, L.P. Clarke and R.P. Velthuizen, 1991. MRI segmentation using supervised and unsupervised methods, In: Nagel JH and Smith WM, editors. Proceedings of the Thirteenth IEEE Engineering in Medicine and Biology Soc. Conference. New York: IEEE, pp: 483-489.

12. Masulli, F. and A. Schenone, 1999. A fuzzy clustering based segmentation system as support to diagnosis in medical imaginh, Artificial intelligence in Medicine, pp: 129-147.
13. Sea Chen, C.A Bouman and M.J. Lowe, 2004. Clustered components analysis for functional MRI, IEEE Transactions on Medical Imaging, pp: 85-98.
14. Bezdek, J.C., 1981. Pattern Recognition with Fuzzy Objective Function Algorithms, New York: Plenum Press.
15. Duda, R.O. and P.E. Hart, 1973. Pattern Classification and Scene Analysis, New York: Wiley.
16. Firenze, F. and P. Morasso, 1993. The capture effect model: A new approach to self-organized clustering, In: Proceedings of the Sixth International Conference on Neural Networks and their Industrial and Cognitive Applications and Exhibition Catalog, NEURO-NIMES 93 Conference. Nimes (France) pp: 45-54.
17. Masulli, F., P. Bogus, A. Schenone and M. Artuso, 1998. Application of MEP-based fuzzy clustering to the segmentation of multivariate medical images, Proceedings of the International Joint Workshop on urrent Issues on Fuzzy Technologies:Methods and Environments for Planning and Programming. Singapore: World Scientific, pp: 266-277.
18. Kohonen, T., 1995. Self-Organizing Maps. Berlin, Springer–Verlag.
19. Krishnapuram, R. and J.M Keller, 1993. A possibilistic approach to clustering, IEEE Trans Fuzzy Syst, pp: 98-110.
20. Krishnapuram, R. and J.M. Keller, 1996. The possibilistic c-means algorithm: Insights and recommendations, IEEE Trans Fuzzy Syst, pp: 85-93.
21. Schenone, A., F. Masulli and M. Artuso, 1997. A neural bootstrap for the possibilistic c-means algorithm, In:Morabito FC, editor. Advances in Intelligent Systems. Amsterdam IOS Press, pp: 59-66.
22. Gath, I. and A.Z.B. Geva, 1989. Unsupervised ptimal fuzzy clustering A, IEE transaction. P.A.M.I, pp: 59-66.
23. Jaynes, E.T., 1957. Information theory and statistical mechanics, Phys. Rev., pp: 20-30.
24. Rose, K., E. Gurewitz and G. Fox, 1990. A deterministic annealing approach to clustering, Pattern Recognit Lett., pp: 89-94.