POPULATION INTENSITY OUTLIERS OR A NEW MODEL FOR BRAIN WM ABNORMALITIES

Tomas-Fernandez, X, Warfield, S.K.

Computational Radiology Laboratory - Childrens' Hospital Boston

ABSTRACT

We present a new automatic method for segmentation of Multiple Sclerosis (MS) lesions in Magnetic Resonance Images. The algorithm performs tissue classification combining a within subject global tissue intensity model and a local tissue intensity model derived from an aligned set of healthy reference subjects. MS lesions are detected as outliers towards the proposed coupled global/local intensity model. Evaluation using BrainWeb synthetic, show our new coupled local/global intensity GMM model to be sensitive towards MS lesions, as well to be robust to noise and intensity inhomogeneity artifacts found MRI scans.

Index Terms— Magnetic Resonance Imaging, Multiple Sclerosis, Segmentation, Bayesian

1. INTRODUCTION

In Van Leemput et al. [1], MS lesions were proposed to be modeled as outliers from a healthy tissue model. The algorithm performed intensity based tissue classification using a stochastic model for normal brain images and detected MS lesions as outliers that were not well explained by the model. Similarly, [2] used a Trimmed Likelihood Estimator (TLE) to estimate a 10 component Gaussian Mixture Model (GMM) and segmented MS lesions as GM intensity outliers on an enhanced Fluid-Attenuated Inversion Recovery (FLAIR) sequence. Additional methods further combine a TLE with a mean shift algorithm [3]. These algorithms detect lesions as intensity outliers from the GMM of all the normal brain tissues.

Intensity based classification relies on contrast between tissue types in feature space and adequate signal compared to image noise. Statistical classification identifies an optimal boundary, in feature space, between tissue types, and the separability of two tissue classes is related to the overlap between classes in feature space. MS lesions MRI intensity feature space overlap with those of brain healthy tissues making difficult to identify a decision boundary, resulting in increased tissue misclassification and multiple MS lesions false positives. In Tomas-Fernandez et al. [4] an extended feature space, consisting in MRI intensities and a voxelwise intensity similarity metric between the MS patient and a reference healthy population, provided a more accurate and specific MS lesion segmentation.

In this paper, we present a novel algorithm for brain tissue segmentation and WM lesions detection in MS. We propose a novel tissue model which combines a patient global intensity model with a population local intensity model derived from a reference of healthy subjects. We evaluated the new combined model using the BrainWeb synthetic database, demonstrating a brain tissue (GM, WM and CSF) segmentation robust to the presence of brain abnormalities as well as a superior MS lesion sensitivity compared to current state-of-the-art algorithms.

2. METHODS

2.1. Local Reference Population Bayesian Intensity Tissue Model

Consider a reference population $\mathbf{P}$ formed by R healthy subjects aligned to the subject of interest. Each reference subject is composed by a multispectral grayscale MRI $\mathbf{V}$ (i.e. T1w, T2w and FLAIR) and the corresponding tissue segmentation $\mathbf{L}$ (i.e. GM, WM and CSF), thus $\mathbf{P} = \{\mathbf{V}, \mathbf{L}\} = \{\mathbf{V}_1, ..., \mathbf{V}_R; \mathbf{L}_1, ..., \mathbf{L}_R\}$. Where each reference grayscale MRI $\mathbf{V_r} = \{\mathbf{V}_{r1}, ..., \mathbf{V}_{rN}\}$ is formed by a finite set of N voxels with $\mathbf{V}_{ri} \epsilon \mathbb{R}^m$. Also each reference tissue segmentation $\mathbf{L_r} = \{\mathbf{L}_{r1}, ..., \mathbf{L}_{rN}\}$ is formed by a finite set of N voxels with $\mathbf{L}_{ri} = \mathbf{e}_k = \{l_{ri1}, ..., l_{riK}\}$ is a K-dimensional vector with each component l_{rik} being 1 or 0 according whether $\mathbf{v}_{ri}$ did or did not arise from the k^{th} class.

