Original Article

www.jmss.mui.ac.ir

A Hybrid Hierarchical Approach for Brain Tissue Segmentation by Combining Brain Atlas and Least Square Support Vector Machine

Keyvan Kasiri, Kamran Kazemi, Mohammad Javad Dehghani, Mohammad Sadegh Helfroush

Department of Electrical and Electronics Engineering, Shiraz University of Technology, Shiraz, Iran

Submission: 12-04-2013 Accepted: 12-09-2013

ABSTRACT

In this paper, we present a new semi-automatic brain tissue segmentation method based on a hybrid hierarchical approach that combines a brain atlas as *a priori* information and a least-square support vector machine (LS-SVM). The method consists of three steps. In the first two steps, the skull is removed and the cerebrospinal fluid (CSF) is extracted. These two steps are performed using the toolbox FMRIB's automated segmentation tool integrated in the FSL software (FSL-FAST) developed in Oxford Centre for functional MRI of the brain (FMRIB). Then, in the third step, the LS-SVM is used to segment grey matter (GM) and white matter (WM). The training samples for LS-SVM are selected from the registered brain atlas. The voxel intensities and spatial positions are selected as the two feature groups for training and test. SVM as a powerful discriminator is able to handle nonlinear classification problems; however, it cannot provide posterior probability. Thus, we use a sigmoid function to map the SVM output into probabilities. The proposed method is used to segment CSF, GM and WM from the simulated magnetic resonance imaging (MRI) using Brainweb MRI simulator and real data provided by Internet Brain Segmentation Repository. The semi-automatically segmented brain tissues were evaluated by comparing to the corresponding ground truth. The Dice and Jaccard similarity coefficients, sensitivity and specificity were calculated for the quantitative validation of the results. The quantitative results show that the proposed method segments brain tissues accurately with respect to corresponding ground truth.

Key words: *Atlas, brain, magnetic resonance imaging, segmentation, support vector machines*

INTRODUCTION

Progress in magnetic resonance imaging (MRI) techniques increases non-invasive study of the structure and functional organization of the brain. MRI provide excellent spatial resolution and tissue contrast and are thus ideally suited for morphological analysis of the brain. Brain tissue segmentation as a tool for delineation of the 3D anatomical structures or tissues plays an important role for numerous applications such as visualization and quantitative analysis of the brain. It can be used for studying the neuro-degenerative disorders such as the schizophrenia[1] or Alzheimer's disease,[2] characterizing morphological differences between subjects based on volumetric analysis of grey matter (GM), white matter (WM) and cerebrospinal fluid (CSF).[3-5] On the other hand, brain segmentation is a preliminary step for the other procedures such as brain registration, warping and voxel-based morphometry.[6]

Manual delineation of the brain structures or tissues from high-resolution 3D images is a tedious task in comparison to accurate and reliable automated segmentation methods. Nowadays, databases contain hundreds of cross-sectional and longitudinal MRI which may require several hours per scan for accurate manual segmentation of several structures. Such segmentation can be very error prone and exhibit nontrivial intra-expert variability during the segmentation of large databases over weeks which is known as "rater drift."[7] On the other hand, studies involving multiple raters face inter-expert variability because of complexity of this task.[8] Furthermore, iterative manual segmentation using the transverse, coronal and sagittal views may result in jagged boundaries, which cause difficulties in shape analysis. Hence, the fact that many applications depend on accurate, robust and cost-effective brain segmentation has inspired much work for developing automatic brain segmentation tools.

A number of techniques have been proposed for semi-automatic or automatic segmentation of brain tissues from cerebral MRI: Statistical-based segmentation,[9-12] geometrical-based segmentation,[13,14] atlas-based segmentation[15] and learning-based segmentation

Address for correspondence:
Dr. Kamran Kazemi, Department of Electrical and Electronics Engineering, Shiraz University of Technology, Shiraz, Iran. E-mail: kazemi@sutech.ac.ir

methods.[16] The use of *a priori* information is a successful concept for robust automatic tissue segmentation from structural MRI.[8] Currently, several common methods have been proposed based on applying probabilistic atlas as *a priori* information.[17-21] The atlas-based brain tissue segmentation starts by registering the brain atlas to the input image. The tissue class labels provided by registered brain atlas are used for the initial brain tissue segmentation.

Brain segmentation based on expectation-maximization (EM) which is originally proposed by Wells *et al*.,[12] in conjunction with brain atlas is a popular statistical classification scheme. This method has been improved by other researchers, e.g.,[17,19,20] to iteratively estimate tissue class label and bias field correction. Furthermore, EM framework has been extended to including a Markov random field (MRF)[19,20] as spatial constraint. However, machine learning methods such as EM and artificial neural network (ANN), which are based on Empirical Risk Minimization cannot control the learning model and therefore result in the phenomena of over-fitting and under-fitting.[22]

In this study, we present a semi-automatic brain tissue segmentation method from cerebral MRI based on support vector machine (SVM).[23] SVM is a supervised learning method that is used for classification and regression based on kernel methods. Recently, SVM has become a popular machine learning tool since it has shown excellent performance in many real-world applications such as a classification problem.[24-26] In contrast to linear classification methods, SVM maps the original parameter vectors into a higher (possibly infinite) dimensional feature space through a non-linear kernel function. Then, it tries to find an optimal hyperplane that minimizes the discrimination error for the training data. In comparison to other machine learning methods (e.g. Bays and ANN), SVM can transforms the problem into a quadratic programming (QP) one. Theoretically, QP will obtain global optimal solution so that it can overcome the local minimum problem. Furthermore, following Structural Risk Minimization principle, it can effectively overcome over-fitting and under-fitting problem and has greater generalization ability. On the other hand, SVM aims to obtain optimal solution under the circumstance of small-sample size, instead of infinite-sample size. Thus, SVM provides good generalization ability that can handle nonlinear classification problems such as brain tissue segmentation.[27] Suykens and Vandewalle[28] proposed a least-square (LS) type of SVM classifier (LS-SVM) by modifying the problem formulation to obtain a linear set of equations in the dual space. These properties suggest that applying LS-SVM may improve the overall accuracy of the brain tissue segmentation from cerebral MRI.

A number of different methods which use SVM solely or in an integrated approach for segmenting the MR brain images have been proposed in previous studies. Quddus *et al*.[29] and Lao *et al*.[30] applied the SVM for WM lesions segmentation from T1 and multi-parametric cerebral MRI. Guo *et al*.[27] applied multi-classification SVM with high dimensional feature vectors for segmentation of T2-MRI.[27] In these researches, training data were generated with the help of experts. Schnell *et al*.[31] developed a fully automated method for classification of high angular resolution diffusion imaging *in vivo* data based on using SVM. In this method, labeled *in vivo* training dataset is prepared using tissue masks created from the SPM5 segmentation of T1-images. Song *et al*.[32] proposed fuzzy nonlinear SVM in conjunction with intensity-based Markov priors for neonatal brain MRI segmentation. In,[33] a combination of atlas prior information, spatial information in MRF model and class probabilities produced by SVM is utilized for segmentation of mouse brain MRI. The MRF provides a statistical model to describe local spatial relationships between classes. Also, SVM was employed by Luts *et al*.[34] to segment and classify the brain tumors in MRI and MR spectroscopic imaging images.

