Automatic Multiple Sclerosis Detection based on Integrated Square Estimation

Jundong Liu

School of Elec. Eng. & Comp. Sci.
Ohio University
Athens, OH 45701

Charles D. Smith, Himachandra Chebrolu

Department of Neurology
University of Kentucky
Lexington, KY 40506

Abstract

This paper presents a fully automatic method for segmentation of Multiple Sclerosis (MS) lesions from multiple sequence MR (T2-weighted and FLAIR) images. Our method treats MS lesions as outliers to the normal brain tissue distribution, and the separation is achieved by minimizing a statistically robust L_2E measure, which is defined as the squared difference between the true density and the assumed Gaussian mixture. Pre- and post-processing procedures including intensity normalization and false positive pruning are designed to remove various signal artifacts. Our method is fully automatic and doesn't require any training, atlas or thresholding steps. The results of our method are compared with lesion delineations by human experts, and a high classification accuracy is demonstrated on 16 datasets containing small to moderate lesion loads.

1 Introduction

Multiple Sclerosis (MS) is a central nervous system (CNS) disease where myelin sheathes of the neurons are destroyed by the immune system. This disease is associated with brain tissue damage (e.g. lesions) that can be observed through Magnetic Resonance Imaging (MRI). These lesions can appear as a hyperintense signal or as a hypointense signal depending on its properties and on the use MRI sequence. Previous research has shown that the FLAIR (Fluid Attenuated Inversion Recovery) sequence contains the most distinctive lesion-healthy tissue differentiation for segmentation of white matter lesions. The radiological criteria for MS include the number of lesion on the MRI, their locations and their sizes, and these quantitative information is also crucial for studying the progression of MS lesions and the effect of drug treatments.

To segment the lesions from MRIs is the prerequisite step for performing various quantitative analyses of the disease. Manual lesion segmentation has been a dominate approach in the past, but it is a tedious and labor intensive task and the segmentation results are inevitably affected by intra and inter-expert variabilities. Automatic segmentation is greatly desired due to its high levels of reproducibility and reliability. For this reason, a number of researchers have been working on developing automatic lesion segmentation solutions in the past decade.

1.1 Related Work

A variety of approaches to MS lesion segmentation have been proposed in the literature. Generally speaking, they can be classified into two groups: outlier-based and class-based methods.

In outlier-based methods [8, 12, 17, 6, 7], MS lesions are treated and detected as the outliers to the normal brain tissue distribution, which is usually modelled with a Finite Gaussian Mixture (FGM) of CSF, GM and WM classes. Van Leemput *et. al.* [8] pioneered this approach. Under their framework, MR field inhomogeneities, parameters of the Gaussian distribution and membership are computed iteratively, with the contextual information being incorporated using a Markov random field. Observed intensity values whose Mahalanobis-distances exceed a predefined threshold are marked as lesions. The thresholds are empirically set in this work.

The approach proposed by Prastawa *et. al.* [12] combines outlier detection and region partitioning together. The salient point of this approach lies in the way contextual information is incorporated: tissue typing is carried out based on regions (connected groups of voxels) instead of individual voxels. Voxels labeling is conducted through maximizing overall relative entropy or Kullback-Leibler divergence between neighboring regions. Samples with Mahalanobis distance greater than a manually chosen threshold are treated as outliers. Souplet *et. al.* [17, 6] also model lesions as a outlier class. A segmentation of the brain into CSF, GM and WM is first performed on the T1-

978-1-4244-3993-5/09/$25.00 ©2009 IEEE

weighted (T1w) and T2-weighted (T2w) sequences. Lesions are then detected through a thresholding step on T2w + FLAIR sequences. A similar procedure was adopted in Garcia-Lorenzo *et. al.* [7, 2]. Their solution consists of three steps: 1) robust estimation of normal appearing brain tissue (NABT) parameters, 2) refinement of outlier detection and 3) application of lesion rule. After the NABT distribution is obtained through a robust Expectation Maximization (EM) algorithm, each voxel in the image can be labeled as an outlier (candidate for lesions) if the Mahalanobis distance for each class is great than a given p-value. The final lesion rule step refines the segmentation result by discriminating White Matter Lesions (WML) from false positives within other tissues and pruning the latter. Manually chosen thresholds are involved in both the outlier detection and WML separation steps.

Class-based methods [22, 3, 4, 5, 11, 19] model the lesions as an independent class to be extracted. In [22], a combination of intensity-based k-nearest neighbor classification (k-NN) and template-driven segmentation (TDS+) was designed to segment different types of brain tissue. Lesions are modeled as one of the expected tissue types, and the class parameters are obtained through a operator-supervised voxel sampling on two randomly selected scans. Since the manual training step is highly data-dependent, it is expected be conducted for each study or data set. A similar approach was proposed in [3]. The segmentation method determines for each voxel in the image the probability of being part of MS-lesion tissue, and the classification is conducted also based on K-NN. Voxel intensities and spatial information are integrated as discriminative features, and voxels are classified based on their proximity to the pre-classified samples in the feature space. It should be noted that manual or semiautomatic training is normally a required step in k-NN based methods, and the value of k (number of classes) has to be determined in advance, either interactively [22] or empirically [3].