At each voxel i, the reference population $\mathbf{P}$ intensity distribution will be modeled as a Gaussian Mixture Model parametrized by $\boldsymbol{\xi}_i = \{\boldsymbol{\pi}_{\mathbf{P}i}; \boldsymbol{\mu}_{\mathbf{P}i}, \boldsymbol{\Sigma}_{\mathbf{P}i}\}$. Where $\boldsymbol{\pi}_{\mathbf{P}i}$, $\boldsymbol{\mu}_{\mathbf{P}i}$ and $\boldsymbol{\Sigma}_{\mathbf{P}i}$ are respectively the population tissue mixture vector, the population mean intensity vector and the population intensity covariance matrix at voxel i. Because $\{\mathbf{V}, \mathbf{L}\}$ are observed variables, $\boldsymbol{\xi}_i$ can be derived using the following expressions:

$$\pi_{Pik} = \frac{1}{R} \sum_{j=1}^{R} p(L_{ij} = e_k) \quad (1)$$

This investigation was supported in part by NIH grants R01 RR021885, R01 EB008015, R03 EB008680 and R01 LM010033.

$$\boldsymbol{\mu}_{Pik} = \frac{\sum_{j=1}^{R} \mathbf{V}_{ij} p(L_{ij} = e_k)}{\sum_{j=1}^{R} p(L_{ij} = e_k)} \quad (2)$$

$$\boldsymbol{\Sigma}_{Pik} = \frac{\sum_{j=1}^{R} (\mathbf{V}_{ij} - \boldsymbol{\mu}_{Pik})^T (\mathbf{V}_{ij} - \boldsymbol{\mu}_{Pik}) p(L_{ij} = e_k)}{\sum_{j=1}^{R} p(L_{ij} = e_k)} \quad (3)$$

where $p(L_{ij} = e_k)$ is the probability of voxel i of the j_{th} reference subject belong to tissue k given by $\mathbf{L}_j$.

Once the local tissue model is estimated from $\mathbf{P}$, the complete likelihood of $\mathbf{Y}$ can be estimated as:

$$f(\mathbf{Y}, \mathbf{Z}|\boldsymbol{\xi}) = \prod_{i=0}^{N} f(\mathbf{Y}_i, \mathbf{Z}_i|\boldsymbol{\xi}_i) = \prod_{i=1}^{N} f(\mathbf{Z}_i = \mathbf{e}_k|\boldsymbol{\xi}_i) f(\mathbf{Y}_i|\mathbf{Z}_i = e_k, \boldsymbol{\xi}_i)$$
$$= \prod_{i=1}^{N} \sum_{k=1}^{K} \frac{f(\mathbf{Z}_i = \mathbf{e}_k|\boldsymbol{\xi}_{ik})}{(2\pi)^{m/2} |\boldsymbol{\Sigma}_{Pik}|^{1/2}} e^{-\frac{1}{2}(\mathbf{Y}_i - \boldsymbol{\mu}_{Pik})^T \boldsymbol{\Sigma}_{Pik}^{-1} (\mathbf{Y}_i - \boldsymbol{\mu}_{Pik})} \quad (4)$$

with $f(\mathbf{Z}_i = \mathbf{e}_k|\boldsymbol{\xi}_{ik}) = \pi_{Pik}$.

2.2. Coupling Global and Local Models

Consider a multispectral grayscale MRI (i.e. T1w, T2w and FLAIR) formed by a finite set of N voxels, our aim is to assign each voxel i to one of K classes (i.e. GM, WM and CSF) considering the observed intensities $\mathbf{Y} = \{\mathbf{y}_1, ..., \mathbf{y}_N\}$ with $\mathbf{y}_i \in \mathbb{R}^m$. Both observed intensities and hidden labels are considered to be random variables denoted respectively as $\mathbf{Y} = \{\mathbf{Y}_1, ..., \mathbf{Y}_N\}$ and $\mathbf{Z} = \{\mathbf{Z}_1, ..., \mathbf{Z}_N\}$.

