

Fuzzy multi-channel clustering with individualized spatial priors for segmenting brain lesions and infarcts

Evangelia I. Zacharaki^{1*}, Guray Erus², Anastasios Bezerianos¹, Christos Davatzikos²

¹Department of Medical Physics, School of Medicine, University of Patras, Patras, Greece

²Department of Radiology, University of Pennsylvania, Philadelphia PA, USA

* (ezachar@upatras.gr)

Abstract. Quantitative analysis of brain lesions and ischemic infarcts is becoming very important due to their association with cardiovascular disease and normal aging. In this paper, we present a semi-supervised segmentation methodology that detects and classifies cerebrovascular disease in multi-channel magnetic resonance (MR) images. The method combines intensity based fuzzy c-means (FCM) segmentation with spatial probability maps calculated from a normative set of images from healthy individuals. Unlike common FCM-based methods which segment only healthy tissue, we have extended the fuzzy segmentation to include patient-specific spatial priors for both pathological conditions (lesions and infarcts). These priors are calculated by estimating the statistical voxel-wise variation of the healthy anatomy, and identifying abnormalities as deviations from normality. False detection is reduced by knowledge-based rules. Assessment on a population of 47 patients from different imaging sites illustrates the potential of the proposed method in segmenting both hyperintense lesions and necrotic infarcts.

Keywords: brain tissue segmentation, lesions, infarcts, fuzzy clustering, outlier detection, MRI

1 Introduction

Brain lesions might be associated with numerous diseases, such as multiple sclerosis (MS), traumatic brain injury, or several cardiovascular symptoms, such as arterial fibrillation, impaired cognition and others [1][2]. Many methods in the literature have been developed for the automatic segmentation of brain lesions, especially MS lesions, using MR imaging [3][4]. Supervised segmentation strategies use features extracted from manual delineations as a training set to build the segmentation model. Those strategies require substantial human effort and are prone to intra- and inter-observer variability. Semi-supervised strategies model the normal data and then detect lesions as outliers that are not well explained by the model [5]. These methods usually derive a prior classification from a digital brain atlas that contains information about the expected location of white matter (WM), gray matter (GM) and cerebrospinal fluid (CSF) [6]. The available digital atlases, such the SPM [7] or ICBM [8] tissue

probability maps, have very low resolution, thus are suitable only for the detection of large lesions.

Although there have been quite many attempts to segment white matter or MS lesions, relatively less attention has been given to the segmentation of stroke related infarct lesions in elderly individuals, and Alzheimer's disease or diabetic patients. Cortical infarcts are characterized by a necrotic part with intensity similar to CSF imposing ambiguity in segmentation. Shen et al. [9] applied a modified Fuzzy c-means (FCM) clustering algorithm to classify voxels in an image according to intensity and prior probability of healthy tissue. They subsequently segment abnormal areas by using similarity to measure residual difference between the calculated tissue probabilities maps and the original prior probabilities. In such an approach lesions will be originally labeled as CSF or GM and will mislead the computation of class parameters.

Similarly to [3][9] we propose to segment infarcts within an outlier detection framework where the inconsistency between image intensity and spatial location is modeled. The proposed framework however differs in that it incorporates individualized spatial priors for both healthy and pathological tissue. The spatial priors for the healthy tissues are calculated by exploiting statistics from a healthy population group, whereas the spatial priors for the abnormal tissues are constructed in an outlier detection scheme [4]. For this purpose, all images from the healthy population group and testing subjects are firstly registered in a common space. The method furthermore classifies the abnormal tissue into gliotic (hyperintense) and necrotic (hypointense) regions based on multi-channel MRI (T1w and FLAIR). The five tissue types are obtained from FCM clusters by maximum membership segmentation followed by the application of some knowledge-based rules which reduce false detections.