Atlas-assisted segmentation framework proposed in [5, 11, 19] makes use of the relative consistent continuity and relationship residing in neighboring anatomical structures within the same group of subjects. Lesions are treated as a subclass within the White Matter tissue, and a topology preservation criterion is employed to guarantee strict topological equivalence between the segmented image and the atlas. However, to ensure the strict correspondence between atlas and the patient images, an atlas that accurately represents the group subjects is likely to be required for each study.

One should note that in the class-based approaches [22, 3, 16, 5, 11, 19], a training procedure, to either calibrate the classifier parameters or to choose the tissue class representatives, is normally required. In order to obtain desired segmentation results, the testing data sets are also expected to be highly similar to the training sets, ideally from the same group. Outlier-based models [8, 12, 17, 6, 7, 2] relax the training requirement, but they usually subsume a thresholding step. Those thresholds, critical for segmentation performance and system reproducibility, usually require certain prior to be set up precisely therefore are often difficult to be determined.

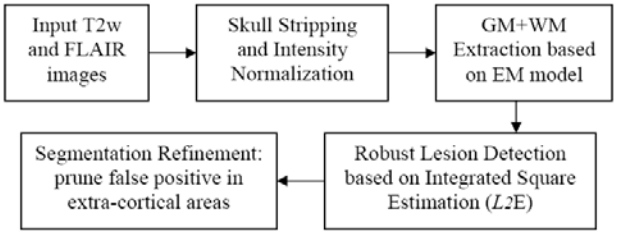


Figure 1: Work flow of our MS lesion detection algorithm

In this paper, we develop a fully automatic method for MS lesions that requires no training, atlas, or thresholding steps. Our method can be regarded as a combination of the class-based and outlier-based approaches. The core algorithm consists of three steps, and the separation of the lesion class from other normal tissue types is achieved by minimizing a statistically robust measure called L_2E criterion, which is defined as the squared difference between the true density and the assumed Gaussian mixture.

2 Method

The method proposed in this article segments MS lesions from two MRI sequences: T2-weighted and FLAIR. The overall procedure is comprised of a sequence of steps, as depicted in Figure 1. Firstly, two preprocessing steps, skull stripping and intensity normalization, allow to reduce image artifacts and to focus on a region of interest. Secondly, an Expectation-Maximization (EM) classification is applied to prune the background and CSF and extract the GM + WM portion of the image. Then, a robust detection algorithm based on Integrated Squared Estimation (L_2E) is applied to separate the lesions from the normal WM + GM portion. Finally, false positives in extra-cortical areas are removed through a refinement step.

Skull stripping deletes the non-brain tissue from the image to simplify the following lesion segmentation. It is performed using the Brain Extraction Tool (BET) of FSL library in our work. Intensity normalization step is to correct the RF field inhomogeneities existing in the acquired MRIs and we used SPM2 for this purpose.

To locate the lesions, the CSF portion of an image needs to be extracted and eliminated. This step is based on FLAIR

images alone, and an EM-based segmentation routine is applied.

2.1 Extract MS lesions from GM+WM

Statistical brain segmentation algorithms usually assume the intensities of the normal brain tissues, CSF, GM and WM, conform to a Finite Gaussian Mixture (FGM) distribution. MS lesion could be directly added as an extra class to the FGM models, but they tend to vary in size, shape and location across different stages of the disease, and not to follow a multivariate Gaussian distribution when multichannel images are used. In light of these difficulties, most outlier-based methods avoid modeling lesions and simply treat them as outliers to the normal tissue types. However, the thresholds to define outliers are either manually or empirically set in these methods, which tend not to work consistently across different data sets.

The new MS lesion segmentation method developed in this paper provides a remedy to the above mention drawbacks. Our method also treats the lesions as outliers, but the separation between normal tissue clusters and outliers is automatically obtained. No manually chosen threshold is involved.

2.1.1 Integrated Square Estimation (L_2E)

In [15], L_2 distance has been investigated as an estimation tool for a variety of parametric statistical models. Estimation through the minimization of the integrated square error, or L_2E error, is shown to be inherently robust. Since then, several works have been published in applying L_2 measure for image registration [9, 20]. A brief introduction of the L_2E measure is given as follows.