The observed data $\mathbf{Y}$ and hidden labels $\mathbf{Z}$ are described by a parametric conditional probability density function defined by $\{\boldsymbol{\psi}, \boldsymbol{\xi}\}$, where $\boldsymbol{\psi}$ parametrize a within subject global GMM and $\boldsymbol{\xi}$ parametrize a local GMM derived from an aligned reference of R healthy subjects $\mathbf{P}$.

Segmenting the observed image $\mathbf{Y}$ is to propose an estimate $\hat{\mathbf{Z}}$ of $\mathbf{Z}$ on the basis of $\mathbf{Y}$, to this purpose, the parameters $\boldsymbol{\psi}$ and $\boldsymbol{\xi}$ need to be estimated somehow. If the underlying tissue segmentation $\mathbf{Z}$ was known, estimation of the model parameters would be straightforward. However, only the image $\mathbf{Y}$ is directly observed, making natural to tackle this problem as one involving missing data making the Expectation-Maximization (EM) algorithm the candidate for model fitting. The EM algorithm finds the parameter $\boldsymbol{\psi}$ that maximize the complete log-likelihood of the data by iteratively maximizing the expected value of the log-likelihood $log(f(\mathbf{Y}, \mathbf{Z}|\boldsymbol{\psi}, \boldsymbol{\xi}))$ of the complete data $\{\mathbf{Y}, \mathbf{Z}\}$, where the expectation is based on the observed data $\mathbf{Y}$ and the estimated parameters $\boldsymbol{\psi}^{(m)}$ and $\boldsymbol{\xi}^{(m)}$obtained in the previous iteration m.

$$log\, L_C(\boldsymbol{\psi}, \boldsymbol{\xi}) = log\,(f(\mathbf{Y}, \mathbf{Z}|\boldsymbol{\psi}, \boldsymbol{\xi}))$$
$$= log\left(\prod_{i=1}^{N} \sum_{k=1}^{K} f(\mathbf{Z}_i = e_k|\boldsymbol{\psi}_k, \boldsymbol{\xi}_{ik}) f(\mathbf{Y}_i|\mathbf{Z}_i = e_k, \boldsymbol{\psi}_k, \boldsymbol{\xi}_{ik})\right)$$

The local intensity model is just dependent on the aligned reference population $\mathbf{P} = \{\mathbf{V}, \mathbf{L}\}$ which is observable. Thus, parameter $\boldsymbol{\xi}$ is constant and independent from the global intensity model parametrized by $\boldsymbol{\psi}$. We can rewrite the complete log-likelihood as:

$$log\, L_C(\boldsymbol{\psi}, \boldsymbol{\xi}) = \log\,(f(\mathbf{Y}, \mathbf{Z}|\boldsymbol{\psi}, \boldsymbol{\xi})) = \log\,(f(\mathbf{Y}, \mathbf{Z}|\boldsymbol{\psi}) f(\mathbf{Y}, \mathbf{Z}|\boldsymbol{\xi}))$$
$$= \sum_{i=1}^{N} \log\left(\sum_{k=1}^{K} f(\mathbf{Z}_i = \mathbf{e}_k|\boldsymbol{\psi}_k) f(\mathbf{Y}_i|\mathbf{Z}_i = \mathbf{e}_k, \boldsymbol{\psi}_k)\right)$$
$$+ \sum_{i=1}^{N} \log\left(\sum_{k=1}^{K} f(\mathbf{Z}_i = \mathbf{e}_k|\boldsymbol{\xi}_{ik}) f(\mathbf{Y}_i|\mathbf{Z}_i = e_k, \boldsymbol{\xi}_{ik})\right)$$

where $\mathbf{e}_k$ is a K-dimensional vector with each component being 1 or 0 according whether $\mathbf{Y}_i$ did or did not arise from the k^{th} class. Thus:

$$log\, L_C(\boldsymbol{\psi}, \boldsymbol{\xi}) = \sum_{i=1}^{N} \sum_{k=1}^{K} z_{ij} \left(\log\left(f(\mathbf{Z}_i = \mathbf{e}_k|\boldsymbol{\psi}_k) + f(\mathbf{Y}_i|\mathbf{Z}_i = \mathbf{e}_k, \boldsymbol{\psi}_k)\right)\right)$$
$$+ \sum_{i=1}^{N} \sum_{k=1}^{K} z_{ij} \left(\log\left(f(\mathbf{Z}_i = \mathbf{e}_k|\boldsymbol{\xi}_{ik}) + f(\mathbf{Y}_i|\mathbf{Z}_i = \mathbf{e}_k, \boldsymbol{\xi}_{ik})\right)\right) \quad (5)$$

