Fully automatic segmentation of multiple sclerosis lesions in brain MR FLAIR images using adaptive mixtures method and markov random field model

Rasoul Khayati\textsuperscript{a}, Mansur Vafadust\textsuperscript{a}, Farzad Towhidkhah\textsuperscript{a,∗}, S. Massood Nabavi\textsuperscript{b}

\textsuperscript{a}Biomedical Engineering Faculty, Amirkabir University of Technology, Tehran, Iran
\textsuperscript{b}Neurology Department, Medical Faculty, Shahed University, Tehran, Iran

Received 4 January 2007; accepted 19 December 2007

Abstract

In this paper, an approach is proposed for fully automatic segmentation of MS lesions in fluid attenuated inversion recovery (FLAIR) Magnetic Resonance (MR) images. The proposed approach, based on a Bayesian classifier, utilizes the adaptive mixtures method (AMM) and Markov random field (MRF) model to obtain and upgrade the class conditional probability density function (CCPDF) and the \textit{a priori} probability of each class. To compare the performance of the proposed approach with those of previous approaches including manual segmentation, the similarity criteria of different slices related to 20 MS patients were calculated. Also, volumetric comparison of lesions volume between the fully automated segmentation and the gold standard was performed using correlation coefficient (CC). The results showed a better performance for the proposed approach, compared to those of previous works.

© 2007 Elsevier Ltd. All rights reserved.

Keywords: Multiple sclerosis; Segmentation; Adaptive mixtures method

1. Introduction

Multiple sclerosis (MS) is one of the progressive central nervous system diseases, which causes morphological and structural changes to the brain. Quantitative assessment of the changes in the brain magnetic resonance (MR) images in association with clinical judgment can provide more accurate assessment of the disease progress. Also, it can provide important information in order to find out the most effective therapeutic regimes for patients. Due to high resolution and good differentiation between soft tissues and other structures, magnetic resonance imaging (MRI) has superiority to the other imaging techniques for the studies of nervous system diseases. MRI has been known as the best paraclinical examination for MS which can reveal abnormalities in 95% of the patients [1]. The traditional interpretation of MR images by a specialist is a difficult and time-consuming task, and result directly depends on the experience of the specialist. A reason for such a difficulty is related to the complexity and visually vague edges of anatomical borders. Therefore, it is desirable to have an automatic segmentation method to provide an acceptable performance.

Segmentation of brain tissues refers to the task of classification of brain into relevant classes such as normal tissue (white matter and gray matter), abnormal tissue (MS lesions) and cerebrospinal fluid (CSF). As the result of this segmentation procedure, the quantitative measure of MS area can be calculated and used for follow up of the disease.

There are many proposed approaches, automatic and semi-automatic, for segmentation of brain into different tissues, including MS lesions. These approaches include a variety of methods such as statistical, fuzzy, neural networks, and fuzzy neural networks, some of which will be reviewed in the following.

Kamber et al. [2] developed a brain tissue model for segmentation of MR images of patients with MS disease. The offered model was three-dimensional, voxel based, which provided...
a priori probabilities for white matter, gray matter and CSF. They suggested restricting search area for MS lesions to only white matter. Johnston et al. [3] presented a semi-automatic segmentation approach based on stochastic relaxation method (SRM) and iterated conditional mode (ICM) [4]. They implemented their algorithm to detect MS lesions in the white matter and also performed a partial volume analysis for the resulted segmentation. Their quantitative analysis was performed by calculating similarity index (SI). They reached to 0.65 for the mean value of SI for selected images. Udupa et al., utilizing fuzzy objects and fuzzy connected sets [5], suggested a semi-automatic segmentation method [6]. In this method, first, sample points, for white matter, gray matter, and CSF was selected by user, which later on was extended and named as fuzzy connected sets. In the next step, the void pixels (i.e., the objects other than the three fuzzy objects defined by the user) were offered to the user, for probable selection as MS lesions. They showed that their methodology was highly reliable and consistent, with a variation coefficient of 0.9% for volume. Admasu et al. [7] suggested an improvement to the Udupa approach by using a neural networks, instead of user, to make decision about the void pixels. The quantitative evaluation was carried out by comparing the quantity of lesions through the proposed method and the manual segmentation by an expert, as well as comparing false positive and false negative percentages. The results showed 90% correlation with the segmentation by the specialist physician. Pachai et al. [8] proposed an automatic segmentation algorithm based on a multi-resolution approach using pyramidal data structure. They showed that systemic pyramidal decomposition in the frequency domain provides a low level, robust, and flexible tool for the analysis of MR images. The total time required for processing a complete MR image, in this method, was much less than those of the semi-automatic and manual approaches. Blonda et al. suggested utilizing fuzzy neural networks for segmentation of MS lesions [9]. In this research, two models, fuzzy learning vector quantization (FLVQ) and fully self-organized simplified resonance theory (FOSART), were used for a semi-automatic segmentation of data set. The researchers showed that FOSART has less performance in comparison with self-organized map (SOM) and FLVQ, in minimizing mean squared error (MSE). However, FORSAT minimizes MSE with fewer epoch numbers and also it is more stable with respect to small changes in input parameters and in the order of the presentation sequence. Guttmann et al. [10] as well as Kikinis et al. [11], carried out computer analysis for the study and quantification of MS lesions by utilizing an adaptive statistical segmentation algorithm, namely expectation maximization (EM) [12], for a semi-automatic segmentation of brain tissues. The burden of disease was evaluated via computation of the volume of individual lesions. Boudraa et al. [13] performed an automatic detection of MS lesions by using fuzzy C-means (FCM) algorithm. The result of this segmentation was compared with manual segmentation via similarity criteria which reached to a mean value of 0.62 for SI. Leemput et al. [14] proposed the use of an intensity-based and stochastic model-aided segmentation approach for fully automatic segmentation of MS lesions. The voxels which were not well explained by the models were considered outliers, and were labeled as MS lesions. Also, they used SI for measuring correspondence with a manual segmentation, and reached to a maximum equals to 0.51. In their research, Nett et al. [15] focused on Bayesian classifier. They improved the statistical classifier by non-parametric modeling of class conditional probabilities using parzen window and by a priori probabilities modeling of the class probabilities through a Markov random field (MRF) model. In addition, quantification was carried out by calculating the total volumes of normal and diseased (lesions) tissues. Ardizzone et al. [16] used a two-channels FCM clustering technique to detect the lesions. The result of their method was compared with proton density (PD) FCM, T2-weighted (T2-w) FCM, and 2D FCM by a physician, through utilizing sensitivity and specificity criteria. The results obtained in this research showed a better efficiency for two-channel FCM, compared with the other methods.