Suppose $y(x)$ is an unknown density function. The parametric approximation of $y(x)$ is $\hat{y}(x|\theta)$. The L_2E minimization estimator for θ is given as:

$$\begin{aligned} \hat{\theta}_{L_2E} &= \arg\min_{\theta} \int [\hat{y}(x|\theta) - y(x)]^2 dx \qquad (1) \\ &= \arg\min_{\theta} \int [\hat{y}^2(x|\theta) - 2\hat{y}(x|\theta)y(x) + y^2(x)] dx \end{aligned}$$

Observing that $y^2(x)$ doesn't contain any θ term, it can thus be dropped from the functional minimization in equation (1). Considering $y(x)$ is a density function, $\int \hat{y}(x|\theta)y(x)$ can therefore be viewed as the expectation of $\hat{y}(x|\theta)$. Putting these two considerations together, equation (1) can be rewritten as:

$$\hat{\theta}_{L_2E} = \arg\min_{\theta} \left[\int \hat{y}^2(x|\theta) dx - \frac{2}{n} \sum_{i=1}^{n} \hat{y}(x_i|\theta) \right] \quad (2)$$

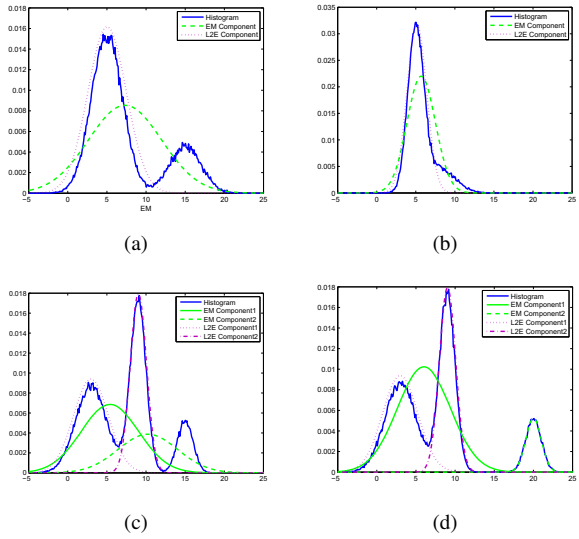


Figure 2: Parameter estimation using EM and FGML2E. True distributions: (a) $0.7N(5,2) + 0.3N(15,2)$; (b) $0.7N(5,1) + 0.3N(8,2)$; (c) $0.41N(3,2) + 0.41N(9,1) + 0.18N(15,1)$; (d) $0.41N(3,2) + 0.41N(9,1) + 0.18N(20,1)$ (Figures are better seen on screen than in black/white print)

2.1.2 Inherent Robustness Properteis of L_2E in Fitting Finite Gaussian Mixture (FGM)

In FGM models, the pixel intensity values for each tissue type are assumed to conform to a Gaussian distribution. Expectation-Maximization (EM) algorithm, the estimator of Maximum Likelihood (ML) measure, is currently a popular solution in many FGM-based segmentation algorithms. However, ML and EM are inherently not robust and potentially influenced by input outliers. In this paper, we adopt the L_2E as the FGM estimator. To demonstrate the robustness of L_2E with respect to outliers, which also serves as the motivation of our model, a group of experiments have been conducted, based on simulated data sets. The detailed formulation of L_2E estimator for FGM will be given later in this paper.

Let K be the number of the Gaussian components. Fig. 2.(a) shows an experiment for a single-mode Gaussian ($K = 1$) case. The input data is made of a single Gaussian plus an outlier portion. The EM estimates, as shown in Fig. 2.(a), are greatly deviated from the true values, while L_2E captures the single Gaussian part very well. Fig. 2.(b) presents the results of the second experiment, in which the outlier portion has bulky overlap with the inlier part. L_2E maintains the ability to capture the major component, without being affected by the outlier.

Fig. 2.(c) shows the result for a two-Gaussian mixture ($K = 2$) case. The data is composed of two Gaussians

and one outlier portion. The result using EM demonstrates the impact of outliers on the global parameter estimation – none of the two Gaussians is estimated correctly by the EM. L_2E, to the contrary, successfully gets hold of the two major Gaussians with great accuracy. A similar experiment is shown in Fig. 2.(d), where the outlier portion is located at a higher position of the intensity spectrum. This time, the outliers, instead of Gaussian components, got captured by the EM algorithm. Therefore if an algorithm is based on EM, the classification results would likely be far away from the true values, if this same data set is applied. L_2E still works perfectly in rejecting the outliers and producing the desired results.

For the examples showed above, if extra classes are assigned, EM algorithm can manage to capture both the inliers and the outliers. However, to separate the inlier portion off the outliers usually requires some prior knowledge or certain post-classification user intervention, which certainly demands extra efforts in practice.