The aim of the presented research is to propose an effective hierarchical semi-automatic segmentation method for classification of GM, WM and CSF using a LS-SVM supervised method in which training is performed using registered brain atlas. In this method, skull stripping, intensity non-uniformity correction and CSF segmentation are performed at the first two stages. Then, *a priori* information provided by a registered probabilistic brain atlas is used to train LS-SVM with radial basis function (RBF) kernel to segment GM and WM.

The rest of the paper is organized as follows. In Section II, materials and methods required for performing and evaluating the proposed method are introduced. Detailed simulation results are provided in Section III. Finally, concluded remarks and discussion are given in Section IV.

MATERIALS AND METHODS

MRI Brain Data

In order to assess the performance of the developed method, an extensive validation was performed based on two imaging data type: Simulated and real MRI.

Brainweb-Simulated Brain Data

The simulated 3D MRI (181 × 217 × 181 voxels of 1 mm^3 isotropic resolution) which are used as test data, are provided by the Brainweb simulated brain database from the McGill University (available from: http://www.bic.mni.mcgill.ca/brainweb). This database provides realistic simulations of MRI acquisition with different levels of intensity non-uniformity and noise. Simulated MRI are generated based on an anatomical model of a normal brain. Five datasets with different noise levels (n) range between 0% and 3% and intensity non-uniformity (rf) with 0%, 20% and 40% are used.

Internet Brain Segmentation Repository-Real Data

In order to evaluate the performance of the proposed method on real MR data, the cerebral MR datasets from 20 normal subjects provided by Center for Morphometric Analysis at Massachusetts General Hospital (Available at http://www.cma.mgh.harvard.edu/ibsr) (IBSR) are used. The real MR brain data and their hand-guided expert segmentation results are available at these datasets. The IBSR provides the performance results from five other automatic segmentation methods that make it convenient to compare the results with those reported by others. These 20 datasets involve different levels of difficulty such as low contrast scans, relatively smaller brain volumes and considerable intensity non-uniformity. This can make it possible to assess the effect of the signal-to-noise ratio, contrast-to-noise ratio, shape complexity and variations in size and intensity non-uniformity on the segmentation results.

LS-SVM

SVM as a subcategory of supervised learning methods is generally used for both classification and regression problems.[35,36] In SVM classifier, as introduced by Vapnik, a decision boundary is defined to classify a set of objects or features by generating input-output mapping functions using a set of labeled training data. The basic idea behind SVM is to use a set of kernels to map original feature space into a high dimensional feature space. Hence, it builds a non-linear discrimination boundary, i.e. complex curve, in the original space through creating an optimal linear discriminating boundary in the high dimensional feature space. Practical expressions are formulated in the dual space in terms of the related kernel function and the solution follows a (convex) QP problem. The LS version of the SVM classifier has been first proposed by Suykens and Vandewalle.[28] The aim of LS-SVM is to construct function y = f(x), which represents the dependence of the scalar output y_i on the input vector x_i given a set of N training data $\{(x_i, y_i)\}_{i=1}^{N}$. The LS-SVM takes the following form in the feature space:

$$y = \sum_{i=1}^{h} w_i \varphi_i(x) + b = w^T \varphi(x) + b;$$
$$w = [w_1, w_2, \ldots, w_h]^T, \varphi = [\varphi_1, \varphi_2, \ldots, \varphi_h]^T \quad (1)$$

where b is a bias term and $\varphi(x)$ is a non-linear mapping function which maps the input data into a higher dimensional feature space whose dimensionality can be infinite. In the LS-SVM, the optimization problem is defined by the following equations:

$$\min_{w,b,e} J_p(w,e) = \frac{1}{2} w^T w + \frac{1}{2} \gamma \sum_{i=1}^{N} e_i^2 \quad (2)$$

with constraints $d_i = w^T \varphi(x_i) + b + e_i$. The e_i is the error in the i^{th} training sample and γ is the penalty factor, which is the trade-off parameter between a smoother solution and the training error. A larger g usually results in higher training accuracy, which may cause to overfit the training data. Equation (2) shows two modifications in comparison to the Vapnik formulation: (1) The inequality constraints are replaced with equality constraints, (2) a squared loss function is taken for this error variable. The Lagrangian form to solve the constrained optimization problem (equation 2) in feature space (primal space) is as follows:

$$L(w,b,e;\alpha) = J_p(w,e) - \sum_{i=1}^{N} \alpha_K \{w^T \varphi(x_i) + b + e_i - y_i\} \quad (3)$$

where α_k values are the Lagrange multipliers. The solution for the constraint optimization problem in the dual space results in the following solution:

$$y(x) = \sum_{i=1}^{N} \alpha_i K(x, x_i) + b \quad (4)$$

Function $K(x_i, x_j)$ is the kernel defined as $K(x_i, x_j) = \phi(x_i)^T \phi(x_j)$ which performs the nonlinear mapping implicitly. The output of the LS-SVM is converted to a posteriori probability, i.e. rang [0,1], using the sigmoid function proposed by Platt[37] (Step 3 of Subsection 2.3).

Proposed Method

In order to increase the robustness of the segmentation method, we propose a hybrid hierarchical model based method to segment the input cerebral MRI into CSF, GM and WM. As shown in Figure 1, the method consists of three steps: (1) Skull striping, (2) CSF segmentation and (3) GM and WM segmentation. According to the significant contrast between gray level intensities of CSF and the other brain tissues, the hierarchical scenario provides powerful and flexible framework for brain tissue segmentation. Figure 2

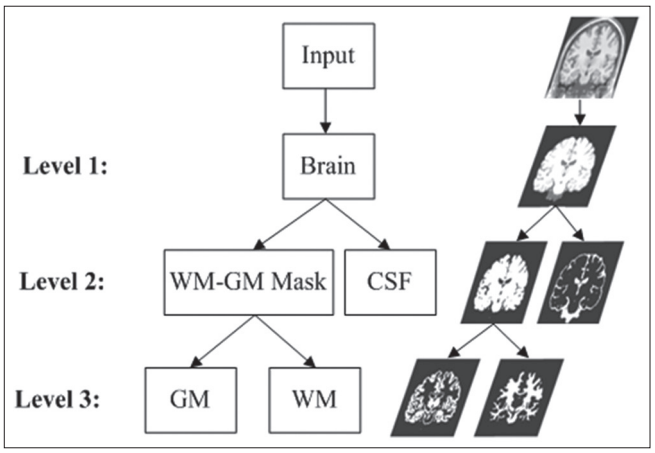

Figure 1: Hierarchical brain tissue segmentation. In the first step, skull stripping is performed to remove non-brain tissues. In the second step, the cerebrospinal fluid is segmented and in the third step, the grey matter and white matter are separated

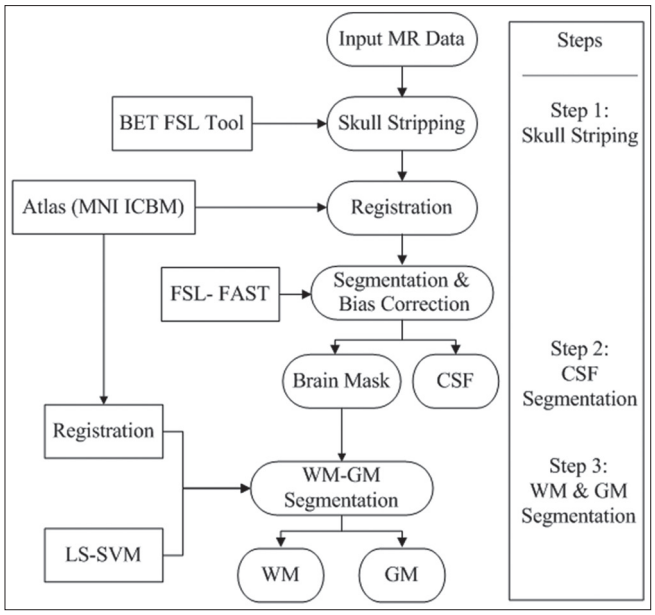

Figure 2: Block diagram of the proposed hybrid hierarchical least-square support vector machine method

illustrates the details of the proposed hybrid hierarchical method.