Anbeek et al. [17] proposed a novel automatic approach for segmentation of white matter lesions in MR images of brain. The introduced algorithm uses different information, including voxel intensity and the spatial information, to classify voxels by a K-nearest neighbor (KNN) classifier. This technique assigns a probability to each voxel for being part of white matter lesion. The SI, then, is used for determination of an optimal threshold on the probability map, to segment the images. They showed the high accuracy of their approach, in comparison with the other methods for similar task. Recently, Admiraal-Behloul et al. [18] suggested a fully automatic segmentation method for quantifying white matter hyperintensity in a large clinical trial on elderly patients. Their introduced algorithm combined information from three different MR images including PD, T2-w and fluid attenuated inversion recovery (FLAIR) and FCM algorithm for clustering process. The approach demonstrated very high volumetric and spatial agreement with expert delineation. A mean value equals to 0.75, was obtained for SI. Also, they suggested that better accuracy was associated with bigger FLAIR (T2) lesion loads.

In previous algorithms suggested by the authors of this paper, an automatic approach was introduced for MS segmentation of brain, in MR-FLAIR images [19]. The proposed approach was based on a modified Bayesian classifier. The class conditional probability density function (CCPDF) of each class, required for classification, was obtained by applying adaptive mixtures method (AMM), on training samples. To improve the performance of the classifier, the selection of the training samples for each tissue class was performed automatically by using a non-linear anisotropic diffusion filter and a k-means classifier. This filter was used for enhancement and smoothing, while preserving the edges. Using a MRF model, to attain and update a priori probability for each class, was another modification applied to the classifier.

Despite the good performance of the proposed method in our previous work [19], it needed to the learning samples which in turn can be considered as a drawback. To overcome this problem, in this paper we estimate the CCPDFs and the a priori probabilities, while no initially training sample or initially guess is required. In the proposed method, distribution function
of the FLAIR image is estimated by a mixture of a large number of normal terms through AMM. Then, the mixture terms are categorized into three classes, as the CCPDFs and the a priori probabilities of the classes. In the next steps, a priori probabilities of the classes as well as mean and variance of CCPDF for each class are attained and updated utilizing MRF model and AMM, respectively.

In the next sections, details of the research procedure including MR imaging type, manual segmentation of MS lesions, brain segmentation, Bayesian classification, utilization of the AMM to estimate the probability density functions, and MRF model for obtaining a priori probabilities of the classes, are explained. Then, the proposed approach is introduced and its efficiency is compared with some of the previous approaches.

2. Methods

2.1. Patients and MR imaging

Twenty definite MS patients, including 16 female and four male with average age of 29±8 years old, were selected in this study according to the revised McDonald criteria 2005 [20]. Mean disease duration for the patients was five years. For all patients the same MR images were obtained via a Siemens 1.5 T scanner. All images were acquired according to full field MRI criteria of MS [20] in T2-w, T1-weighted (T1-w), Gadolinium-enhanced T1-weighted, and FLAIR in axial, sagittal and coronal surfaces.