Comparing with the M-estimators, L_2E differs in the sense that it is not formulated with any scale parameter. This property can again be counted as an advantage of L_2E, as the success of the M-estimators is heavily dependent on the often elusive setup of the scale parameter. Another proof of the superiority of L_2E comes from ([21],[15]). The authors compared L_2E with 15 other robust estimators, and L_2E often came out on top, particularly with the outlier-abundant, heavy-tailed data sets.

2.1.3 MS Lesion Detection using L2E ($MSLD - L_2E$)

The remaining problem is how to separate the clusters of normal tissues (GM and WM) from the outliers (MS lesions). This is achieved through a 3-step procedure:

Step 1: Based on the one-dimensional histogram of a combined T2w + FLAIR image, model the lesion part with an independent class and make a preliminary separation between lesions and the normal tissues.

Step 2: Use multivariate Gaussian fitting to capture the GM and WM portion in T2w/FLAIR two-dimensional joint histogram.

Step 3: Conduct automatic outlier (MS-lesions) detection based on the Gaussian components from step 2 and the separation line obtained in step 1.

In step 1, aiming to combine the information from T2w and FLAIR, we adopt an integrated image as the segmentation basis, whose intensity is given by

$$I_{comb} = \sqrt{I_{T2_w}^2 + I_{FLAIR}^2} \quad (3)$$

Since the histogram of the I_{comb} is one dimensional, it's easy to fit it with a 1D finite Gaussian mixture of three classes: GM, WM and MS-lesions. Let $\theta = \{v_1, \mu_1, \sigma_1, v_2, \mu_2, \sigma_2, v_3, \mu_3, \sigma_3\}$ be the combined vector representing the portions, means and standard deviations of the three Gaussian components. The optimal θ is obtained to minimize the L_2E metric given by

$$L_2E_{1d}(\theta) = \sum_{k=1}^{3} \frac{v_k^2}{2\sqrt{\pi}\sigma_k^2} - \frac{2}{n}\sum_{k=1}^{3}\sum_{i=1}^{n}\left[v_k\phi(x_i|\mu_k, \sigma_k^2)\right] \quad (4)$$

where n is the number of voxels.

An example of the 3-step lesion detection procedure is shown in Figure 3. Fig. 3(a), 3(b) and 3(c) are 2D slices of T2w, FLAIR and the combined image (with background and CSF removed), respectively. The histogram of the combined image appears as a single bump with a long tail. The long tail corresponds to an agglomeration of the lesion voxels, while the single bump is actually the summation of two highly overlapping Gaussians, one for WM, and other for the cortical GM. Our L_2E-based FGM algorithm can accurately capture the 3 components, especially the lesion part away from the other two components. Fig. 3(e) shows the fitting result. Black curve is the constructed histogram of I_{comb} ; green line shows the overall fitting, with three individual components plotted in red. The separation line is set as the intersection of the rightmost two components (cortical GM and the lesions) in the histogram. Fig. 3(f) is the zoom-in version of 3(e).

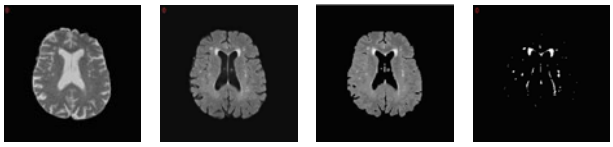
The one-dimensional image I_{comb} is convenient to use, however, it doesn't contain as much information as the two-dimensional joint T2w/FLAIR histogram. In step 2, we use the latter to better capture and describe the GM and WM portions. This time, however, the lesions part no longer conform to a Gaussian distribution, so we only focus on GM and WM, and the number of classes is reduced to 2. Let $\phi(\boldsymbol{x}|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ denotes the k-th ($k = 1, 2$) component multivariate normal density, the parametric distribution assumed by FGMs is as follows: $y(x|\boldsymbol{\theta}) = \sum_{k=1}^{K} w_k\phi(\boldsymbol{x}|\boldsymbol{\mu}_k, \Sigma_k)$ where $\boldsymbol{\theta} = \{w_1, \boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1, w_2, \boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2\}$ represents the portions, means, and covariance matrices of the two Gaussian components. The optimal $\boldsymbol{\theta}$ is obtained to minimize the L_2E metric given by

$$L_2E_{2d}(\theta) = \sum_{k=1}^{2}\sum_{l=1}^{2} w_k w_l \phi(\mathbf{0}|\boldsymbol{\mu}_k - \boldsymbol{\mu}_l, \boldsymbol{\Sigma}_k + \boldsymbol{\Sigma}_l) - \frac{2}{n}\sum_{i=1}^{n}[w_1\phi(x|\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1) + w_2\phi(x|\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)] \quad (5)$$

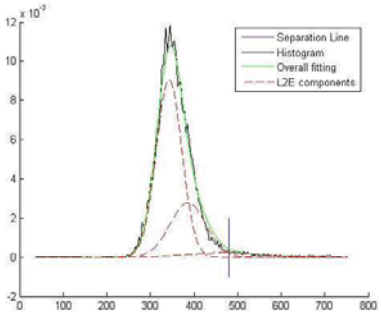
where $w_1 + w_2 = 1$.

