Journal of Medical Imaging

MedicalImaging.SPIEDigitalLibrary.org

Locally adaptive magnetic resonance intensity models for unsupervised segmentation of multiple sclerosis lesions

Alfiia Galimzianova
Žiga Lesjak
Daniel L. Rubin
Boštjan Likar
Franjo Pernuš
Žiga Špiclin

SPIE.

Alfiia Galimzianova, Žiga Lesjak, Daniel L. Rubin, Boštjan Likar, Franjo Pernuš, Žiga Špiclin, "Locally adaptive magnetic resonance intensity models for unsupervised segmentation of multiple sclerosis lesions," *J. Med. Imag.* **5**(1), 011007 (2017), doi: 10.1117/1.JMI.5.1.011007.

Journal of Medical Imaging 5(1), 011007 (Jan–Mar 2018)

Locally adaptive magnetic resonance intensity models for unsupervised segmentation of multiple sclerosis lesions

Alfiia Galimzianova,[a,b,*] Žiga Lesjak,[a] Daniel L. Rubin,[b] Boštjan Likar,[a,c] Franjo Pernuš,[a] and Žiga Špiclin[a,c]
[a]University of Ljubljana, Faculty of Electrical Engineering, Ljubljana, Slovenia
[b]Stanford University, School of Medicine, Palo Alto, California, United States
[c]Sensum, Computer Vision Systems, Ljubljana, Slovenia

Abstract. Multiple sclerosis (MS) is a neurological disease characterized by focal lesions and morphological changes in the brain captured on magnetic resonance (MR) images. However, extraction of the corresponding imaging markers requires accurate segmentation of normal-appearing brain structures (NABS) and the lesions in MR images. On MR images of healthy brains, the NABS can be accurately captured by MR intensity mixture models, which, in combination with regularization techniques, such as in Markov random field (MRF) models, are known to give reliable NABS segmentation. However, on MR images that also contain abnormalities such as MS lesions, obtaining an accurate and reliable estimate of NABS intensity models is a challenge. We propose a method for automated segmentation of normal-appearing and abnormal structures in brain MR images that is based on a locally adaptive NABS model, a robust model parameters estimation method, and an MRF-based segmentation framework. Experiments on multisequence brain MR images of 30 MS patients show that, compared to whole-brain MR intensity model and compared to four popular unsupervised lesion segmentation methods, the proposed method increases the accuracy of MS lesion segmentation. [DOI: 10.1117/1.JMI.5.1.011007]

Keywords: multiple sclerosis; unsupervised segmentation; locally adaptive models.

Paper 17197SSPRR received Jun. 30, 2017; accepted for publication Oct. 9, 2017; published online Nov. 1, 2017.

1 Introduction

Characterization, diagnosis, and prognosis of many neurological diseases rely on paraclinical symptoms observed on brain magnetic resonance (MR) images, which can be extracted and quantified by image analysis methods and are then referred to as disease imaging markers. Such analysis generally involves segmentation of brain MR images into normal-appearing and abnormal structures. For instance, multiple sclerosis (MS) is a disease characterized by focal inflammatory lesions disseminated in the brain parenchyma and in the spinal cord; therefore, an accurate segmentation of these lesions in the MR images is required to obtain the established imaging markers, such as lesion volume, count, and location.[1] Recent research shows that atrophy of normal-appearing brain structures (NABS) is also important for characterizing the progression of MS disease and response to pharmacological treatment.[2] Segmentation of MS lesions can be performed manually; however, this process is subjective, tedious, and time-consuming, especially with recent trends toward high-resolution isotropic three-dimensional (3-D) brain MR imaging and large patient cohorts imaged in clinical studies. For the task of accurate and reproducible segmentation of MS lesions, which can also be obtained with high efficiency, we consider here the automated methods.

Automated segmentation of brain MR images is a challenging task because of MR acquisition imperfections (MR bias field and image noise), complex brain anatomy, and varying manifestations of the abnormal structures. MR imaging provides a high contrast among the major brain structures, such as gray matter (GM), white matter (WM), and compartments of cerebrospinal fluid (CSF), which are jointly referred to as NABS. Because of high contrast between the three NABS in MR images, several recent methods perform segmentation of the NABS and abnormal structures through estimation of the underlying MR intensity model of NABS. The intensity model, however, is difficult to estimate in the presence of spatial MR intensity variations that result from the MR bias field,[3] non-stationary noise,[4] and structural intensity nonuniformity.[5]