We selected the FLAIR images, especially axial ones, with lesions in deep, priventricular, subcortical, and cortical white matters (supratentorial lesions). More lesion load and higher accuracy of FLAIR in revealing of these MS lesions were the reason for this selection [21]. Also, FLAIR is especially helpful for priventricular lesions closely apposed to an ependymal surface, where they may be obscured by the high signal CSF on T2-w images [21].

Scan parameters of, repetition time (TR)/echo time (TE)/inversion time (TI), for FLAIR images were 9000/144/2500 ms, TR/TE for T1-w images were 424/10 ms and TR/TE for T2-w images were 3820/105 ms. Each image volume (patient data) consisted of averagely 40 slices with a 256 × 256 scan matrix.

As a preprocessing step, the images first had to be scaled. Therefore the range between zero intensity and maximum intensity, \( M \), in the original 12-b data (\( I \)) was scaled to a new intensity (\( I_s \)) between 0 and 255 (8-bit) which is obtained by: 
\[
I_s = \frac{I}{M} \times 255.
\]
This preprocessing enables us to make use of all dynamic range of 256 levels, and also immune the images from the variation in their intensity. The pixel size was 1 mm\(^2\), and the slice thickness was 3 mm with any gap.

For a better evaluation of the proposed method, we investigated supratentorial FLAIR slices, 20 slices per patient in average. The selection of slices was based on presence of lesions, with clear boarders, which enable careful outlining in manual segmentation by our colleagues, a neurologist and a radiologist. At least five slices from each of four patients (at least 20 slices), and 12 slices (in average) from each of the rest, 16 patients, were selected. Thus, totally 217 slices were used.

2.2. Manual segmentation of MS lesions and brain segmentation

In this study, the neurologist and the radiologist who were not aware of the results of the computerized methods in this research, were asked to perform manual segmentation of MS lesions in FLAIR images with visual inspection of the corresponding T1-w and T2-w images. The results of this step for all the selected slices, provided us binary segmented images, considered as gold standard, for usage in evaluation step [17]. To compare the segmentation of lesions for patients with different lesion volumes (LV), three patient categories: patients with, small (LV < 4 cc), moderate (4 cc < LV < 18 cc), and large lesion (LV > 18 cc) load, in our selected slices, were composed [18].

Also, the brain segmentation was performed using a fully automatic object-oriented approach [22]. This method was based on the regional–spatial characteristics of brain in MR images. This algorithm consists of five steps. At first, original image is converted to a binary image. Secondly, morphological opening on the binary image is performed and tiny regions are eliminated. In the third step, three rectangular masks showing the cerebral regions are produced and the regions in the binary image which have overlap with these rectangles are preserved and, the rest are eliminated. In the fourth step, final mask is generated by dilation of selected regions and filling tiny holes. Finally, an image, which includes only cerebral tissues, is obtained by applying the resulted mask on the original image.

2.3. Bayesian classification

The statistical classification is fundamentally based upon the application of Bayes rule for the quantification of decision process. The Bayes rule, Eq. (1), indicates how the posterior probability of a class is calculated, given feature measurements [23].

\[
p(o_j|x) = \frac{p(x|o_j)p(o_j)}{p(x)}.
\]  

In this equation, \( o_j \) refers to class \( j \)th and \( x \) is the feature vector. \( p(x|o_j) \), \( p(o_j) \), and \( p(x) \) are the CCPDF, the a priori probability and the a posteriori probability of class \( o_j \), respectively. The feature vector \( x \) is formed by the intensity of the FLAIR image pixels. The quantity \( p(x) \) is known as the evidence, and serves only as a scale factor, which makes Eq. (1) a true probability, with values between zero and one. It can normally be ignored when comparing the result of Eq. (1) for all classes, as it is constant for all classes. Therefore, in this case Eq. (1) can be simplified as

\[
p(o_j|x) = p(x|o_j)p(o_j).
\]  

This equation is maximum a posteriori (MAP) estimate of Eq. (1). Assuming the number of classes equals to \( n_c \), for a
given \( x \), the MAP is computed for all classes and \( x \) is assigned to the class with maximum MAP [23]:

Decision \( \omega_j \) if \( p(\omega_j|x) > p(\omega_i|x) \)

for \( i = 1, \ldots, n_c, \ i \neq j \)  (3)

It is reminded that the Bayesian decision rule is the optimal rule as long as, the probability models for \( p(x|\omega_j) \) and \( p(\omega_j) \) are available. Because, these models are not known for brain tissues, they are required to be estimated before any further computation. The efficiency of classification directly depends on the quality of the obtained models. The estimation procedure is described in the next sections.