The first two steps of the proposed method consist of skull stripping, intensity non-uniformity correction and CSF segmentation which are performed using freely available software tools developed in the Oxford Centre for functional MRI of the Brain (FMRIB) (FSL version 4.1).[38,39] The FMRIB automated segmentation tool (FAST) is part of the FSL library, which is composed of modules for structural MRI analysis such as intensity non-uniformity correction and tissue segmentation.

Step 1: Skull Stripping

Intracranial segmentation commonly referred to as "skull-stripping," removes extracerebral tissue such as skull, eyeballs and skin. Skull stripping facilitates image processing such as surface rendering, cortical flattening, image registration and tissue segmentation. Thus, as the first step of the hierarchical method, we applied "Brain Extraction Tool" (BET)[38,39] integrated in FSL software to take an image of a head and remove all non-brain parts of the image.

Step 2: Intensity Non-uniformity Correction and CSF Segmentation

MRI suffers from the non-homogeneity in the radio-frequency field, which results in non-biological intensity non-uniformities across the acquired brain image. Thus, in the second step of the proposed hierarchical brain segmentation method, FAST toolbox integrated in FSL software is used to segment CSF, while also correcting for intensity non-uniformity. Accurate intensity non-uniformity correction requires segmentation knowledge while perfect segmentation requires uniformity in the intensity mapping of the image. The segmentation routine implemented in FAST toolbox is based on hidden MRF model and an associated EM algorithm. In this method, the histogram of the input image is modeled as a mixture of Gaussians with mean and variance for each class. The segmentation allows a reconstruction of the image; subtracting this from the real image gives an estimate of the non-uniformity. This whole process is then iterated between segmentation and intensity non-uniformity correction until reaching the convergence.[40] The resulting outputs are intensity non-uniformity corrected version and segmented CSF, GM and WM from the input data. In this step, the GM and WM masks are combined and the result is considered as a brain mask. This mask is then fed to the next level to segment GM and WM using LS-SVM classifier.

Step 3: GM and WM Segmentation

Segmentation of GM and WM from obtained GM-WM mask in the second step is performed using LS-SVM classifier. The training samples for learning the machine are chosen from *a priori* information available in registered brain atlas ICBM. The registration process consists of global and local steps. The registration is begins with 12-parameter affine transformation to correct position and overall shape differences between the input image and the atlas. Then, the affine registered atlas proceeds with a nonlinear registration using a spatial transformation model consisting of a linear combination of low-spatial frequency discrete cosine transform functions.[17,6] The hereby obtained registration parameters W are then used to map the GM and WM *a priori* information to the input image.

This step starts by assigning an initial label for one of the GM and WM classes, to each voxel inside GM-WM mask. The labeling process is achieved by converting the GM and WM *a priori* information provided by registered brain atlas into discrete version through assigning the most probable class to each voxel. The result of this step is initial approximation masks for GM and WM, which are used to select the training samples for each class. Owing to the large number of samples, only about 0.5% of the image voxels are selected for the training samples. Thus, reducing the computation time and computer memory usage will be achieved. The training samples are selected randomly in a uniform manner for all classes with equal number of samples in each class.

Besides designing the classifier, the segmentation result relies on the feature vector extracted from MRI data. The LS-SVM supervised learning method uses four dimensional feature vectors. Two types of features are used for training and test: Voxel intensity and voxel spatial information. Let X be the feature vector of input data to be segmented and Y be the associated target vector. As the first feature, the grey level intensity from the input MRI is selected. However, class overlapping between different tissues of the

MR signal is a critical issue when employing SVM for MRI segmentation.[33] Thus, the voxel coordinates (x, y, z) are selected as the second feature group for training and test. *A priori* information about tissue probability in each voxel is selected as the target vector during the training process. These vectors consist of two class probabilities: GM and WM.

$$X = \{I(x,y,z), x, y, z\}, Y = \{P_{GM}, P_{WM}\} \quad (5)$$

The output of this step for each voxel is a vector consisting of two elements of GM and WM probabilities. In this step, the KULeuven's LS-SVMlab MATLAB/C toolbox (available from: http://www.esat.kuleuven.ac.be/sista/lssvmlab) is used to handle the training and testing procedures. The RBF is chosen as the LS-SVM kernel as follows:

$$K(x, x_i) = \exp\left\{-\frac{||x - x_i||^2}{2\sigma^2}\right\} \quad (6)$$

where σ is the kernel variance.

After the training step, the feature vector $X=\{I(x,y,z),x,y,z\}$ is determined for each voxel of the input image and is used as input feature vector for trained LS-SVM to label the voxels of the input image. We consider the probability that each voxel belongs to each of the particular tissue types instead of assigning one tissue class to that voxel. However, the output of the SVM decision function does not provide posterior probabilities. Platt[37] proposed a method to map the SVM outputs into a posterior probability, i.e. range [0,1], by applying an additional sigmoid function on the outputs. Therefore, the sigmoid function (Eq 7) is used to map the LS-SVM output to the posterior probability presentation.

$$A(y_i, x_i) = \frac{1}{1 + \exp(\alpha f_k(x_i) + \beta)} \quad (7)$$

In this equation, α and β are determined from the training data and f_k is the LS-SVM decision function for the class k.

Evaluation Method

In order to quantitatively assess the accuracy of the proposed method, four commonly used measures, i.e. Dice coefficient, Jaccard coefficient, sensitivity and specificity, are used to compare the results of semi-automatic segmentation (A) with corresponding gold standard (G). The segmented probabilistic grey level coded masks are converted into binary masks by assigning the most probable tissue to each voxel. Furthermore, we compared our hybrid hierarchical LS-SVM technique with segmentation results of two other methods: FSL-FAST and hierarchical LS-SVM. In the latter method, after skull stripping using BET tool, all brain tissues consisting of CSF, GM and WM are segmented using LS-SVM classifier. In other words, after skull stripping and registration of the atlas to the input image, GM and WM *a priori* information are merged so as to provide the new prior information. Then, the classifier is designed to produce class probability for CSF and GM-WM. The input and output feature vectors for this step are as follows:

$$X = \{I(x,y,z), x, y, z\}, Y = \{\rho_{WM-GM}, \rho_{CSF}\} \quad (8)$$

The obtained GM-WM mask is then segmented into GM and WM with the same method as described in the previous subsection.