Methods for MS lesion segmentation that are based on a whole-brain (WB) MR intensity model of NABS can either simultaneously perform the bias field correction and segmentation[6] or perform intensity bias correction as a preprocessing step.[7] However, because the intensity bias is also structure dependent,[5] it cannot be easily compensated without an accurate NABS intensity model at hand. To implicitly compensate the intensity bias of different origins, a number of methods for the segmentation of NABS, which perform local MR intensity modeling of NABS, were proposed.[8–10] For example, the NABS segmentation method by Scherrer et al.[8] is based on the estimation of local and cooperative Markov random fields (MRFs) on nonoverlapping cubic subvolumes that estimate the local intensity distribution and thus handle the nonuniformity without explicit bias field modeling. Conversely, in the method by Shattuck et al.,[10] the bias field is estimated and corrected prior to segmentation by the estimation of local MR intensity

*Address all correspondence to: Alfiia Galimzianova, E-mail: alfiia@stanford.edu

2329-4302/2017/$25.00 © 2017 SPIE

models of NABS in overlapping subvolumes. However, application of locally adaptive NABS segmentation methods[8–10] to MR images containing abnormal structures, such as lesions, is not straightforward.

For the segmentation of both NABS and lesions, locally adaptive intensity models of NABS and of lesions were introduced by Harmouche et al.,[11] where, based on training datasets, the intensity models were learned individually on several different anatomical subregions of the brain. A method for post-processing the NABS and lesion segmentation was proposed by Biediger et al.,[12] which aims to improve segmentation by local region-growing cellular automata, using as seed points the voxels previously labeled as lesions. The main idea is to improve the boundaries of the segmented lesions based on local properties of the MR images.

In this paper, we propose a method for unsupervised segmentation of normal-appearing and abnormal structures in brain MR images of MS patients. The proposed method employs a robust unsupervised mixture estimation based on confidence-level (CL) outlier detection,[13] which is executed on multiple overlapping subvolumes of brain MR image. The subvolumes are positioned on a rectangular lattice that spans across the coregistered multisequence MR images, and in each subvolume, a simplified local NABS intensity model is estimated. The estimated local models on the lattice are interpolated to form a local model at each voxel and, then, the fuzzy membership maps of NABS and MS lesions are computed and used to initialize the MRF-based segmentation. The proposed method was evaluated on 30 sets of 3T MR images of MS patients, which were acquired and annotated using the same protocol consistently. The experiments of this preliminary study indicate that the locally adapted modeling improves over the commonly used WB approach and outperforms four other tested unsupervised methods.

2 Methods

Consider a dataset of brain MR images with M sequences represented as a nonoriented graph $(\mathcal{V}, \mathcal{E})$, in which the brain voxels are graph vertices $\mathcal{V}$ and codependent voxels are represented by the vertices connected by graph edges $\mathcal{E}$. Let real-valued vectors $Y = \{y_j\}_{j\in\mathcal{V}} \subset \mathbb{R}^\nu$ represent ν-variate observations, or features, obtained from M different sequences of MR, and categorical variables $Z = \{z_j\}_{j\in\mathcal{V}} \subset \mathbb{Z}$ represent the hidden labels or image segmentation. From a probabilistic point of view, the hidden labels can be obtained by maximum a posteriori (MAP) estimation of hidden label probabilities given the image sets, i.e., $Z^* = \arg\max_{Z\in\mathbb{Z}} P(Z|Y)$. We model this posterior as an MRF, wherein we assume positivity $P(Z|Y) > 0 \;\; \forall\, Z \in \mathbb{Z}$ and the Markovianity $P(z_j|Z_{\mathcal{V}\setminus\{j\}}, Y) = P(z_j|Z_{N_j}, Y)$ on the neighborhood $\mathcal{N}_j = \{i \in \mathcal{V} | (i,j) \in \mathcal{E}\}$ of each voxel $j \in \mathcal{V}$. According to Hammersley–Clifford theorem,[14] the distribution of such a field has the form of $P(Z|Y) = \frac{1}{\sum_{Z'\in\mathbb{Z}} \exp[-U(Z'|Y)]} \times \exp[-U(Z|Y)]$ defined by its energy function $U(Z|Y)$, and therefore, the hidden labels can be found as $Z^* = \arg\min_{Z\in\mathbb{Z}} U(Z|Y)$.