2.4. Estimation of probability density function by AMM

To estimate \( p(x|\omega_j) \) the AMM [24] is used. The main idea of the AMM is utilizing all the data, one by one, and determining the distance of each observation to each term density in the model. If the distance to each term is more than a threshold, a new term is created. However, if the distance is less than the threshold for all terms, the estimated parameters will be updated

The rule in Eq. (6) states that, we create a new term if the smallest squared MD is more than the threshold. We chose \( \xi_c = 1 \), as it is recommended for univariate case [24]. Therefore, a new term is created if a new observation is more than one standard deviation away from the mean of each term.

4. If the minimum distance of the new point is less than the threshold \( \xi_c \), then update the existing terms using

\[
\hat{p}_i^{(n+1)} = \frac{n \hat{p}_i^{(n)}}{n+1}, \quad i = 1, \ldots, N. 
\]

The rule in Eq. (6) states that, we create a new term if a new observation is more than one standard deviation away from the mean of each term.

The required steps for AMM are as follows:

1. Initialize the AMM using the first point of data \( X^{(1)} : \hat{\mu}_1^{(1)} = X^{(1)}, \hat{\Sigma}_1^{(1)} = 1 \), in which \( I \) is the identity matrix. In the case of univariate, the variance of the first term is one.

2. For a new point of data \( X^{(n+1)} \), calculate the squared Mahalanobis distances (MD) from each of existing terms using

\[
\text{MD}_i^2(X^{(n+1)}) = (X^{(n+1)} - \hat{\mu}_i^{(n)})^T (\hat{\Sigma}_i^{(n)})^{-1} (X^{(n+1)} - \hat{\mu}_i^{(n)}),
\]

3. If the minimum squared distance of the new point from all the terms is more than a threshold \( \xi_c \), i.e.,

\[
\min_i \{ \text{MD}_i^2(X^{(n+1)}) \} > \xi_c
\]

4. If the minimum squared distance of the new point from all the terms is more than a threshold \( \xi_c \), then update the existing terms using

\[
\hat{p}_i^{(n+1)} = \frac{n \hat{p}_i^{(n)}}{n+1}, \quad i = 1, \ldots, N. 
\]

5. Continue steps 2–4 for all data points (voxels).

The following features can be mentioned as the advantages of the AMM algorithm in comparison with the EM algorithm:

1. The AMM does not require an initial knowledge for the number of terms.
2. It does not require an initial guess for the term parameters.
3. As the whole data are not used simultaneously for updating the estimation of term parameters, the required calculation is much lower than that of EM especially while we have a massive data set.
4. In contrast to the EM algorithm, the convergence of AMM and also number of iterations do not depend on the tolerance (i.e., a preset threshold for algorithm termination), initial parameters, and the data load.

2.5. Computation of the a priori probability by MRF model

An advantage of MRF models [25] is the use of neighborhood information to improve the a priori probabilities \( p(o) \) [26]. MRF models an image or a volume as a random field, a structured collection of random variables, \( \Omega = \{ \Omega_1, \ldots, \Omega_m \} \), and defined on the set \( S \). There are two conditions for considering a random field as MRF [25]:

1. \( p(o) > 0, \forall o \in \Omega \), Condition of positivity,
2. \( p(o_i|o_{S-\{i\}}) = p(o_i|o_{N_i}) \), Condition of Markovianity.

The condition of positivity is implicitly satisfied with the application of the MRF model of a priori probabilities. The condition of Markovianity states that the probability of an observation \( o_{j} \), given the other random variables in the field, is equal to the probability of the observation \( o_{i} \), given a neighborhood around the sampled location, denoted as \( N_i \). This statement fits well with respect to surrounding local areas of a pixel in order to help modeling of \( p(o_i) \). This condition alone is not amenable in classifier computations. It is possible to find a solution by use of Gibbs random field (GRF). The Gibbs distribution is in the form of the Eq. (15):

\[
p(o) = \frac{1}{Z} \exp \left( -\frac{1}{T} U(o) \right),
\]

\[
Z = \sum_{\Omega} \exp \left( -\frac{1}{T} U(o) \right),
\]

\[
U(o) = \sum_{c} V_c(o).
\]

The constant \( T \), in Eq. (15), stands for the temperature constant which normally supposes to be one. The quantity \( Z \) is a normalizing constant which guarantees that \( p(o) \) is always smaller or equal to one. The function \( U(o) \) in Eq. (17) is the energy function and is obtained by the sum of the function \( V_c(o) \), over all \( C \) possible cliques. Each \( V_c(o) \) is called clique potential. A clique is a grouping of samples in a neighborhood system, such that the grouping includes voxels that are neighbors of another in the same system. The Hammersley Clifford theorem defines that if and only if a random field \( \Omega \) on \( S \) is a MRF with respect to neighborhood system \( N \), then \( \Omega \) is a GRF on \( S \) with respect to a neighborhood system \( N \) [25]. This result allows the given conditional probability as a Markovianity condition of a MRF to be converted to the non-conditional probability of a Gibbs distribution, given in Eq. (15).