Similarity Metrics

Two similarity metrics are used for quantitative evaluation of the proposed method: Dice coefficient[41] or similarity index and Jaccard coefficient.[42] These metrics represent spatial overlap between two binary images and their values range between 0 (no overlap) and 1 (perfect agreement) as they are expressed as a percentage in the following:

$$D = \frac{2|A \cap G|}{|A| + |G|} \times 100 \quad (9)$$

$$J = \frac{|A \cap G|}{|A \cup G|} \times 100 \quad (10)$$

Success and Error Rate

The sensitivity (true positive fraction [TPF]) refers to the ability to correctly identify appropriate tissue in the segmented mask. It is defined as follows.

$$TPF = \frac{TP}{TP + FN} \quad (11)$$

The specificity (true negative fraction [TNF]) refers to the ability of the proposed segmentation method to correctly remove non-desired voxels.

$$TNF = \frac{TN}{TN + FP} \quad (12)$$

Sensitivity and specificity are the measures that are computed based on true positive, true negative, false negative and false positive.

RESULTS

In this section, the results of the proposed hybrid hierarchical model-based brain tissue segmentation method are presented. The LS-SVM parameters (penalty factor and kernel variance σ) and sigmoid parameters for mapping the LS-SVM output (α and β) were obtained through the simulations. The α and β were determined in such a way that we could have an appropriate mapping between the

LS-SVM output space and [0,1] range. Therefore, we opted α and β to be -5 and 2.5 during all experiments for the LS-SVM learning method. The penalty factor and σ were chosen to be 10 and 0.3.

Simulated Data Test

The first experiment was carried out based on Brainweb simulated data using the proposed hybrid hierarchical LS-SVM method, FSL-FAST method and hierarchical LS-SVM segmentation method. Figure 3 shows the input MRI, the ground truth, the results of the proposed method, the FSL-FAST and the hierarchical LS-SVM for 0%, 1%, 3% noise and 0%, 20%, 40% RF level for a selected slice. In this experiment, 3500 samples for each class were randomly selected for training. As the training and testing process of the learning algorithm directly depend on the training samples, results obtained through the segmentation procedure will slightly vary in different tests. Therefore, it is needed to run the whole process for a large number of times and pick the average of results for the assessment of the algorithm.

Table 1 presents quantitative results of the proposed hybrid hierarchical LS-SVM method for the simulated MRI (with different levels of noise and intensity non-uniformity) in terms of Dice and Jaccard similarity coefficients and sensitivity and specificity. Tables 2 and 3 illustrate the results of brain tissue segmentation using FSL-FAST and hierarchical LS-SVM, respectively. Figure 3 illustrates the results of the proposed method in comparison to other methods.

If the result of segmentation for a selected segmentation method can be classified as a high quality set of results by defining an arbitrarily lower bound on the Dice coefficient, such as Dice >0.90, then the proposed method demonstrates good performance for segmenting WM and GM from the simulated images. As can be seen, for the

Table 1: Quantitative results of the proposed hybrid hierarchical LS-SVM method of Brainweb datasets in term of D, J, sensitivity (TPF) and specificity (TNF)

Noise/rf (%, %)	CSF (%)				GM (%)				WM (%)			
	D	J	TPF	TNF	D	J	TPF	TNF	D	J	TPF	TNF
0, 0	84.65	73.39	99.79	93.85	93.38	85.49	92.64	95.44	94.87	88.80	93.26	98.72
1, 0	83.72	72.00	99.80	93.67	92.97	85.10	91.80	95.20	94.53	88.72	92.84	97.98
3, 0	83.83	72.17	99.74	93.54	92.56	84.96	90.05	95.71	93.73	88.93	92.67	97.61
3, 20	82.46	70.16	99.62	93.76	91.87	84.81	88.95	95.84	93.07	87.70	92.39	96.52
3, 40	82.64	70.42	99.53	93.91	91.07	83.50	87.29	95.14	92.72	87.44	92.21	96.29

CSF – Cerebrospinal fluid; GM – Grey matter; WM – White matter; LS-SVM – Least-square support vector machine; TNF – True negative fraction; TPF – True positive fraction; D – Dice coefficient; J – Jaccard coefficient

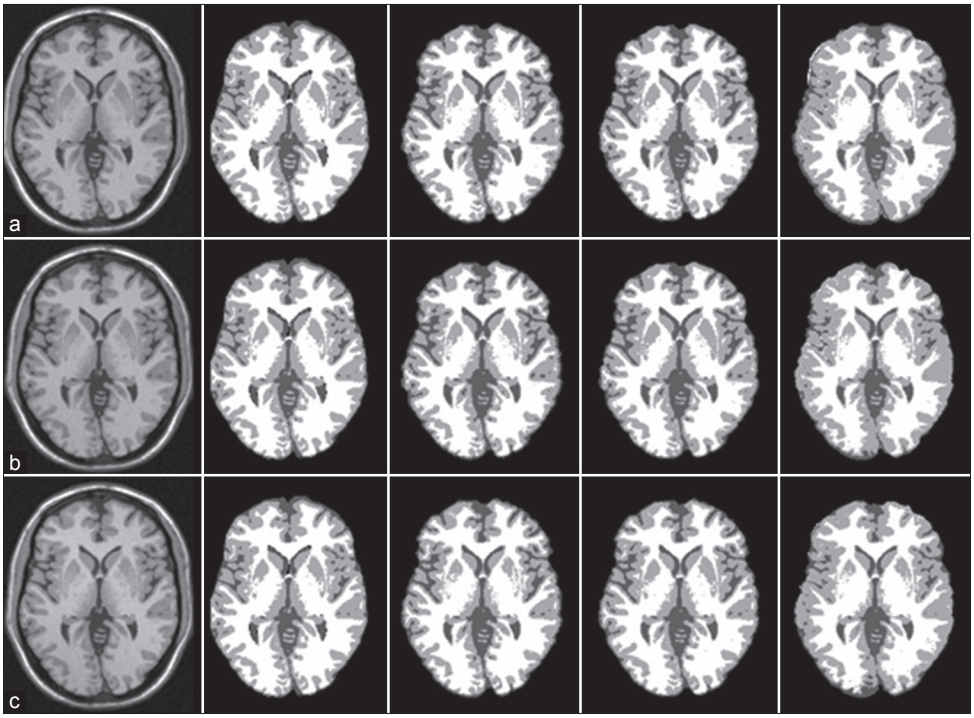

Figure 3: Segmentation results of 73rd slice of the Brainweb database using proposed hybrid hierarchical least-square support vector machine, FSL-FMRIB's automated segmentation tool and hierarchical LS-SVM methods

Table 2: Quantitative results of FSL-FAST segmentation of brainweb datasets

Noise/rf (%, %)	CSF (%)				GM (%)				WM (%)			
	D	J	TPF	TNF	D	J	TPF	TNF	D	J	TPF	TNF
0, 0	84.65	73.39	99.79	93.85	89.25	80.59	89.18	89.95	92.50	86.04	86.05	99.99
1, 0	83.72	72.00	99.80	93.67	90.17	82.10	88.80	92.22	94.33	89.26	89.34	99.95
3, 0	83.83	72.17	99.74	93.54	90.84	83.22	86.96	95.73	95.87	92.08	94.15	98.69
3, 20	82.46	70.16	99.62	93.76	90.79	83.14	86.92	95.58	95.69	91.73	94.20	98.37
3, 40	82.64	70.42	99.53	93.91	90.55	82.74	86.42	95.71	95.38	91.17	94.41	97.81