E-Step: Because $\boldsymbol{\xi}$ is constant, the E-step of the coupled model requires the computation of the conditional expectation of $\log\,(L_C(\boldsymbol{\psi}, \boldsymbol{\xi}))$ given $\mathbf{Y}$, using $\boldsymbol{\psi}^{(m)}$ for $\boldsymbol{\psi}$.

$$Q(\boldsymbol{\psi}, \boldsymbol{\xi}; \boldsymbol{\psi}^{(m)}, \boldsymbol{\xi}^{(m)}) \sim Q(\boldsymbol{\psi}; \boldsymbol{\psi}^{(m)}) = E_{\boldsymbol{\psi}^{(\mathbf{m})}}\{\log\, L_C(\boldsymbol{\psi}, \boldsymbol{\xi})|\mathbf{Y}\}$$

The complete log-likelihood is linear in the hidden labels z_{ij}. The E-step requires the calculation of the current conditional expectation of $\mathbf{Z}_i$ given the observed data $\mathbf{Y}$:

$$E_{\boldsymbol{\psi}^{(\mathbf{m})}}(\mathbf{Z}_i = \mathbf{e}_k|\mathbf{Y}, \boldsymbol{\xi}_{ij}) = f(\mathbf{Z}_i = \mathbf{e}_j|\mathbf{Y}_i, \boldsymbol{\xi}_{ij}, \boldsymbol{\psi}_j^{(m)})$$
$$= \frac{f(\mathbf{Y}_i|\mathbf{Z}_i = \mathbf{e}_j, \boldsymbol{\xi}_{ij}, \boldsymbol{\psi}_j^{(m)}) f(\mathbf{Z}_i = \mathbf{e}_j|\boldsymbol{\xi}_{ij}, \boldsymbol{\psi}_j^{(m)})}{\sum_{k=1}^{K} f(\mathbf{Y}_i|\mathbf{Z}_i = \mathbf{e}_k, \boldsymbol{\xi}_{ik}, \boldsymbol{\psi}_k^{(m)}) f(\mathbf{Z}_i = \mathbf{e}_k|\boldsymbol{\xi}_{ik}, \boldsymbol{\psi}_k^{(m)})} \quad (6)$$

Parameter $\boldsymbol{\psi}^{(m)}$ and $\boldsymbol{\xi}$ are independent, thus:

$$f(\mathbf{Y}_i|\mathbf{Z}_i = \mathbf{e}_j, \boldsymbol{\xi}_{ij}, \boldsymbol{\psi}_j^{(m)}) \sim f(\mathbf{Y}_i|\mathbf{Z}_i = \mathbf{e}_j, \boldsymbol{\psi}_j^{(m)}) f(\mathbf{Y}_i|\mathbf{Z}_i = \mathbf{e}_j, \boldsymbol{\xi}_{ij})$$
$$f(\mathbf{Z}_i = \mathbf{e}_j|\boldsymbol{\xi}_{ij}, \boldsymbol{\psi}_j^{(m)}) \sim f(\mathbf{Z}_i = \mathbf{e}_j|\boldsymbol{\psi}_j^{(m)}) f(\mathbf{Z}_i = \mathbf{e}_j|\boldsymbol{\xi}_{ij})$$