To estimate \( p(o) \) in our model we used the simple equation for the \( U(o) \) energy function proposed by Therrien [27], and utilized by Nett et al. [15]. This equation is a linear combination of products of elements in the cliques:

\[
U(o_{k,l}) = \omega_{k,l} \cdot (\alpha + \beta_1 \cdot (\omega_{k-1,l} + \omega_{k+1,l})
+ \beta_2 \cdot (\omega_{k,l-1} + \omega_{k,l+1})).
\]

In this equation, \( \omega_{k,l} \) shows the sample at indices \((k, l)\) on the labeled image. The parameters \( \alpha \) and \( \beta_1 \) and \( \beta_2 \) are the parameters used for adjustment of the relative weights or contributions of neighborhood interaction.

2.6. Fully automatic segmentation of MS lesions using AMM and MRF model

Based on the explanations mentioned above, the block diagram of our approach for fully automatic segmentation of MS lesions is shown in Fig. 1.

This approach includes nine steps:

1. Estimation of input image distribution as a mixture of a large number of normal terms by the AMM (i.e., \( \hat{\mu}_{\text{terms}} = \{ \hat{\mu}_1, \ldots, \hat{\mu}_N \} \), \( \hat{\mu}_{\text{terms}} = \{ \hat{\mu}_1, \ldots, \hat{\mu}_N \} \), and \( \hat{\sigma}^2_{\text{terms}} = \{ \hat{\sigma}^2_1, \ldots, \hat{\sigma}^2_N \} \). It is reminded that for univariate case (one feature), the covariance matrix is replaced by the variance of each term.
2. Selection of thresholds for terms grouping. The thresholds value is calculated according to \( \text{thr}_{1,2} = \mu_{\text{in}} \pm 1.5 \sigma_{\text{in}} \), in which \( \mu_{\text{in}}, \sigma_{\text{in}} \) are mean and standard deviation of the input image, respectively.

3. Grouping the normal terms of the estimated distribution with the selected thresholds into three classes as the CCPDFs of classes. Indeed, initial values for parameters (i.e., \( \hat{\mu}_{\text{init}}(\omega_j) \), \( \hat{\mu}_{\text{init}}(\omega_j) \) and \( \hat{\sigma}_{\text{init}}^2(\omega_j) \)) are estimated. The \( \hat{\mu}_{\text{init}}(\omega_j) \) is calculated according to

\[
\hat{\mu}_{\text{init}}(\omega_j) \quad = \begin{cases} 
\text{mean}\{\mu \in \hat{\mu}_{\text{term}} \mid \hat{\mu}_{\text{term}} < \text{thr}_1\} & \text{for } j = 1 \\
\text{mean}\{\mu \in \hat{\mu}_{\text{term}} \mid \text{thr}_1 \leq \hat{\mu}_{\text{term}} \leq \text{thr}_2\} & \text{for } j = 2 \\
\text{mean}\{\mu \in \hat{\mu}_{\text{term}} \mid \hat{\mu}_{\text{term}} < \text{thr}_2\} & \text{for } j = 3 
\end{cases},
\]

in which, \( j \) denotes \( j \)th class. Also, the variances \( (\hat{\sigma}_{\text{init}}^2(\omega_j)) \) and the coefficients \( (\hat{\mu}_{\text{init}}(\omega_j)) \) of these classes are equal to mean of the variances and sum of the coefficients of corresponding terms, respectively.

4. Bayesian classification of the input image using initial values of parameters.

5. Extraction of the gray level values for each tissue class resulted from previous step.

6. Estimation of the CC PDF (i.e., \( \hat{\mu}_k(\omega), \hat{\sigma}_k^2(\omega) \)) for each tissue class, through the AMM.

7. Estimation of a priori probabilities (i.e., \( \hat{p}_k(\omega) \)) for tissue classes, by MRF model stated in Eqs. (15)–(18).

8. Calculation and evaluation of the new value for termination tolerance. The new value of termination tolerance is calculated from

\[
tol = \max \left( \frac{\max(\text{abs}(\hat{\mu}_k(\omega) - \hat{\mu}_{k-1}(\omega))),}{\max(\text{abs}(\hat{\mu}_k(\omega) - \hat{\mu}_{k-1}(\omega))),} \right) \quad \text{(19)}
\]

in which \( \hat{\mu}_k(\omega), \hat{\mu}_k(\omega), \) and \( \hat{\sigma}_k^2(\omega) \) are the row vectors of the a priori probabilities, mean, and variance of the tissue classes in \( k \)th iteration, respectively.

9. Repetition of steps 4–8 or termination of the computation.

### 2.7. Postprocessing

In manual segmentation, the possible lesions with sizes as small as one or two pixels are not usually considered as MS lesions by experts. However, as the gray level of neighboring pixels in an image usually are highly correlated, in our approach, the one-pixel object is considered as a possible noise if its gray level is higher than the average gray levels of its neighbor pixels (in a \( 3 \times 3 \) neighborhood) plus a margin of 25 (which is selected by trial). This object is recognized as noise and removed in the next step. Regarding the two-pixel objects, in contrast with the manual procedure, these objects may be identified as seed points of MS lesions as recommended by the neurologist.