CSF – Cerebrospinal fluid; GM – Grey matter; WM – White matter; FAST – FMRIB automated segmentation tool; FMRIB – Functional magnetic resonance imaging of the brain's; TNF – True negative fraction; TPF – True positive fraction; D – Dice coefficient; J – Jaccard coefficient

Table 3: Quantitative results of the hierarchical LS-SVM algorithm of brainweb datasets

Noise/rf (%, %)	CSF (%)				GM (%)				WM (%)			
	D	J	TPF	TNF	D	J	TPF	TNF	D	J	TPF	TNF
0, 0	81.11	73.75	97.5	94.94	90.42	82.67	91.35	92.41	92.53	84.95	94.16	95.06
1, 0	79.93	72.54	96.80	92.86	89.02	81.07	89.73	91.81	91.76	84.03	93.97	94.58
3, 0	77.72	69.23	96.74	92.12	88.12	80.98	88.16	91.27	91.18	83.65	93.03	94.61
3, 20	76.03	68.09	95.51	91.31	86.83	79.09	87.20	91.76	90.24	81.12	92.18	94.01
3, 40	74.79	65.98	94.73	90.52	86.55	78.29	86.10	90.84	89.08	80.73	91.47	93.21

CSF – Cerebrospinal fluid; GM – Grey matter; WM – White matter; TNF – True negative fraction; TPF – True positive fraction; D – Dice coefficient; J – Jaccard coefficient; LS-SVM – Least-square support vector machine

proposed method, the minimum Dice similarity coefficient for segmented GM in different noise and intensity non-uniformity levels of the simulated MRI data is 91.07% whereas the lowest specificity coefficients were 89.95% and 90.84% for FSL-FAST and hierarchical LS-SVM respectively. The same results were obtained for Jaccard, sensitivity and specificity. Furthermore, quantitative comparison between the obtained results for segmented WM and CSF with different methods shows the same results for GM. These results confirm that the proposed method can accurately segment brain tissues in comparison with well-developed methods such as FSL-FAST.

Real Data Test

As the proposed method is designed to work for real medical applications, we applied the proposed method to segment the 20 normal subjects of IBSR T1-weighted brain scans. Since these datasets are clinical datasets which are the result of real MRI scanning, they involve different levels of difficulty and cover various problems related to segmentation of real MR data. Small brain volumes, sudden variations in gray level intensity, large intensity non-uniformity that is destructive in some of the datasets and large noise levels are the real factors affecting the segmentation process. This database is considered more noticeable and reliable for assessment of segmentation methods compared with the simulated databases.

During the segmentation process, 1100 samples per class were randomly chosen for training LS-SVM learning algorithm. Figures 4 and 5 show the original slices, hand-guided segmentations and the results of the proposed semi-automatic segmentation method, FSL-FAST and hierarchical LS-SVM for selected subjects 111-2 and 11-3, respectively.

Figure 6 illustrates the segmentation results in terms of the Dice similarity coefficient for the proposed hybrid hieratical LS-SVM method for GM and WM segmentation. In addition, the average results for all of the 20 datasets for the proposed method, FSL-FAST and hierarchical LS-SVM are shown in Table 4. As can be seen, the method can segment the GM and WM with mean similarity coefficient (Dice) of 78.96% and 78.22%, respectively. As a result, the proposed hybrid hierarchical method can segment GM and WM accurately in comparison to FSL-FAST and hierarchical LS-SVM. Furthermore, comparing the proposed method with the results of other techniques reported at the IBSR website [Table 5] shows that the proposed hybrid hierarchical method can yield an acceptable result in average for GM and WM segmentation based on Jaccard similarity metric.

The computational complexity of the core of the proposed method is loaded by the LS-SVM classifier. As reported in,[43] the computational complexity of the LS-SVM method is $O[l.n^2]$, where n is the number of training samples and l is the feature vector size. The advantage of this method in terms of computational complexity is to use a very short length feature vector and a small number of samples for the learning procedure. As n (the number of training samples) is chosen to be %0.5 of the total brain voxels N and only four features are selected, then the computational complexity in the training phase will be less than $O[N^2]$. In the test phase of the problem, since the SVM approach

Table 4: Quantitative evaluation of the proposed hybrid hierarchical LS-SVM brain tissue segmentation method in the IBSR datasets

Method	WM (%±standard)				GM (%±standard)			
	D	J	TPF	TNF	D	J	TPF	TNF
Hybrid hierarchical LS-SVM	78.22±3.50	63.95±4.79	83.10±2.99	98.83±0.21	78.96±3.70	64.57±5.33	70.92±5.03	99.33±0.09
FSL-FAST	77.00±9.88	63.42±10.81	87.80±15.37	98.52±0.44	75.63±5.48	61.11±6.91	65.10±6.98	99.55±0.54
Hierarchical LS-SVM	76.86±2.96	62.50±3.87	83.21±2.73	98.70±0.02	76.66±4.01	62.33±5.24	68.16±5.11	99.36±0.08

GM – Grey matter; WM – White matter; LS-SVM – Least-square support vector machine; FAST – FMRIB automated segmentation tool; FMRIB – Functional magnetic resonance imaging of the brain's; TNF – True negative fraction; TPF – True positive fraction; D – Dice coefficient; J – Jaccard coefficient; IBSR – Internet brain segmentation repository

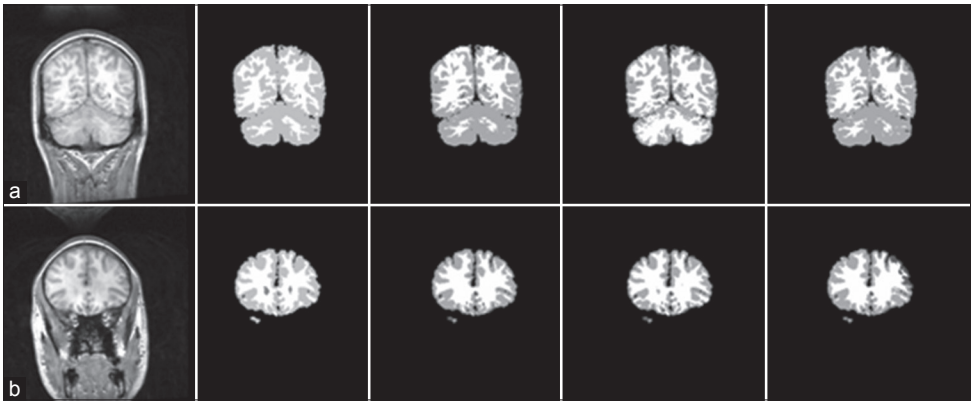

Figure 4: Sample slices from segmentation of internet brain segmentation repository 111-2 dataset using proposed hybrid hierarchical least-square support vector machine, FSL-FMRIB's automated segmentation tool and hierarchical LS-SVM methods