M-Step: Because $\boldsymbol{\xi}$ is constant, the Maximization step will consist in the maximization of $Q(\boldsymbol{\psi}, \boldsymbol{\xi}; \boldsymbol{\psi}^{(m)}, \boldsymbol{\xi}^{(m)})$ with respect to $\boldsymbol{\psi}$, which results in the following update equations:

$$f(\mathbf{Z}_i = \mathbf{e}_k|\boldsymbol{\psi}_k^{(m+1)}) = \frac{1}{N} \sum_{i=1}^{N} f(\mathbf{Z}_i = \mathbf{e}_k|\mathbf{Y}_i, \boldsymbol{\psi}^{(m)}, \boldsymbol{\xi}_{ik}) \quad (7)$$

$$\boldsymbol{\mu}_k^{(m+1)} = \frac{\sum_{i=1}^{N} \mathbf{Y}_i f(\mathbf{Z}_i = \mathbf{e}_k|\mathbf{Y}, \boldsymbol{\psi}^{(m)})}{\sum_{i=1}^{N} f(\mathbf{Z}_i = \mathbf{e}_k|\mathbf{Y}, \boldsymbol{\psi}^{(m)})} \quad (8)$$

$$\boldsymbol{\Sigma}_k^{(m+1)} = \frac{\sum_{i=1}^{N} f(\mathbf{Z}_i = \mathbf{e}_k|\mathbf{Y}, \boldsymbol{\psi}^{(m)})(\mathbf{Y}_i - \boldsymbol{\mu}_k^{(m)})^T (\mathbf{Y}_i - \boldsymbol{\mu}_k^{(m)})}{\sum_{i=1}^{N} f(\mathbf{Z}_i = \mathbf{e}_k|\mathbf{Y}, \boldsymbol{\psi}^{(m)})} \quad (9)$$

2.3. MS Lesion Detection

Intuitively from (5), the local intensity model downweights the likelihood of those voxels with an abnormal intensity given the reference population. Because MS lesions show

an abnormal intensity level compared to a healthy subject in the same location, we assume that MS lesions correspond with brain areas with low likelihood, making feasible MS lesions detection as outliers toward our coupled global/local intensity model. To extract the voxel outliers and to estimate the parameters of the different brain tissues in a robust way, we used the Trimmed Likelihood Estimator (TLE).

The TLE was proposed by [5] as a modification of the Maximum Likelihood Estimator in order to be robust to the presence of outliers. Using the TLE, the complete log-likelihood (5) can be expressed as:

$$\log L_C(\boldsymbol{\psi}, \boldsymbol{\xi}) \; = \; \log \left(\prod_{i=1}^{h} f(Y_{\nu(i)}, Z_{\nu(i)} | \boldsymbol{\psi}, \boldsymbol{\xi}_i) \right)$$

with $\nu = \{\nu(1), ..., \nu(N)\}$ being the corresponding permutation of indices sorted by their probability $f(Y_{\nu(i)}, Z_{\nu(i)} | \boldsymbol{\psi}, \boldsymbol{\xi}_i)$, and h is the trimming parameter corresponding to the percentage of values included in the parameter estimation.

The trimming proportion h will define a set of outlier voxels $\mathcal{C}$ which among others will encompass the MS lesions. Because the set of outlier voxels $\mathcal{C}$ contained no only MS lesions but also lesion false positives, a Graph-Cuts algorithm [6] was used to segment $\mathcal{C}$ into MS lesions or MS lesion false positives.

3. VALIDATION

3.1. Reference Population

Data from 15 volunteers was acquired on a 3T clinical MR scanner from GE Medical Systems (Waukesha, WI, USA) using an 8-channel receiver head coil. MR scans were acquired using three different pulse sequences: a T1-weighted MPRAGE (Magnetization Prepared Rapid Acquisition Gradient Echo) sequence; a T2-weighted scan from an FSE (Fast Spin Echo) sequence; and a FLAIR scan, also run with an FSE sequence. The T1w sequence was axially acquired; T2w and FLAIR sequences were sagitally acquired. All sequences were acquired with a matrix size of 256x256 and a field of view of 28 cm. Slice thickness was 1.3 mm for the T1w-MPRAGE sequence; 1 mm for the T2w-FSE sequence; and 2 mm for the FLAIR-FSE sequence. The MPRAGE parameters were TR 10/TE 6 ms with a flip angle of 8. For the FSE, the paramenters were TR 3000/TE 140 ms with an echo train length of 15.