Although the focus of the paper is to detect stroke-related infarcts and separate the hyperintense part of the infarct lesion (IL) from the hypointense infarct necrosis (IN), the method is also applied for the segmentation of other pathologies with high FLAIR signal appearance, such as white matter lesions, and these are marked also as IL. The method has been assessed on 47 brain MR images with lesions and compared against expert manual segmentations.

2 Methods

Brain abnormalities manifest themselves due to either unusual intensity characteristics (evident in appropriate image modalities) or due to inconsistency between tissue intensity and spatial location. Under this scope the method applies a modified FCM algorithm which combines intensity based similarity with prior anatomical knowledge. Specifically, image segmentation is performed in two steps. First the healthy and pathological regions are extracted by applying a semi-supervised approach using healthy images as training data. Then the segmented regions are smoothed and used as patient-specific prior tissue probability maps in a modified FCM algorithm. The first step is described in section 2.1, whereas the modified FCM algorithm is presented in section 2.2.

2.1 Construction of patient-specific spatial priors

The methodology for calculating prior tissue probability at each voxel is illustrated in Fig. 1 and can be separated into two branches, the construction of spatial priors for normal classes and abnormal classes.

The spatial priors for abnormal classes are constructed following an outlier detection approach. Assuming that voxel intensities are statistically independent and follow a multivariate normal distribution we build a statistical atlas from a training set of healthy subjects. Let n be the number of parameters used (MR sequences). For each voxel the statistical model consists of a $1 \times n$ vector with the mean gray level intensities and a $n \times n$ covariance matrix. Abnormalities are then detected as deviations from this model using as distance metric the voxel-wise Mahalanobis distance. This distance map, expressing probability of each voxel to be outlier, is then thresholded by a spatially varying threshold. The threshold is determined for each voxel by the ability of the statistical model to explain the training data. Specifically the average voxel-wise Mahalanobis distance is calculated for the healthy population group. Let's note it $D_{healthy}$. The applied voxel-wise threshold is $a \cdot D_{healthy}$, where a is a scalar (e.g. $a = 3$). This methodology extracts all abnormal regions including ILs, and INs. The abnormal regions are subsequently classified into ILs and INs based on intensity characteristics. Voxels with FLAIR intensities higher than the average intensity are classified as ILs, whereas hypointense voxels are classified as INs.

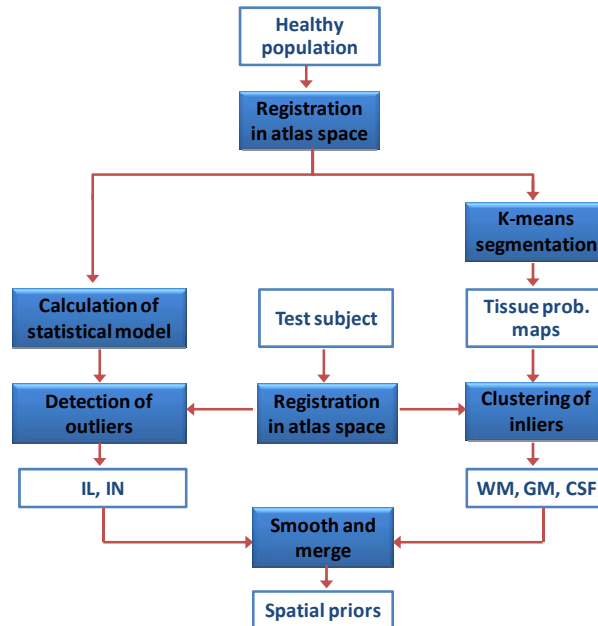


Fig. 1. Methodology of the calculation of spatial priors for normal tissue (white matter (WM), gray matter (GM) and cerebrospinal fluid (CSF)), and abnormal tissue (hyperintense infarct lesion (IL) and infarct necrosis (IN))

The healthy tissue priors are calculated by segmenting the patient's T1w image into WM, GM and CSF using the k-means algorithm. Since abnormalities are present, the algorithm is not randomly initialized because abnormal tissue might influence the cluster center estimation of healthy tissue. The estimation is thus guided by the healthy training data. Specifically, the healthy images are segmented by k-means, one by one, updating each time the estimate of the clusters centers location. The average (over all subjects) center location for WM, GM and CSF clusters is finally used as initial estimate in the segmentation process of the patient's image. Therefore, the effect of the presence of abnormalities is reduced.