Figure 4 is the FGM fitting result using the 2D L_2E metric in eqn. (5). All the voxel pairs whose Mahalanobis dis-

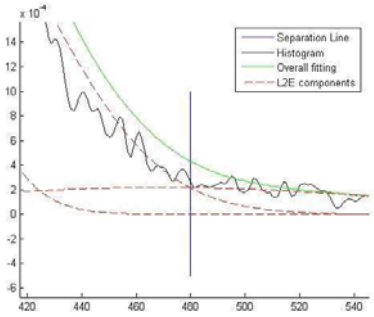
tance exceeds certain threshold will be counted as preliminary lesions/outliers, where the threshold is determined in step 3.



(a) (b) (c) (d)



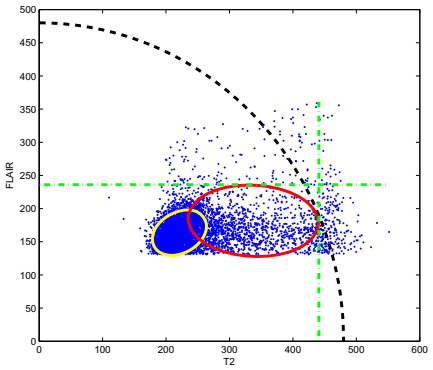
(e)



(f)

Figure 3: MS lesion detection using our $MSLD - L_2E$ method. The input FLAIR (a), T2w (b), the combined image (c) after background and CSF removal. The segmentation result is shown in (d). (e) is an illustration of the fitting result using 3-class L_2E. (f) is a zoom-in version of (e). The multi-spectral segmentation on FLAIR and T2w is given in Figure 4. Pictures are better seen on screen than in black/white print.

It should be noted that, since $I_{comb} = \sqrt{I_{T2_w}^2 + I_{FLAIR}^2}$, the separation threshold obtained in step 1 corresponds to a quarter circle in the 2D T2w/FLAIR joint histogram (the black dashed line shown in Fig. 4). A reasonable Mahalanobis threshold can therefore be set based on the contact point of the quarter circle and the closest Gaussian ellipse. This point can be computed analytically or through an iterative procedure. In our implementation, we gradually increase the σ of the



(a)

Figure 4: Automatic threshold detection based on the multi-spectral segmentation of FLAIR and T2w.

two Gaussian components. When any of the Gaussian ellipses start to inscribe the quarter circle, we stop and the returned σ is chosen as the Mahalanobis threshold. The voxel intensities beyond the extreme points/lines are counted into the lesion class. The extreme lines are marked with green color in Fig. 4. The resulting segmentation is given in Fig. 3(d).

2.1.4 Numerical Solution

To minimize the energy in Eqn.(4) and Eqn.(5), the constrained minimization routine *fmincon* in Matlab Optimization Toolbox is employed. The underlying numerical method is sequential quadratic programming. The objective function of L_2E (Eqns (4) and (5)) is implemented in a Matlab .m file, and sent as an input to *fmincon*.

2.2 Refinement step to prune false positives

In addition to lesions, the segmentation results obtained from $MSLD - L2E$ may contain other hyper-intense signals like bony artifacts and flow artifacts. The latter are generated by pulsatile CSF flow that causes incomplete nulling of CSF signal intensities. This type of artifacts are mainly located around the interface of CSF and cortical GM, so a dilated CSF mask could potentially erase them up. However, periventricular areas have to be spared from the erasion, because the lesions surrounding the ventricles are most clearly depicted in FLAIR images and have very little chance to be flow artifacts [14].

To eliminate the extra-cortical CSF artifacts, a region of interest needs to be defined. Our solution is to develop two expanded masks, corresponding to extra-cortical and periventricular areas respectively, based on the available FLAIR CSF segmentation.Two morphological operations,