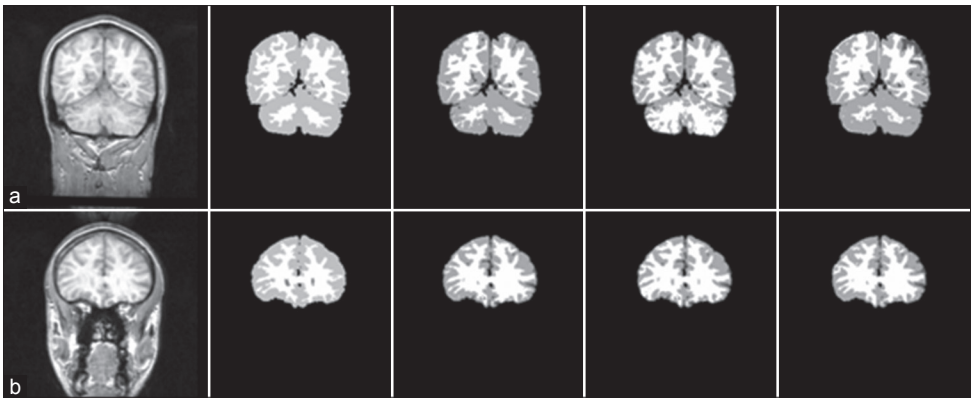

Figure 5: Sample slices from segmentation of internet brain segmentation repository 11-3 dataset using proposed hybrid hierarchical least-square support vector machine, FSL-FMRIB's automated segmentation tool and hierarchical LS-SVM methods

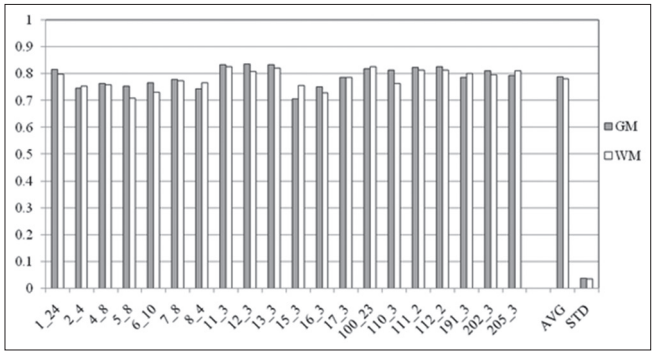

Figure 6: Dice coefficients of the grey matter and white matter segmentation for internet brain segmentation repository dataset using proposed hybrid hierarchical least-square support vector machine method

uses the kernel trick and an implicit mapping to the high dimensional space, the computational complexity of the problem for segmenting each voxel will only depend on the size of feature vector and therefore is not comparable to the complexity of the training procedure. Overall, the order of the computational complexity of the problem is determined by the training phase complexity, which will be less than $O[N^2]$.

The results attest to the similarity between the segmented GM and WM and the corresponding manually segmented tissues (i.e. the ground truth images). The Dice coefficient (D), Jaccard coefficient (J), sensitivity (TPF) and specificity (TNF) were computed after extracting

Table 5: Segmentations results for IBSR dataset in terms of average coefficient

Method	GM	WM
Adaptive MAP	0.564	0.567
Biased MAP	0.558	0.562
Fuzzy c-means	0.473	0.567
Maximum a posteriori probability	0.550	0.554
Maximum-likelihood	0.535	0.551
tree-structure k-means	0.477	0.571
FSL-FAST	0.611	0.634
Hierarchical LS-SVM	0.623	0.625
Hybrid hierarchical LS-SVM	0.645	0.639
Manual (4 brains averaged over 2 experts)	0.876	0.832

IBSR – Internet brain segmentation repository; GM – Grey matter; WM – White matter; MAP – Maximum aposteriori probability; LS-SVM – Least-square support vector machine; FAST – FMRIB automated segmentation tool; FMRIB – Functional magnetic resonance imaging of the brain's

the tissues of interest from T1-weighted MRI. In each section, the values reported in the "Hybrid Hierarchical LS-SVM" line were obtained by applying the proposed hybrid hierarchical LS-SVM method for brain tissue segmentation. Values reported in the "FSL-FAST" and "Hierarchical LS-SVM" lines were obtained by using FSL-FAST and Hierarchical LS-SVM methods for GM and WM segmentation.

DISCUSSION

In this paper, we presented a hybrid hierarchical method for semi-automatic model-based brain tissue segmentation from cerebral MRI. In the first step, the skull was removed. Then the CSF was extracted in the second step; these two steps were performed using the FSL-FAST tool. In the third step, the brain mask obtained from the first two steps was segmented into GM and WM using the LS-SVM classifier integrated in the KULeuven's LS-SVMlab MATLAB/C toolbox. On the other words, the first two steps are done in one environment and the results of them are used in other environment for the third step. The training process of the classifier was performed using two feature groups: Sample position and grey level intensity. The training samples were selected randomly from the registered GM and WM *a priori* information. Since the CSF and GM are of close intensities in the T1-weighted MRI data, simultaneous classification of them may cause more misclassification. Thus, applying a hierarchical method for removing the skull, CSF extraction and then GM and WM segmentation increase the segmentation accuracy.

In our segmentation method, the LS-SVM in conjunction with the brain atlas was used to segment the brain tissues. Since the LS-SVM classifier provides a good generalization over different types of the input data, acceptable segmentation accuracy was achieved. Comparing to the standard SVM which requires solving a QP problem, the LS-SVM problem could be formulated as a system of linear equations. Therefore, the proposed method is more efficient, especially for a large-scale problem. Thus, using the LS-SVM as the learning method instead of the traditional SVM could considerably decrease the overall computational cost. Furthermore, using the brain atlas for selecting training samples reduces manual intervention.

Since the training and testing process of the learning algorithm directly depend on the training samples, results obtained through the segmentation procedure will vary slightly in different tests. Furthermore, different approaches for choosing training samples may affect the performance of the algorithm. Dedicating a large part of the samples to the edges and boundaries will help the learning algorithm accurately segment the boundary regions. However, this is at the cost of losing the generality of the learning method and occurring misclassification in smooth areas. In this study, after performing several tests over various sampling methods while keeping the balance between class samples, uniform random selection of training samples were opted to train the learning machine.

To validate the proposed method, two different databases were used: The Brainweb simulated data and the IBSR real MRI data. These two databases include different images with different degrees of noise level and intensity non-uniformity. Visual inspection of the segmented GM and WM demonstrated that the method can accurately extract the major tissues of the GM and WM. The quantitative results showed that the proposed hybrid hierarchical model-based method can extract the location of the brain tissues in the head with a high degree of accuracy. Figure 7 illustrates the average Dice similarity coefficient between the segmented GM and WM from various images with different qualities in the databases along with their corresponding ground truth. Furthermore, the corresponding sensitivity and specificity coefficients [Tables 1 and 4] confirmed that the new technique can segment GM and WM accurately.

As can be seen in the results and as concluded in Figure 7, it is obvious that segmentation using the FSL-FAST leads to more accurate CSF extraction [Table 1]. Results shown in the Figure 7 prove that applying FSL-FAST tool for the skull stripping and CSF extraction in the first two steps and then feeding the remaining brain mask to the LS-SVM classifier for the GM/WM segmentation lead to accurate segmentation. Although FSL-FAST technique offers more accurate results than the hierarchical LS-SVM method for the whole process, the combination of them leads to an increase in the accuracy of the segmented Brainweb datasets tissues. It is deduced that, precise CSF extraction can cause considerable as improvement in GM separation. However, a slight decline in the accuracy of WM segmentation for simulated datasets can be observed.