All crisp tissue segmentations are converted to prior probabilities by gaussian smoothing. If Ω is the set of voxel locations in the image domain and C the number of classes, e.g. $C = 5$, all probability maps are collected together and normalized to satisfy $\sum_{i=1}^C p_{ij} = 1$, for each voxel $j \in \Omega$. If a class is missing and not detected in the outlier detection step described above, e.g. necrosis might not be present in every image, C is correspondingly reduced. These maps are utilized as individualized spatial priors in equation (1) as shown in the next section.

2.2 Modified FCM for healthy and abnormal tissue segmentation

The FCM algorithm [10] is based on minimizing an objective function with respect to a fuzzy membership U and set of cluster centroids V . Let $\mathbf{x}_j \in R^n$, be the feature representation (e.g. multi-channel intensities with $n = 2$) of voxel $j \in \Omega$. FCM clusters the data by computing the fuzzy membership $u_{ij} \subseteq U$ at each voxel j to the i -th class, such that $u_{ij} \in [0,1]$ and $\sum_{i=1}^C u_{ij} = 1, \forall j \in \Omega$. The number of classes C is assumed to be known. In the conventional approach the objective function expresses the degree of intensity based dissimilarity between the data values \mathbf{x}_j and the cluster centroids, $\mathbf{v}_i \subseteq V$. In this modified approach the objective function J_m includes an additional term reflecting the distance between the fuzzy membership u_{ij} and prior tissue probability, p_{ij} at each voxel j , which is calculated as described in the previous section. Thus J_m is expressed as follows:

$$J_m(U, V) = \sum_{j \in \Omega} \sum_{i=1}^C [u_{ij}^m d_v^2(\mathbf{x}_j, \mathbf{v}_i) + w_i (u_{ij} - p_{ij})^2] \quad (1)$$

where d_v is a distance function, $m \in (1, \infty)$ is a weighting exponent on each fuzzy membership controlling the degree of fuzziness and $w_i, i = 1 \dots C$, is a normalization weight controlling the contribution of each distance term per cluster. The use of different weights w_i allows controlling the penalty on dissimilarity for each class. For unbiased results w_i can take the default value of 1, as performed in this study. For the two terms in the objective function to be comparable (intensity dissimilarity and spatial inconsistency) the intensity values \mathbf{x} should be scaled in the range $[0,1]$, similarly to the probability range.

The constrained optimization of J_m is expressed using Lagrange multipliers and u_{ij} and \mathbf{v}_i are determined by setting the derivative of J_m to zero. For $m=2$ (as usually chosen in image segmentation), the fuzzy membership and the cluster center are calculated as follows:

$$\mathbf{v}_i = \frac{\sum_j u_{ij}^2 \mathbf{x}_j}{\sum_j u_{ij}^2}, \quad (2)$$

$$u_{ij} = w_i p_{ij} q_{ij} - \frac{\sum_i w_i p_{ij} q_{ij} - 1}{\sum_i q_{ij}} q_{ij}, \quad (3)$$

where $q_{ij} = \frac{1}{d_v^2(\mathbf{x}_j, \mathbf{v}_i) + w_i}$. The algorithm iteratively optimizes the objective function in Eq. (1) with the continuous update of u_{ij} and \mathbf{v}_i , until convergence is reached. The prior tissue probability maps are utilized as the membership at first iteration.