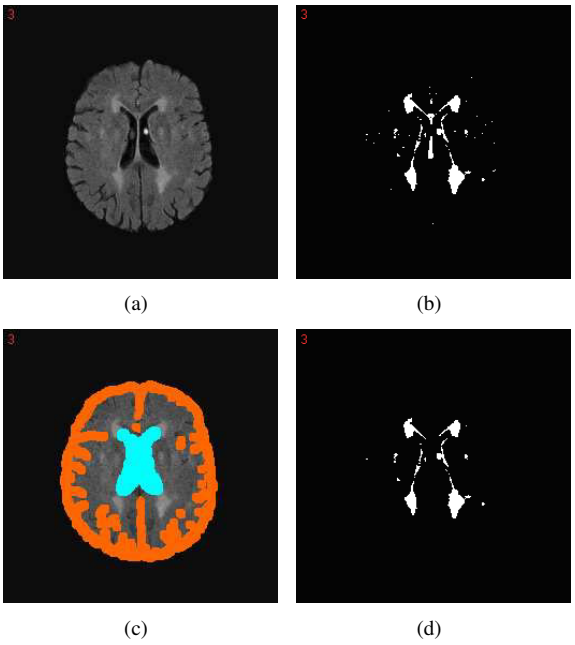


Figure 5: Refinement to prune false positives. (a) FLAIR image; (b) segmentation results before artifact pruning; (c) extra-cortical (orange) and peri-ventricular (light blue) CSF masks; (d) final lesion segmentation.

erosion and dilation, are applied. Erosion operation shrinks an object by one voxel. The separation of ventricular CSF from other previously connected portions can be achieved by repeatedly applying the erosion operation on the binary CSF segmentation. A series of dilation operations on the disconnected extra-cortical CSF and ventricular CSF can bring them back to their original size and beyond.

Figure 5 depicts an example of the artifact reduction procedure. Fig. 5(a) is a 2D FLAIR slice. 5(b) is the segmentation result from our $MSLD - L_2E$ procedure. The orange area in 5(c) shows the expanded extra-cortical CSF mask, where detected "lesions" will be regarded as false positives and removed. Blue mask delineates the periventricular areas where the $MSLD - L_2E$ responses should be preserved.

In some data sets, third ventricle area also tends to have extra-bright FLAIR artifacts. To remove these artifacts, we identify the brain middle plane based the method proposed in MindBoggle [10]. A deformation registration [18] procedure is applied to capture the third ventricle, and false-positives lesions falling in this area are then deleted. Finally, lesions in non-ventricular area those size is smaller than 4 voxels are regarded as noise responses and removed. Fig. 5(d) shows the final segmentation of the lesions after all the artifacts removal steps.

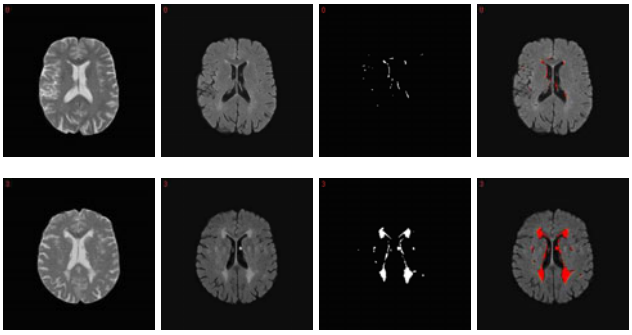


Figure 6: Classification of patients with small (first row) and moderate (second row) lesion loads. From left to right: T2w, FLAIR, the MS lesion segmentation result and superimposed on the FLAIR image.

3 Experimental Results

The experiment we conducted is based on 16 set of T2 and T2w-FLAIR data sets obtained from the University of Kentucky Hospital. The data are 0.94mm x 0.94mm x 3mm in spacing and 256 x 256 x 41 in volume size. Based on the lesion size estimated by eye, we divided the 16 data into 2 groups: patients with small lesions load, and with moderate lesion load. In Figure 6, example images are shown of the classification results from these two groups. For each patient category, the following images are shown: T2w, FLAIR, segmentation from our $MSLD - L_2E$ model, and the FLAIR image with the segmentation superimposed

3.1 Evaluation measures

To assess the accuracy and consistency of the results, a comparison between the automatic segmentations and manual segmentations from an expert neuro-scientist has been performed. The performance metrics used are: *Volume Difference, Dice Coefficient, Sensitivity* and *Specificity*.

The volume difference captures the absolute percent volume difference to the expert manual segmentation. Dice coefficient measures the similarity of two sets and ranges from 0 for sets that are disjoint to 1 for sets are identical. It is defined as:

$$K(S_1, S_2) = \frac{2 \times |S_1 \cap S_2|}{|S_1| + |S_2|} \tag{6}$$

Segmentation sensitivity of certain tissue type A is the probability of the voxels in A (in the ground truth) being correctly classified by the estimated segmentation. Specificity, on the other hand, is the probability that a voxel u(X) is not labeled as A given u(X) doesn't belong to A.

Table 1 shows the MS lesion segmentation results from our L_2E-based scheme. The average Dice coefficients

(similarity index) for the patients with small lesion load and moderate lesion load are 0.6336 and 0.8405 respectively, which are higher than the reported numbers in [3] and [12]. Although no direct comparison with other works can be made at this point due to the different testing data used, the high similarity indexes obtained for both subgroups are good indications of the accuracy and robustness of our solution.