After image acquisition, the T2w and FLAIR images were aligned to the T1w scan. In addition, all acquired MR images were re-oriented to an axial orientation, then the intra-cranial volume, CSF, GM and WM tissues were manually segmented by a trained expert.

To achieve accurate alignment between healthy volunteers and a patient with MS we used the algorithm proposed by Commowick et al. [7]. This nonlinear registration algorithm is not intrinsic to our method, and it is possible to use other non-linear registration approaches.

Because the intensity levels of the subject of interest and the reference population need to be in a comparable range, we used a linear transformation to match the median intensity of each modality of each reference subject to those found in the scans of the subject of interest.

3.2. Synthetic Data Validation

We used the BrainWeb synthetic MS brain phantom provided by the McConnell Brain Imaging Center (Montreal, Qc, Canada) [8]. The MS brain phantoms are based in the original BrainWeb healthy phantom, which has been extended with three different MS lesion volumes: mild ($0.4cm^3$), moderate ($3.5cm^3$) and severe ($10.1cm^3$). Each MS phantom was provided with their own ground truth which enables the validation of our algorithm's brain tissue segmentation (i.e. GM, WM and CSF) as well as MS lesion segmentation.

For each MS phantom (mild, moderate and severe), we generated both T1w and T2w MR images using the default acquisition parameters with different levels of noise ($n = 1\%, 3\%, 5\%, 7\% \, and \, 9\%$) and intensity inhomogeneity ($rf = 0\%, 20\% \, and \, 40\%$). Each image was acquired with an isotropic resolution of $1mm^3$. The brain volume was extracted for all images using the brain mask provided in the anatomical ground truth.

3.2.1. MS Lesion Segmentation

Our MS lesion detection algorithm relies in $\mathcal{R}$, which corresponds to the size of the neighborhood used to generate the local intensity model, and h which is the trimming parameter in the TLE estimation and also defines the sensitivity of the algorithm to detect MS lesions.

To estimate the optimal parameters for our segmentation algorithm, we employed the T1w and T2w MR images with 3%noise and 20% inhomogeneity with the three available lesion loads (mild, moderate and severe). MS lesion detection were performed using different values for $\mathcal{R}$ and h.

MS lesion delineation were evaluated computing the Dice Score (DSC) [9] between our estimated segmentation (Seg) and the BrainWeb ground truth (Ref). Figure 1 show DSC results for each synthetic phantom (mild, moderate and severe) using different values of TLE percentage h and neighborhood size $\mathcal{R}$. DSC values are higher for lower h values, independently of the $\mathcal{R}$ value, but starts worsening after $h > 6\%$ due to an increase of false positive lesions.

3.2.2. Noise and Intensity Inhomogeneity

Based in the optimal $\mathcal{R} = 2$ and $h = 5\%$ parameters found in the previous experiment, we evaluated the MS lesion segmentation achived by the algorithm under different conditions of noise and intensity inhomogeneity using the DSC.

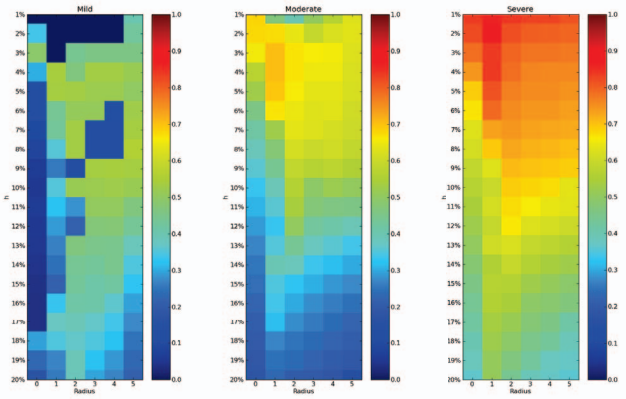


Fig. 1. BrainWeb DSC values for the MS lesion automatic segmentation varying the local neighborhood size $\mathcal{R}$ and TLE percentage h. From left to right: mild, moderate and severe lesion loads.