In regular FCM, d_v is the Euclidean distance. In order not to restrict the detection to data classes with hyper-spherical shape, we used the Mahalanobis distance expressed as follows:

$$d_v^2(\mathbf{x}_j, \mathbf{v}_i) = (\mathbf{x}_j - \mathbf{v}_i)' \cdot \Sigma^{-1} \cdot (\mathbf{x}_j - \mathbf{v}_i) \quad (4)$$

where Σ^{-1} is the inverse covariance matrix of the total dataset in this application. An improved implementation (to be investigated in the future) would include a different covariance matrix for each class and thus require iterative updates of Σ_i^{-1} based on the class membership at each iteration [11].

2.3 Knowledge-based rules

Knowledge-based rules are finally applied as post-processing steps to automatically reduce false positives. Abnormal voxels that are well explained by the statistical model of healthy anatomy, i.e. exhibiting small Mahalanobis distance, are assumed to be false detection and are eliminated. Single ischemic voxels are also removed. Two rules are applied for necrosis. Small necrotic clusters that lie across the GM-CSF interface are expected to be GM or CSF wrongly detected as IN due to partial volume effect. Necrosis is also constraint to be in proximity with a gliotic hyperintense region (IL). In contrast to other lesion segmentation methods [6] we didn't restrict the ILs to be situated inside WM in order not to further reduce generality.

2.4 Data description and preprocessing

The data consist of transaxial T1w and FLAIR scans obtained from different imaging sites. The acquisition of FLAIR (template space) was performed with a 3-mm slice thickness, no slice gap, voxel size=0.94×0.94×3 mm³, and in plane matrix 256×256. The images which didn't show any visible pathologies (72 in total) were used as the training data for statistical modeling of the healthy tissue. Different sets of data were selected for testing the method including lesions with or without necrosis. The first set (dataset1) consists of 35 diabetic patients with ILs (in 33 subjects) and necrotic infarcts (in 2 subjects). The second set (dataset2) involves 12 subjects selected from an ageing population (all subjects have both ILs and INs).

The abnormality masks used for assessing the method were manually delineated by a single rater for dataset1 and by two independent raters for dataset2. Hyperintensities in FLAIR images with respect to surrounding healthy white matter tissue were marked as IL. Cortical infarcts were separated into two regions. Their necrotic part

was segmented as IN, while the hyperintense rim, which usually surrounds the necrotic part, was marked as IL.

All images (from the testing subject and the normal subjects used for building the statistical atlas) were preprocessed in the same way by applying skull stripping, bias correction, global histogram matching and registration. For each subject the T1w image was first registered to the FLAIR image by affine registration. Then, for comparison of brains across different subjects, all subjects were spatially normalized to one of the subjects chosen as a template, by registering their T1w images with the template's T1w image using the HAMMER registration algorithm [12]. HAMMER algorithm achieves registration by hierarchically finding correspondences on edge voxels with high (feature-based) similarity. Voxel pairs with low similarity, such as the ones around pathology, are not included in the deformation process. Therefore, although the algorithm is designed for images without pathology, it has been observed that it is also quite robust when the images include small localized abnormalities. The FLAIR images and the expert-defined masks were subsequently warped to the template space using the deformation fields estimated for their corresponding T1 images.

2.5 Assessment

The results of the automated method were compared with lesion and infarct delineations by human experts. The segmentation of IL was assessed using the Dice coefficient as measure of spatial overlap. The Dice coefficient over two segmentations, S_1 and S_2 , is the volume of intersection of the two segmentations divided by the mean of the two segmentation volumes: $Dice = 2 \cdot |S_1 \cap S_2| / (|S_1| + |S_2|)$.

The necrotic part of infarcts is usually too small to be appropriately assessed by criteria measuring volume overlap, such as the Dice coefficient. Thus necrotic voxels were grouped into clusters based on connectivity and segmentation assessment was performed based on the number of correctly classified clusters. A necrotic cluster is considered as a correct detection if the voxels within the cluster partially overlap with manually segmented necrosis. Clusters with a volume smaller than a threshold V_{min} were not considered in the evaluation process.