### 2.8. Evaluation

Results of the lesion segmentation based on the proposed method are compared with the gold standard. The similarity criteria (SI) [28], overlap fraction (OF) and extra fraction (EF), [29] are calculated for the selected slices. The SI is a criterion for the correctly classified lesion area relative to the total areas of lesion, in both the gold standard and in the segmented image. The OF and EF specify, respectively, the areas which have been correctly and falsely classified as lesion area relative to the lesion area in the gold standard. The similarity criteria are defined by (see Fig. 2):

\[
\text{SI} = \frac{2 \times \text{TP}}{2 \times \text{TP} + \text{FP} + \text{FN}}, \quad (20.1)
\]

\[
\text{OF} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad (20.2)
\]

\[
\text{EF} = \frac{\text{FP}}{\text{TP} + \text{FN}}. \quad (20.3)
\]

In these equations, TP stands for true positive voxels, FP for false positive voxels, and FN for false negative voxels. SI and OF for a good segmentation should be close to 1 and EF should be close to 0. Practically, a value for SI more than 0.7 represents a very good segmentation in this field [30]. Also, the mean values of similarity criteria are categorized to the three patient categories and volumetric comparison of lesions volume between the fully automated segmentation and the gold standard are performed using correlation coefficient (CC).

### 3. Results

The proposed algorithm was implemented on different FLAIR images using a Pentium four processor (2.66 GHz) and 512 MB RAM. In Fig. 3(a) and (b), a typical original FLAIR image, as a sample slice of a patient with moderate lesion load, and segmentation of brain, in it, from the other tissues have been shown, respectively.
In Fig. 4(a), results of steps 1 and 2 of the proposed algorithm, including estimated distribution obtained through the AMM with maximum 25 Gaussian terms overlaid on distribution of the sample slice and selection of the thresholds, have been shown. The calculated MSE, less than $6 \times 10^{-7}$, show an acceptable result. In Fig. 4(b), (c), and (d), results of step 3, including mixture of normal terms obtained by AMM, grouping them into three classes with the selected thresholds, as the distribution functions of the classes, and the resulted distribution by thresholding, overlaid on distribution of the sample slice with MSE less than $3 \times 10^{-5}$, have been shown, respectively.

To perform steps 4–8 of the proposed algorithm, the initial values for parameters were set as follows: the maximum iteration to 10 and the tolerance for stopping iteration to $10^{-5}$. The parameters of MRF model, i.e., $\alpha$, $\beta_1$, and $\beta_2$ have been experimentally set to 0.1, 0.01, and 0.01 for the best result.

Finally, the result of lesion segmentation in the sample slice (Fig. 3(a)) after six iterations and updating the parameters (the a priori probability, mean, and variance of each class), has been shown in Fig. 5(c). As it is seen, there is a good correlation between the input image and the resulted image indicating the acceptable performance of the suggested algorithm in detecting the lesion borders as well as CSF regions. Finally, the segmented lesions resulted from the manual segmentation and the postprocessing step have been shown in Fig. 5(b) and (d), respectively.

Fig. 6(a) shows the decision regions in Bayesian classification. It is reminded that the overlap (see Fig. 6(a)) between the three classes is a reasonable value (the mean of probability of errors in all data set is less than 0.05). It was found that the overlap for singular lesions with high contrast was less than that for scattered lesions with lower contrast. Fig. 6(b) shows the distribution of input image overlaid on its final estimation with MSE less than $1 \times 10^{-6}$.

In Figs. 7 and 8, results of applying the proposed algorithm to the image of a patient with a small lesion load and the image of a patient with a large lesion load have been shown, respectively.

The evaluation of the results performed qualitatively and quantitatively as follows. The quality performance of the results was confirmed by the neurologist and the radiologist separately. Then, in quantitative evaluation step, the similarity criteria (i.e., SI, OF, and EF) were calculated for all selected slices. Mean values of the similarity criteria as well as mean value of segmentation time ($T$) are given in Table 1 for each patient data and for all images in data set (last line
Fig. 6. (a) Decision regions in Bayesian classification: CSF (green), normal tissue (blue), lesions (yellow), (b) distribution of the input image (red), overlaid on its final estimation (blue).

Fig. 7. Result of applying the proposed algorithm to the image of a patient with a small lesion load: (a) input image, (b) manual segmentation of lesions, (c) result of fully automatic segmentation, and (d) extracted lesions after postprocessing.

Fig. 8. Result of applying the proposed algorithm to the image of a patient with a large lesion load: (a) input image, (b) manual segmentation of lesions, (c) result of fully automatic segmentation, and (d) extracted lesions after postprocessing.

of Table 1) and are compared with the similar values for the former automatic approach suggested in [19].