Results in Figure 2, show that our coupled local/global intensity model MS lesion segmentation is robust to the presence of noise in the image. The accuracy of our algorithm was reduced for images with high level of intensity inhomogeneity. This intensity artifact will cause MS lesions to have different intensity profiles, making the graph cuts algorithm to discard some of the lesions detected as outliers by the TLE.

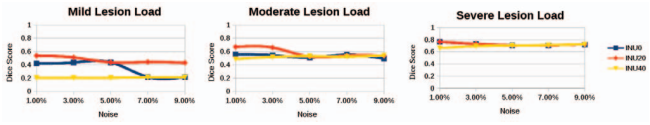


Fig. 2. BrainWeb MS lesion DSC results for mild, moderate and severe lesion load for all levels of noise and intensity inhomogeneity.

4. DISCUSSION

In this paper we proposed a novel automatic algorithm for segmentation of MS patients in MRI. The algorithm proposes a novel model which combines a within subject global tissue intensity model and a local intensity model derived from an aligned set of healthy reference subjects. MS lesions are segmented as outliers towards our new coupled local/global intensity GMM. Outliers include MS lesions but also voxels that do not follow the proposed model such as vessels or intensity artifacts. A graph-cuts algorithm is used to classify the model outliers into MS lesions and false positives.

BrainWeb MS lesion DSC results for mild, moderate and severe lesion load for different levels of noise and intensity inhomogeneity prove show that our coupled local/global intensity model MS lesion segmentation is robust to the presence of noise in the image. In the case of Mild lesion load, the accuracy of our algorithm was heavily impacted for a high level of intensity inhomogeneity. For Mild lesion load, intensity inhomogeneity artifact caused MS lesions to have different intensity profiles, making the graph cuts algorithm to discard some of the MS lesions detected as ouliers by the TLE. Such effect, suggest the need of pre-processing the MS patient MRI to correct for intensity inhomogeneities.

5. REFERENCES

[1] K Van Leemput, F Maes, D Vandermeulen, a Colchester, and P Suetens, "Automated segmentation of multiple sclerosis lesions by model outlier detection.," *IEEE transactions on medical imaging*, vol. 20, no. 8, pp. 677–88, Aug. 2001.

[2] J.C. Souplet, *Évaluation de l'atrophie et de la charge lésionnelle sur des séquences IRM de patients atteints de sclérose en plaques*, Ph.D. thesis, Universite de Nice-Sophia Antipolis, 2009.

[3] D Garcia-Lorenzo, S Prima, D Arnold, D Collins, and C Barillot, "Trimmed-Likelihood Estimation for Focal Lesions and Tissue Segmentation in Multi-Sequence MRI for Multiple Sclerosis.," *IEEE transactions on medical imaging*, , no. c, pp. 1–13, Feb. 2011.

[4] X. Tomas-Fernandez and Simon K. Warfield, "A new classifier feature space for an improved Multiple Sclerosis lesion segmentation," in *2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*. Mar. 2011, pp. 1492–1495, IEEE.

[5] N Neykov, P Filzmoser, R Dimova, and P Neytchev, "Robust fitting of mixtures using the trimmed likelihood estimator," *Computational Statistics & Data Analysis*, vol. 52, no. 1, pp. 299–308, Sept. 2007.

[6] DM Greig and BT Porteous, "Exact maximum a posteriori estimation for binary images," *of the Royal Statistical Society. Series*, vol. 51, pp. 271–279, 1989.

[7] Olivier Commowick, *Creation et Utilisation d'Atlas Anatomiques Numeriques pour la Radiotherapie (Design and Use of Anatomical Atlases for Radiotherapy)*, Ph.D. thesis, Universite de Nice-Sophia Antipolis, 2007.

[8] R K Kwan, A C Evans, and G B Pike, "MRI simulation-based evaluation of image-processing and classification methods.," *IEEE transactions on medical imaging*, vol. 18, no. 11, pp. 1085–97, Nov. 1999.

[9] Lee R. Dice, "Measures of the Amount of Ecologic Association Between Species," *Ecology*, vol. 26, no. 3, pp. 297, July 1945.