Data Set	Vol. Diff [%]	Dice Coeff.	Sensi.	Specif.	Lesion Size
1	20.3600	0.6936	0.6998	0.9998	S
2	1.5700	0.7174	0.7231	0.9997	S
3	32.5200	0.6323	0.5394	0.9998	S
4	4.2400	0.9281	0.9084	0.9998	M
5	25.7000	0.7023	0.7029	0.9999	S
6	30.0000	0.4938	0.4196	0.9997	S
7	6.6200	0.3400	0.3417	0.9999	S
8	5.5800	0.6861	0.6880	0.9997	M
9	22.3600	0.6236	0.7284	0.9997	S
10	22.9700	0.7327	0.6484	0.9991	S
11	9.0160	0.8361	0.8398	0.9998	M
12	11.7800	0.8474	0.7974	0.9998	M
13	28.6500	0.6743	0.5777	0.9998	S
14	18.9600	0.8606	0.7790	0.9999	M
15	63.5200	0.5304	0.3619	0.9999	S
16	6.3800	0.8844	0.9126	0.9997	M
Ave. (S)	23.56	0.6336	0.6003	0.9998	
STD. (S)	17.9645	0.1564	0.1848	0.0001	
Ave. (M)	9.32	0.8405	0.8209	0.9998	
STD. (M)	5.4316	0.0824	0.0853	0.0000	
Ave. (All)	19.38	0.6989	0.6667	0.9997	
STD. (All)	15.52	0.1561	0.1774	0.0001	

Table 1: MS lesion segmentation results from our L_2E scheme. "S" stands for *Small*, and "M" for *Moderate*.

Since our automatic segmentation procedure scrutinizes an image more thoroughly than a human rater, and inevitably there are artifacts that can not be completely eliminated through the refinement step, it shouldn't be a surprise that the lesion result produced by $MSLD - L_2E$ is close to a superset of the manual segmentation. This is reflected in the high specificity and relative low sensitivity scores list in Table 1.

The obtained Dice Coefficients and Sensitivity also suggest that the proposed method produces better results for patients with a moderate lesion load than for patients with a small lesion load. This can be explained by the fact that small errors have a relatively larger effect on a smaller reference area. We expect the values to be increased when our method is applied to the patients with larger lesion load.

3.2 Program running time

The algorithm presented in this paper is implemented with Matlab 7 R14, under an IBM Thinkpad T60 laptop (Intel Core Duo T2300 / 1.66 GHz processor, 2.0GB of RAM). The average running time for Kentucky data is around 2 minutes per set. The L_2E parameters update routine is implemented purely using Matlab m-files so far. We expect a significant reduce in the whole running time after the core level set part is implemented with mex-C.

4 Discussion and future work

We have described a fully automatic algorithm for MS-lesion detection that requires virtually no user interaction. The segmentation models lesion voxels in an additional class to the mixture of the normal brain tissues, CSF, GM and WM. Neither training nor thresholding is needed in our model.

It should be noted that most of the methods for MS lesion segmentation described in the literature are semi-automated rather than fully automated methods. Some of efforts are to facilitate the tedious task of manually delineating by human raters, and to reduce the associated intra and inter-expert variabilities. Other models, though rely on certain automatic classifiers to make final labeling, often involve an interactive process during the training stage. While these methods have proven to be useful, they remain inconvenient or sometimes impractical when massive amount of scans need to be analyzed. In contrast, our method is fully automated and the only user interaction needed is feeding data to our system. As the algorithm requires no training, it has the potential to be directly applied to images generated using different MR scanners or different scanning parameters. Extension to other lesion characteristics or different MR scanning sequences can be done in a fairly straightforward way.

Another advantage of our approach lies in the fact that it involves no thresholding step. Most outlier-based methods avoid explicit lesions modeling because MS lesions tend to vary greatly in size, shape and location, and some of them don't even contains a sufficient number of voxels for estimating the model parameters easily. However, without explicit modeling, either soft or hard rejection, a predetermined threshold has to be used to decide the separation line/plane between the normal tissue and the outlier voxels. Since the thresholds are often data-dependent, manually chosen values tend to not work consistently across different data sets. Our method overcomes this problem, thanks to the strong capture capability of L_2E estimation, and achieves great flexibility and broad applicability.

References

[1] F. Admiraal-Behloul, D.M.J. van den Heuvel, H. Olofsen M.J.P. van Osch, J. van der Grond, M.A. van Buchem, and J.H.C. Reiber, "Fully automatic segmentation of white matter hyperintensities in MR images of the elderly", NeuroImage 28 (2005) 607 - 617.

[2] L. S. Ait-Ali, S. Prima, P. Hellier, B. Carsin, G. Edan, and C. Barillot, "STERM: A Robust Multidimensional Parametric Method to Segment MS Lesions in MRI", MICCAI 2005, pp. 409-416, 2005.