3 Results

3.1 Segmentation of lesions and infarcts

Fig. 2 (last row) illustrates some examples of IL and IN segmentation by the proposed method. The results are compared against manual delineations (IL available for subjects 3 and 4, and IN available for subjects 1, 2, 3). In case of large lesions overall good agreement is observed. In order to quantitatively assess lesion segmentation, Dice coefficients have been calculated for both datasets and plotted against the total lesion load (or average volume in the case of two raters) for all subjects in Fig. 3. The inter-rater variability is also calculated for dataset2 by defining the one of the two masks as ground truth and assessing the other. It can be observed that the agreement

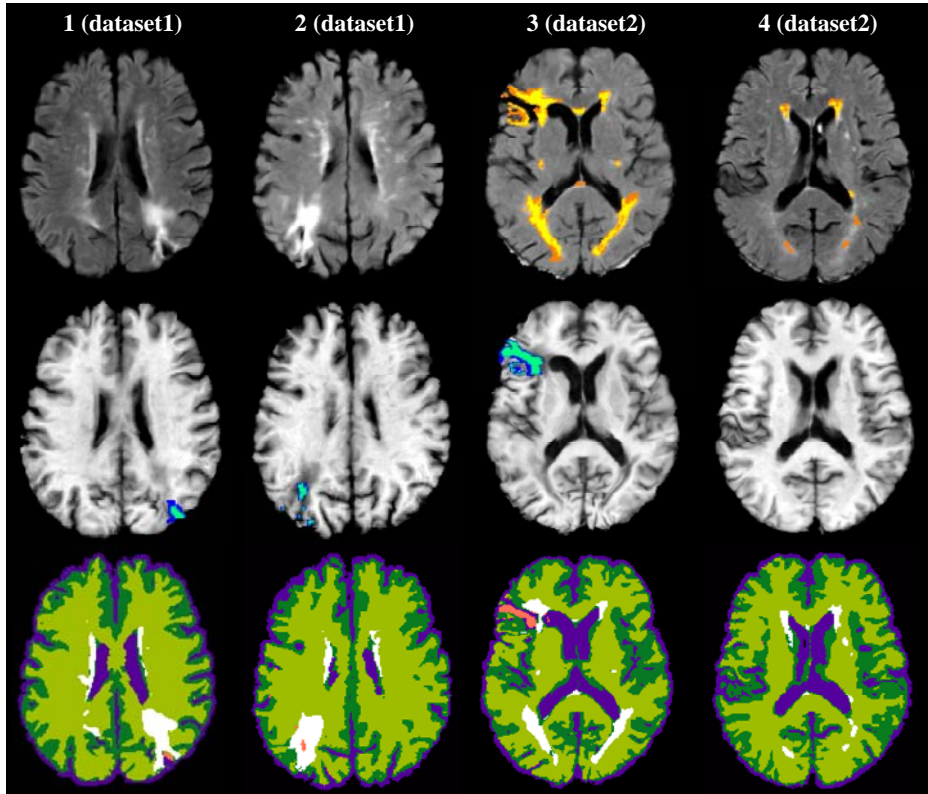


Fig. 2. Visual segmentation assessment on 4 subjects. From top to bottom: FLAIR with manually outlined IL (available only for subjects 3 and 4), T1w with manually outlined IN, automated segmentation (blue: CSF, dark green: GM, light green: WM, pink: IN, white: IL). The manual segmentations are shown in color scale with darker color indicating the presence of IL or IN in neighboring slices.

between automatic and manual segmentation increases with increasing lesion load. This is expected since small lesion clusters might be missed by the method, due to partial volume effects, smoothing and/or warping of the images and manual masks. We should note that any further volumetric measurements for clinical assessment should be performed in the original space by warping back the images with segmented lesions.