The results in this table show that the performance of the fully automatic segmentation suggested in this paper is better than that of the automatic approach. Comparing the two approaches, the fully automatic (proposed) approach has averagely improved 3.65%, 0.46% and 9.63% the criteria SI, OF and EF, respectively. However, the mean of the segmentation
time in the fully automatic approach is 3.75 s more than that of automatic approach. This longer time is expected, because in each iteration, of the fully automatic approach, in addition to the a priori probability, the mean and variance of each class should be updated.

As it is seen, the OF value is lower for the fully automatic approach in more than 50% of all cases. In these cases, the probable reason may be the presence of some single small lesions, which have vague borders and low contrast with the normal tissue. These lesions are not segmented well in fully automated segmentation. These types of lesions may be over estimated because of their confusion with normal tissues. Therefore, the false positive results increased and leaded to the higher EF compared to our previous segmentation method. The problem of OF and EF may be improved, using a preprocessing step, such as using a nonlinear anisotropic diffusion filter which enhances an image while preserves the edges [19].

Also, the EF value is higher for the fully automatic approach in 30% of all cases. In these cases, the probable reason may be the presence of the scattered confluent lesions, which usually have vague borders and can not be segmented well in fully automated segmentation. These types of lesions may be over estimated estimated because of their confusion with normal tissues. Therefore, the false positive results increased and leaded to the higher EF compared to our previous segmentation method. The problem of OF and EF may be improved, using a preprocessing step, such as using a nonlinear anisotropic diffusion filter which enhances an image while preserves the edges [19].

The mean values of similarity criteria in Table 1 are categorized to the three patient categories and new results have been shown in Table 2.

As can be seen in this table, SI, OF and EF are improved with an increase in lesion load. It is noticeable that lesion volumes segmented automatically are relatively smaller than those found by the experts. A human operator tends to be conservative and draws boundaries well around region of interest whereas the fully automatic method does not have this problem.

The results of volumetric comparison of lesions between the fully automated segmentation and the gold standard are presented from Table 3.

As it is seen in this table, according to the value of CC, accuracy of the fully automated segmentation is increased for patients with large lesion load.

4. Discussion

In this paper a new approach, for fully automatic segmentation of brain tissues in MR FLAIR images of MS patients,
is proposed. Calculation and optimization of CCPDFs and *a priori* probabilities of classes, which lead to improvement in our approach, are done by utilization of AMM (with no initially supplied training samples) and MRF, respectively. Our proposed approach is evaluated via similarity criteria (i.e., SI, OF, and EF) in a data set of MR FLAIR images of 20 MS patients. These results were compared with the results of automatic method [19] and also the results, previously, reported by other researchers such as Johnston et al. [3], Boudraa et al. [13], Leemput et al. [14], and Zijdenbos et al. [31], which used similar methods of evaluation (i.e., SI). It is reminded that these researchers have used manual segmentation for evaluation of their methods. We, too, used manual segmentation for evaluation. Therefore comparison of our method with these methods is reasonable. This comparison is done in Table 4.

As it is seen in Table 4, the MS lesion segmentation algorithm used in this paper improves 4%, 10%, 13%, 24%, and 7%, the results reported in [19,3,13,14,31]. Because the duration times for segmentation were not reported in the previous researches, mentioned above, no comparison was performed in this regard. Anbeek et al. [17] and Admiraal-Behloul et al. [18], made use of FLAIR images for segmentation of white matter lesions in patients with (Mean ± SD: 65.6 ± 7.7) years old. In comparison, we used FLAIR images for segmentation of MS lesions in younger patients (Mean ± SD: 29 ± 8).

The method proposed by Anbeek et al. [17] was a supervised pixel classification, in which a set of images was used as training set to build and tune the segmentation algorithm. To be truly effective, supervised training algorithms require a representative samples covering most of the cases (ideally all) in order to perform well in practice.

To initialize and guide the FCM algorithm, Admiraal-Behloul et al. [18] used brain templates, where prior distributions of the tissue types were supposed to be known. The success of a template-based segmentation algorithm depends on the outcome of the template. Compared to the above-mentioned methods our proposed algorithm does not need to any training set or any template.

For the patients with small lesion load, Anbeek et al. [17] and Admiraal-Behloul et al. [18], reached to values of 0.5 and 0.7 for SI, respectively, while, we obtained a value of 0.725 for SI, according to Table 5. However, averagely speaking, a decrease about 5% in value of SI for all patients was seen in our proposed approach, compared to Anbeek et al. [17] and no improvement attained compared to Admiraal-Behloul et al. [18].