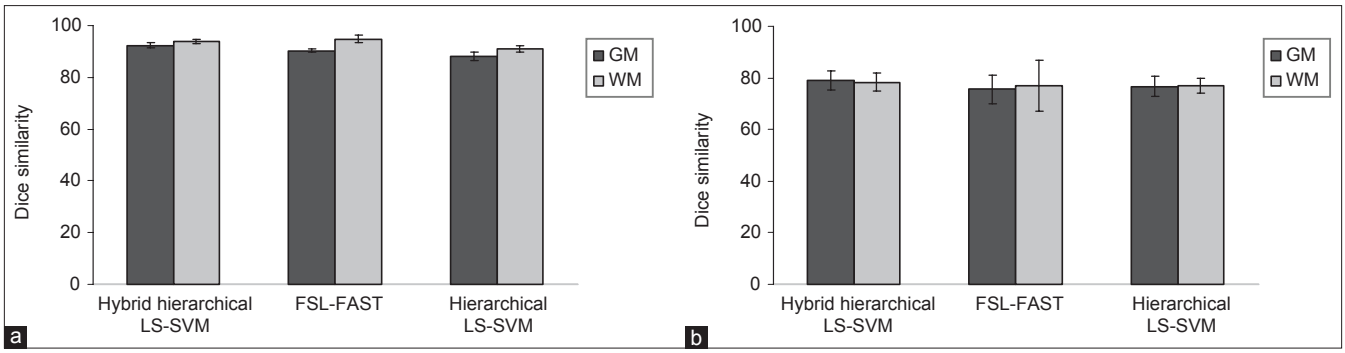

Figure 7: Comparison between the mean Dice coefficients values for grey matter and white matter segmentation with the proposed hybrid hierarchical least-square support vector machine method and two methods, FSL-FMRIB's automated segmentation tool and hierarchical LS-SVM. Bar charts present the mean Dice coefficients values for segmentation in terms of (a) Brainweb and (b) Internet brain segmentation repository datasets

The experimental results based on real clinical data [Table 4] show that the proposed hybrid hierarchical method provides better results in comparison with FSL-FAST and hierarchical LS-SVM. The obtained similarities between the segmented WM and the corresponding ground truth based on Dice similarity metric are 78.22 ± 3.50, 77.00 ± 9.88 and 76.86 ± 2.96 for the proposed method, FSL-FAST and the hierarchical LS-SVM, respectively. The similarity coefficients for GM are 78.96 ± 3.70, 75.63 ± 5.48 and 76.66 ± 4.01, respectively. These results show less variation in the result of segmenting different brain scans as well as improvement in the similarity between segmented tissues and their corresponding ground truth. This fact verifies the robustness of the proposed method in segmenting MRI with different characteristics of noise and intensity non-homogeneities. This is evident because of using the SVM classifier that provides good generalization in comparison to other classifiers such as maximum likelihood.

CONCLUSION

The new hybrid hierarchical model-based brain tissue segmentation technique can accurately segment the brain tissues from the cerebral MRI images. The results based on simulated MRI on different noise and intensity non-uniformity levels shows that the proposed method segments the GM better than others and offers a marginal improvement for WM. However, it gives better results for IBSR real database. Therefore, the obtained results suggest that the LS-SVM is a promising technique for image classification in a medical imaging application. This new approach can be used as a part of a neurological tissue analysis framework, such as statistical morphological analysis and head model creation for source localization.

ACKNOWLEDGMENTS

We would like to welcome the FMRIB Analysis at University of Oxford for providing the FSL software used in the automatic segmentation of the images. The synthetic MRI were provided by the McConnell Brain Imaging Center of the Montreal Neurological Institute, McGill University and are available at http://www.bic.mni.mcgill.ca/brainweb/. The 20 normal MR brain datasets and their manual segmentations were provided by the Center for Morphometric Analysis at Massachusetts General Hospital and are available at http://www.cma.mgh.harvard.edu/ibsr/.

REFERENCES

1. Tanabe JL, Amend D, Schuff N, DiSclafani V, Ezekiel F, Norman D, *et al*. Tissue segmentation of the brain in Alzheimer disease. AJNR Am J Neuroradiol 1997;18:115-23.
2. Apostolova LG, Dinov ID, Dutton RA, Hayashi KM, Toga AW, Cummings JL, *et al*. 3D comparison of hippocampal atrophy in amnestic mild cognitive impairment and Alzheimer's disease. Brain 2006;129:2867-73.
3. Lawrie SM, Abukmeil SS. Brain abnormality in schizophrenia. A systematic and quantitative review of volumetric magnetic resonance imaging studies. Br J Psychiatry 1998;172:110-20.
4. McCarley RW, Wible CG, Frumin M, Hirayasu Y, Levitt JJ, Fischer IA, *et al*. MRI anatomy of schizophrenia. Biol Psychiatry 1999;45:1099-119.
5. Shenton ME, Kikinis R, Jolesz FA, Pollak SD, LeMay M, Wible CG, *et al*. Abnormalities of the left temporal lobe and thought disorder in schizophrenia. A quantitative magnetic resonance imaging study. N Engl J Med 1992;327:604-12.
6. Ashburner J, Friston KJ. Voxel-based morphometry-The methods. Neuroimage 2000;11:805-21.
7. Spinks R, Magnotta VA, Andreasen NC, Albright KC, Ziebell S, Nopoulos P, *et al*. Manual and automated measurement of the whole thalamus and mediodorsal nucleus using magnetic resonance imaging. Neuroimage 2002;17:631-42.
8. Cocosco CA, Zijdenbos AP, Evans AC. A fully automatic and robust brain MRI tissue classification method. Med Image Anal 2003;7:513-27.
9. Bezdek JC, Hall LO, Clarke LP. Review of MR image segmentation techniques using pattern recognition. Med Phys 1993;20:1033-48.
10. Held K, Rota Kops E, Krause BJ, Wells WM 3rd, Kikinis R, Müller-Gärtner HW. Markov random field segmentation of brain MR images. IEEE Trans Med Imaging 1997;16:878-86.
11. Liew AW, Yan H. An adaptive spatial fuzzy clustering algorithm for 3-D MR image segmentation. IEEE Trans Med Imaging 2003;22:1063-75.
12. Wells WM, Grimson WE, Kikinis R, Jolesz FA. Adaptive segmentation of MRI data. IEEE Trans Med Imaging 1996;15:429-42.
13. Li C, Kao CY, Gore JC, Ding Z. Minimization of region-scalable fitting energy for image segmentation. IEEE Trans Image Process 2008;17:1940-9.
14. Wang L, Chen Y, Pan X, Hong X, Xia D. Level set segmentation of brain magnetic resonance images based on local Gaussian distribution fitting energy. J Neurosci Methods 2010;188:316-25.