The segmentation of infarct necrosis is assessed based on the number of correctly or incorrectly detected clusters, defined as true positives (TP) and false positives (FP), over all subjects. The effect of the cluster volume threshold, V_{min} , on segmentation accuracy, is shown in Fig. 4. On dataset1 (Fig. 4a) seven necrotic clusters were in total manually outlined for the two subjects. Two of them are missed by the automated method (for $V_{min}=0$). The volume of each of the missed clusters is equal to a single voxel; thus it is reasonable to miss these clusters. The assessment of detection rate is different for dataset2 where two manual masks (R1 and R2) are available. First we

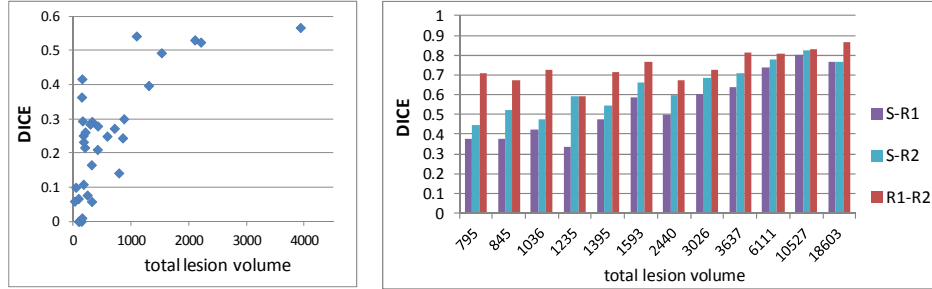
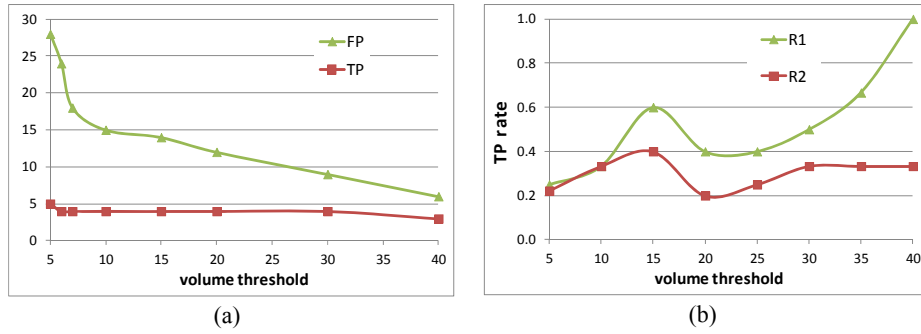


Fig. 3. Dice scores in relation to total IL volume (in voxels) for each subject in dataset1 (left) and dataset2 (right). Comparison of our method (S) with manual segmentations by raters R1 and R2, and inter-rater variability (R1 vs R2) are shown for dataset2. In this case the total volume represents the average volume.



	<i>dataset1</i> (2 subjects)	<i>dataset2</i> (13 subjects)	
		R1	R2
# clusters	7	22	22
median vol. (# voxels)	209	3	5

(c)

Fig. 4. Detection accuracy of IN versus cluster volume threshold. (a) TPs and FPs for dataset1 where the total number of IN clusters is 7. (b) TP rate for dataset2 by defining manual segmentation R1 and R2 as ground truth (c) Statistics of clusters (number and median volume)

assumed R1 is ground truth and calculated the sensitivity (TP rate) of the automated method as the ratio of correctly detected clusters by the method versus the number of clusters correctly detected by the 2nd rater (after volume thresholding). Then we assumed R2 is ground truth and compared our method against the accuracy of the 1st rater. The sensitivity is plotted in Fig. 4b. The results show that the method performs more similar to rater 2 in necrosis segmentation, and more similar to rater 1 in IL segmentation (Fig. 3, right). The total number of clusters (summed over all subjects) and their median volume (in voxels) are shown in Fig. 4c. The statistics show that the

delineation of the necrotic part of the infarcts in dataset2 included many unconnected components making automated detection more difficult.