[3] P. Anbeek, K. L. Vinchen, M. J. P. van Osch, R. H.C. Bisschops and J. van der Grond, "Probabilistic segmentation of white matter lesions in MR imaging", NeuroImage, 21 (2004) 1037 - 1044.

[4] P. Anbeek, K. L. Vinchen and M. A. Viergever, "Antomated MS-Lesion Segmentation by K-Nearest Neighbor Classification", 3D Segmentation in the Clinic: A Grand Challenge II: MS lesion segmentation. June, 2008.

[5] P. L. Bazin and D.L. Pham. Statistical and topological atlas based brain image segmentation. MICCAI, 2007.

[6] G. Dugas-Phocion, M. A. Gonzalez, C. Lebrun, S. Chanalet, C. Bensa, G. Malandain, N. Ayache, "Hierarchical Segmentation of Multiple Sclerosis Lesions in Multi-Sequence MRI", ISBI'08,

[7] D. Garcia-Lorenzo, S. Prima, S. P. Morrissey and C. Barillot, "A Robust Expectation-Maximization Algorithm for Multiple Sclerosis Lesion Segmentation", 3D Segmentation in the Clinic: A Grand Challenge II: MS lesion segmentation. June, 2008.

[8] K. Van Leemput, F. Maes, D. Vandermeulen, A. Colchester and P. Suetens, "Automated Segmentation of Multiple Sclerosis Lesions by Model Outlier Detection", IEEE TMI, vol. 20, No. 8, August 2001, pp. 677 - 689.

[9] J. Liu, B. C. Vemuri, J. L. Marroquin, "Local Frequency Representations for Robust Multimodal Image Registration". IEEE TMI, Vol. 21, No. 5, pp. 462-469, May 2002.

[10] A. Klein and J. Hirsch. "Mindboggle: a scatterbrained approach to automate brain labeling". NeuroImage. 24(2): 261-280. 2005.

[11] J. Morra, Z. Tu, A. Toga, P. Thompson. "Automatic Segmentation of MS Lesions Using a Contextual Model for the MICCAI Grand Challenge", 3D Segmentation in the Clinic: A Grand Challenge II: MS lesion segmentation. June, 2008.

[12] M. Prastawa and Guido Gerig, "Automatic MS Lesion Segmentation by Outlier Detection and Information Theoretic Region Partitioning", 3D Segmentation in the Clinic: A Grand Challenge II: MS lesion segmentation. June, 2008.

[13] D. Rey, G. Subsol, H. Delingette, and N. Ayache. "Automatic detection and segmentation of evolving processes in 3D medical images: Application to multiple sclerosis". Medical Image Analysis, 6(2):163-179, June 2002.

[14] M. A. Sahraian and W. W. Radue, "MS Lesions in Fluid Attenuated Inversion Recovery Images", MRI Atlas of MS Lesions, pp. 35-44. Springer, 2007.

[15] D. W. Scott, "Parametric Statistical Modeling by Minimum Integrated Square Error", Technometrics, vol. 43(3), 2001.

[16] N. Shiee, P. Bazin and D. Pham, "Multiple Sclerosis Lesion Segmentation Using Statistical and Topological Atlases", 3D Segmentation in the Clinic: A Grand Challenge II: MS lesion segmentation. June, 2008.

[17] J. C. Souplet, C. Lebrun, N. Anyche, G. Malandain, "An Automatic Segmentation of T2-FLAIR Multiple Sclerosis Lesions", 3D Segmentation in the Clinic: A Grand Challenge II: MS lesion segmentation. June, 2008.

[18] J. P. Thirion and G. Calmon. Deformation analysis to detect and quantify active lesions in threedimensional medical image sequences. IEEE TMI, 18(5):429-441, 1999.

[19] Z. Tu, K. Narr, I. Dinov, P. Dollar, P. Thompson, and A. Toga. "Brain anatomical structure parsing by hybrid discriminative/generative models". IEEE TMI, 2008.

[20] B. Jian, B. C. Vemuri, "A Robust Algorithm for Point Set Registration Using Mixture of Gaussians". ICCV 2005: 1246-1251.

[21] W. Wojciechowski, "Robust Modeling", doctoral dissertation, Rice University, Houston, 2001

[22] Y. Wu, S. K. Warfield, I. L. Tan, W. M. Wessl III, D. S. Meier, R. A. Van Schijndel, F. Barkhof and C. Guttmann, "Automated segmentation of multiple sclerosis lesion subtype with multichannel MRI", NeuroImage 32 (2006) 1025 - 1215.

[23] A. Zijdenbos, R. Forghani, and A. Evans, "Automatic Quantification of MS lesions in 3D MRI Brain Data Sets: Validation of INSECT", MICCAI'98, pp. 439-448, 1998.