Compared to Admiraal-Behloul et al. [18] method, our method improves the CC for patients with small and large lesion load about 2% and 1%, respectively, according to Table 3. No improvement in CC for patients with moderate lesion load was gained. Furthermore, the lesions smaller than six voxels were excluded by Admiraal-Behloul et al. [18], while, we do not exclude the lesions smaller than two voxels, until, the postprocessing condition is satisfied. If we ignore the lesions smaller than six voxels, the results of fully automated segmentation will be improved, because the possible lesions with sizes as small as one or two pixels are not usually considered as MS lesions by experts, in manual segmentation.

Finally, our findings about lesions load in FLAIR images, mentioned in Tables 2 and 3, are consistent with previous studies by Anbeek et al. [17] and Admiraal-Behloul et al. [18]. They suggested that better SI and CC were associated with bigger T2-w lesion loads.

Because of partial volume effect, the edges of the tissues or lesions are not well defined and consequently their correct delineation are not easy. This difficulty becomes more important when the operator delineates small or irregular lesions. Therefore, the correction of the partial volume effect is necessary [13]. The most prominent partial volume effect could be seen at the interface of lateral ventricles, especially in T2-w and PD images, and also in subarachnoid CSF spaces in T1-w enhanced images. Since we made use of FLAIR images and theoretically in FLAIR images, there is suppression of CSF signals, in these regions, we could ignore the partial volume artifact in our study. However, we expect using some corrective measures, such as morphological operators, connectivity principles and integration of explicit anatomical models of ventricles, which are useful and reduce this artifact [32]. As a future research, we intend to use fractal analysis for MS lesion segmentation and hope it makes better and more accurate results.

### 5. Summary

In this paper a new approach, based on a Bayesian classifier, for fully automatic segmentation of MS lesions in FLAIR-MR images is proposed. In the proposed method, distribution function of a FLAIR image is estimated by a mixture of a large number of normal terms by AMM. Then, the mixture terms are categorized into three classes, as the CCPDFs and the *a priori* probabilities of the classes. In the next steps, *a priori* probabilities of the classes as well as parameters of the classes (i.e.,

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SI</strong></td>
<td>≥ 0.65</td>
<td>≥ 0.62</td>
<td>0.51</td>
<td>0.68</td>
<td>≥ 0.71</td>
<td>≥ 0.75</td>
</tr>
</tbody>
</table>

### Table 5

*SI* values for the fully automatic method and the other methods

<table>
<thead>
<tr>
<th>Patient category</th>
<th>Anbeek et al. [17]</th>
<th>Admiraal-Behloul et al. [18]</th>
<th>Fully automatic</th>
</tr>
</thead>
</table>
means and variances) are attained and updated, utilizing MRF model and AMM, respectively, and without any need for training samples.

To compare the performance of the proposed approach with those of previous approaches including manual segmentation, the similarity criteria of different slices related to 20 MS patients were calculated. Also, volumetric comparison of lesions volume between the fully automated segmentation and the gold standard is performed using CC. The results showed a better performance for the proposed approach, compared to those of previous works.

Finally, our findings about lesions load in FLAIR images, are consistent with that of other researchers, who suggest that, the better SI and CC are associated with bigger T2-w lesions load.

Conflict of Interest Statement

None declared.

Acknowledgments

The authors wish to thank Dr. D. Etemadi of Koorosh Diagnostics and Medical Imaging Center, Tehran, Iran, for supplying the MRI data and manually segmented lesions, on which the algorithm was tested.

References


Rasoul Khayati was born in Salmass, Iran, in 1966. He received a M.Sc. degree from the Amirkabir University of Technology, Iran, in 1993. Currently, he is continuing his study toward Ph.D. at the same University. His research interests focus on biomedical image processing, medical imaging systems, and biological systems modeling.
Mansur Vafadust was born in Khash, Iran, in 1955. He received a M.Sc. degree from the Amirkabir University of Technology, Iran, in 1989 and a Ph.D., from the University of New South Wales, Australia, in 1994. Currently, he is a faculty member of Biomedical Eng. at the Amirkabir University of Technology. His research interests focus on biomedical image processing and medical imaging systems.

Farzad Towhidkhah was born in Kermanshah, Iran, in 1959. He received a M.Sc. degree from the Amirkabir University of Technology, Iran, in 1989 and a Ph.D., from the University of Saskatchewan, Canada, in 1996. Currently, he is a faculty member of Biomedical Eng. at the Amirkabir University of Technology. His research interests focus on biological system modeling, human motor control, and intelligent control.

Sayed Massoood Nabavi received his M.D. and national board of Neurology degrees from Tehran University of medical sciences, Iran from 1986 to 1998. He passed many pereceptorship and fellowship programs on Neurology and specially on MS in European countries from 1999 to 2006 (Germany and Austria). He is currently an active member of ENS (European Neurological Society), Prof. of Neurology and head of Neurology Department in Shahed University, Tehran, Iran. His research interests focus on different fields of Multiple Sclerosis and some other brain disorders.