4 Discussion and Conclusions

We presented a method that uses multi-channel MR images to segment healthy and abnormal brain tissue employing a statistical model, together with a modified FCM algorithm. The segmentation of enhancing lesions approached the accuracy of manual segmentation for high lesion load. The detection of infarcts was only slightly worse than the performance of human experts. However more validation is required because there was a considerable disagreement in spatial correspondence between expert segmentations. The assessment of the method on synthetic databases might provide useful insight. Also the sensitivity of the method to the selected parameters should be investigated.

The method was applied to segment and classify vascular pathology; it can nevertheless also be applied to detect lesions that have not vascular origin. Adaptation of the knowledge-based rules might be required in this case.

Extensions of the FCM algorithm have been proposed in the literature to reduce the algorithm's sensitivity to noise and intensity inhomogeneity. The extensions incorporate spatial regularisation, topological constraints or bias field correction, as briefly reviewed in [13]. In this paper we did not apply any of those extensions since we focused on the lesion and infarct segmentation. It is expected however that the improved modeling of the healthy tissue will improve also the segmentation of the abnormal tissue.

Acknowledgments

This research was supported by a Marie Curie International Reintegration Grant within the 7th European Community Framework Programme.

References

1. de Groot, J.C., et al.: Cerebral white matter lesions and cognitive function: the Rotterdam Scan Study. *Annals of Neurology* 47(2), 145-151 (2000)
2. Rovaris, M., et al.: Relevance of Hypointense Lesions on Fast Fluid-Attenuated Inversion Recovery MR Images as a Marker of Disease Severity in Cases of Multiple Sclerosis. *Am. J. Neuroradiology* 20(5), 813 - 820 (1999)
3. Prastawa, M., Gerig, G.: Automatic MS Lesion Segmentation by Outlier Detection and Information Theoretic Region Partitioning. In: *Proceedings of the Grand Challenge II Workshop of MICCAI 2008*, vol. 11(W5) (2008)
4. Lladó, X., et al.: Segmentation of multiple sclerosis lesions in brain MRI: A review of automated approaches. *Information Sciences* 186, 164–185 (2012)
5. Zacharaki, E.I., Bezerianos, A.: Abnormality segmentation in brain images via distributed estimation. *IEEE Trans Inf Technol Biomed*, in press (2012)

6. Van Leemput, K., Maes, F., Vandermeulen, D., Colchester, A., Suetens, P.: Automated segmentation of multiple sclerosis lesions by model outlier detection. *IEEE Trans Med Imag* 20, 677–88 (2001)
7. Statistical Parametric Mapping, <http://www.fil.ion.ucl.ac.uk/spm/>
8. International Consortium for Brain Mapping, <http://www.loni.ucla.edu/ICBM/>
9. Shen, S., Szameitat, A.J., Sterr, A.: An improved lesion detection approach based on similarity measurement between fuzzy intensity segmentation and spatial probability maps, *Magnetic Resonance Imaging* 28, 245–254 (2010)
10. Bezdek, J.C.: *Pattern recognition with fuzzy object function algorithms*. Norwell (MA): Kluwer Academic Publishers, pp. 256 (1981)
11. Liu, H-C., Jeng, B-C., Yih, J-M., Yu, Y-K.: Fuzzy C-Means Algorithm Based on Standard Mahalanobis Distances, In: *Proceedings of the 2009 International Symposium on Information Processing*, pp. 422-427 (2009)
12. Shen, D., Davatzikos, C.: HAMMER: hierarchical attribute matching mechanism for elastic registration. *IEEE Trans Med Imag* 21, 1421-39 (2002)
13. Ji, Z.-X., Sun, Q.-S., Xia, D.-S.: A modified possibilistic fuzzy c-means clustering algorithm for bias field estimation and segmentation of brain MR image, *Computerized Medical Imaging and Graphics* 35, 383–397 (2011)