15. Collins DL, Holmes CJ, Peters TM, Evans AC. Automatic 3D model-based neuroanatomical segmentation. Hum Brain Mapp 1995;3:190-208.
16. Hall LO, Bensaid AM, Clarke LP, Velthuizen RP, Silbiger MS, Bezdek JC. A comparison of neural network and fuzzy clustering techniques in segmenting magnetic resonance images of the brain. IEEE Trans Neural Netw 1992;3:672-82.
17. Ashburner J, Friston KJ. Unified segmentation. Neuroimage 2005;26:839-51.
18. Marroquin JL, Vemuri BC, Botello S, Calderon F, Fernandez-Bouzas A. An accurate and efficient Bayesian method for automatic segmentation of brain MRI. IEEE Trans Med Imaging 2002;21:934-45.
19. Van Leemput K, Maes F, Vandermeulen D, Suetens P. Automated model-based bias field correction of MR images of the brain. IEEE Trans Med Imaging 1999;18:885-96.
20. Van Leemput K, Maes F, Vandermeulen D, Suetens P. Automated model-based tissue classification of MR images of the brain. IEEE Trans Med Imaging 1999;18:897-908.
21. Zhou Y, Bai J. Atlas-based fuzzy connectedness segmentation and intensity nonuniformity correction applied to brain MRI. IEEE Trans Biomed Eng 2007;54:122-9.
22. Wu Y, Guo L. Multi-layer support vector machine and its application. In: International Conference IEEE Transactions on Machine Learning and Cybernetics.; 2006. p. 3627-31.
23. Cristianini N, Shawe-Taylor J. An Introduction to Support Vector Machines. Cambridge, UK: Cambridge Univ. Press; 2000.
24. Osuna E, Freund R, Girosi F. Training support vector machines: Application to face detection. In: Proceeding Computer Vision and Pattern Recognition. Conf Proc IEEE Computer Society 1997; 1997:130-6.
25. Pontil M, Verri A. Support vector machines for 3-D object recognition. IEEE Trans Pattern Anal Mach Intell 1998;20:637-46.
26. Scholkopf B, Kah-Kay S, Burges CJ, Girosi F, Niyogi P, Poggio T, *et al*. Comparing support vector machines with Gaussian kernels to radial basis function classifiers. IEEE Trans Signal Process 1997;45:2758-65.
27. Guo L, Liu X, Wu Y, Yan W, Shen X. Research on the segmentation of MRI image based on multi-classification support vector machine. Conf Proc IEEE Eng Med Biol Soc 2007;2007:6020-3.
28. Suykens JA, Vandewalle J. Least squares support vector machine classifiers. Neural Process Lett 1999;9:293-300.
29. Quddus A, Fieguth P, Basir O. Adaboost and support vector machines for white matter lesion segmentation in MR images. Conf Proc IEEE Eng Med Biol Soc 2005;1:463-6.
30. Lao Z, Shen D, Liu D, Jawad AF, Melhem ER, Launer LJ, *et al*. Computer-assisted segmentation of white matter lesions in 3D MR images using support vector machine. Acad Radiol 2008;15:300-13.
31. Schnell S, Saur D, Kreher BW, Hennig J, Burkhardt H, Kiselev VG. Fully automated classification of HARDI *in vivo* data using a support vector machine. Neuroimage 2009;46:642-51.
32. Song Z, Awate SP, Licht D, Gee JC. Clinical neonatal brain MRI segmentation using adaptive nonparametric data models and intensity-based Markov priors. Springer, Medical Image Computing and Computer Assisted Intervention (MICCAI), 2007.
33. Bae MH, Wu T, Pan R. Mix-ratio sampling: Classifying multiclass imbalanced mouse brain images using support vector machines. Expert Syst Appl 2010;37:4955-65.
34. Luts J, Laudadio T, Idema AJ, Simonetti AW, Heerschap A, Vandermeulen D, *et al*. Nosologic imaging of the brain: Segmentation and classification using MRI and MRSI. NMR Biomed 2009;22:374-90.
35. Vapnik V. The Nature of Statistical Learning Theory. New York: Springer-Verlag; 1995.
36. Vapnik V, Golowich S, Smola A. Support vector method for function approximation, regression estimation, and signal processing. In: Mozer M, Jordan M, Petsche T, editors. Neural Information Processing Systems. Vol. 9. Cambridge, MA: MIT Press; 1997.
37. Platt JC. Probabilistic output for support vector machines and comparisons to regularized likelihood methods. Advances in Large Margin Classifiers. Cambridge, MA: MIT Press; 1999.
38. Smith SM. Fast robust automated brain extraction. Hum Brain Mapp 2002;17:143-55.
39. Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE, Johansen-Berg H, *et al*. Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage 2004;23 Suppl 1:S208-19.
40. Zhang Y, Brady M, Smith S. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. IEEE Trans Med Imaging 2001;20:45-57.
41. Dice LR. Measures of the amount of ecologic association between species. Ecology 1945;26:297-302.
42. Jaccard P. The distribution of flora in the alpine zone. New Phytol 1912;11:37-50.
43. Available from: http://www.esat.kuleuven.be/sista/lssvmlab/downloads/tutorialv1_8.pdf.

How to cite this article: Kasiri K, Kazemi K, Dehghani MJ, Helfroush MS. A Hybrid Hierarchical Approach for Brain Tissue Segmentation by Combining Brain Atlas and Least Square Support Vector Machine. J Med Sign Sens 2012;3:195-232-43.

Source of Support: Nil, **Conflict of Interest:** None declared

BIOGRAPHIES

Keyvan Kasiri received the B.Sc. degree in Electrical and Computer Engineering (ECE) from the IsfahanUniversity of Technology, Iran, in 2007 and the M.Sc. degree in Electrical Engineering from the Shiraz University of Technology, Shiraz, Iran in 2010. He joined the Vision and Image Processing Research Group, University of Waterloo, Canada, in 2011.Keyvan is currently working towards the Ph.D. degree in the Department of Systems Design Engineering at the University of Waterloo. His research interests include different aspects of signal and image processing, analysis of biomedical images, pattern recognition,and machine learning.

E-mail: kkasiri@uwaterloo.ca

Kamran Kazemi received the B.S. and M.S. degrees in Electrical engineering from Ferdowsi University of Mashhad and K. N. Toosi University of technology, Tehran in 2000 and 2002, respectively. He performed his Ph.D. degree in Electrical Engineering and Biomedical Engineering as cooperation between K. N. Toosi University of Technology, Tehran, Iran and University of Picardie Jules Verne (UPJV) Amiens, France. Actually he is working as assistant professor in department of Electrical and Electronics Engineering, Shiraz University of Technology, Shiraz, Iran. His research interests include different aspects of biomedical signal and image processing, pattern recognition, and machine learning.

E-mail: kazemi@sutech.ac.ir

Mohammad Javad Dehghani received the BS and MS degrees from Isfahan University of Technology, Iran, in 1988 and 1992 respectively, the Ph.D. degree from IIT Madras in 2003, all in communication engineering. In 1993, he joined the faculty of Electronic Department, School of Electronic Technology, Shiraz University, Iran. Since 2003 he has been on the faculty of the Department of Electrical Engineering, Shiraz University of Technology, Shiraz, Iran. His activities have included Image and signal processing, digital filter structures, filter banks, wavelet based signal processing and wireless communication. He is a Member of the IEEE.

E-mail: dehghani@sutech.ac.ir

Mohammad Sadegh Helfroush received the B.S. and M.S. degrees in Electrical engineering from Shiraz University of Shiraz and Sharif University of technology, Tehran in 1993 and 1995, respectively. He performed his Ph.D. degree in Electrical Engineering from Tarbiat Modarres University, Tehran, Iran. Actually he is working as associate professor in department of Electrical and Electronics Engineering, Shiraz University of Technology, Shiraz, Iran. His research interests include different aspects of signal and image processing, pattern recognition, and machine learning.

E-mail: ms_helfroush@sutech.ac.ir