In graphical notation, the energy function takes the following form:

$$U(Z|Y) = \sum_{j\in\mathcal{V}} -\log\, U_{\mathcal{V}}(z_j) + \eta \sum_{(j,i)\in\mathcal{E}} U_{\mathcal{E}}(z_j, z_i), \tag{1}$$

where $U_{\mathcal{V}}$ and $U_{\mathcal{E}}$ are the data and smoothness terms, respectively. The smoothness term $U_{\mathcal{E}}$ weighted by a coefficient η was modeled by Potts model

$$U_{\mathcal{E}}(z_j, z_i) = \begin{cases} 0, & \text{if } z_j \neq z_i \\ 1, & \text{if } z_j = z_i \end{cases}. \tag{2}$$

The data term $U_{\mathcal{V}}$ was defined as a linear combination of intensity potential f_I and location potential f_L, which are dependent on the intensity features $Y^I = \{y_j^I \subset y_j, y_j^I \in \mathbb{R}^M\}_{j\in\mathcal{V}}$ and the spatial features $Y^L = \{y_j^L \subset y_j, y_j^L \in \mathbb{R}^3\}_{j\in\mathcal{V}}$, respectively, i.e.,

$$U_{\mathcal{V}}(z_j|Y) = -f_I(z_j|Y^I) - \gamma f_L(z_j|Y^I, Y^L), \tag{3}$$

where γ is a weight coefficient.

In the following, we describe in detail both the intensity and location potentials for the task of NABS and MS lesion segmentation (see also Fig. 1).

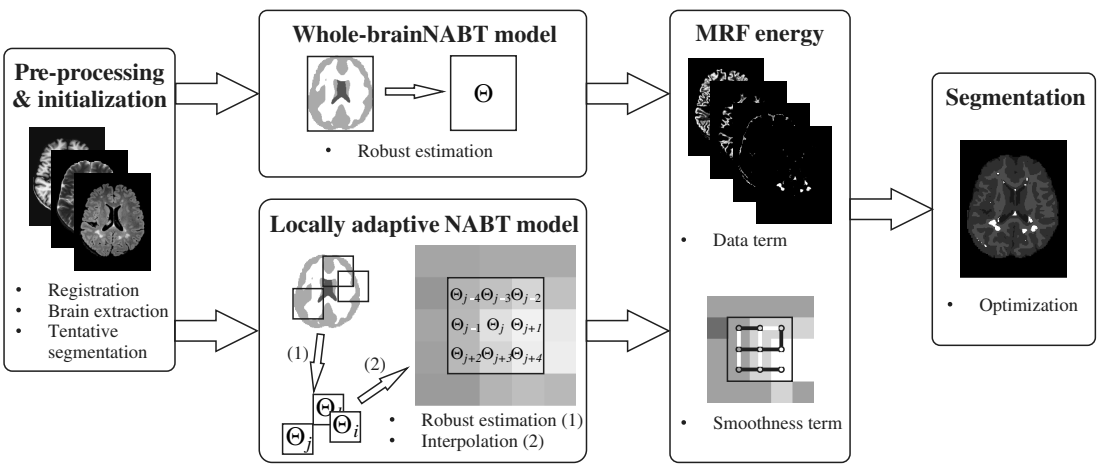

Fig. 1 The workflow of the proposed segmentation method based either on WB NABS intensity model (upper row) or on the locally adaptive NABS model (lower row).

2.1 *Intensity Potential Based on Locally Adaptive NABS Intensity Model*

Because only a few major brain structures have high contrast in the structural MR images, they can be compactly represented by a structural intensity model. We considered GM, WM, and CSF to represent a local NABS model, in which each voxel intensity y_j^I is modeled by the probability distribution of a Gaussian mixture model (GMM) with voxel-wise parameter vectors Θ_j as

$$p(y_j^I|\Theta_j) = \sum_{k\in\{\mathrm{GM,WM,CSF}\}} \pi_k^j g(y_j^I|\mu_k^j, \Lambda_k^j), \tag{4}$$

where parameters of GMM at voxel j are the means μ_k^j and covariances Λ_k^j of the component k and component weights $\pi_k^j \in (0,1)$, which sum to one. These parameters form a mixture parameter vector $\Theta_j = \{\pi_k^j, \mu_k^j, \Lambda_k^j\}_{k\in\{\mathrm{GM,WM,CSF}\}}^{j\in\mathcal{V}}$. If all local NABS parameters are equal throughout all image voxels, the model becomes a stationary WB GMM, commonly used in brain MR intensity modeling.[6,7,15] The assumption that NABS parameters may vary across the brain volume leads to a locally adaptive brain intensity model.[8,10]

The voxel-wise local model parameters estimated for the given intensity features, i.e., $\Theta_j(Y^I) = \{\pi_k^j, \mu_k^j, \Lambda_k^j\}_{k=\{\mathrm{GM,WM,CSF}\}}^{j\in\mathcal{V}}$ of the mixture [Eq. (4)], define the intensity potential of the data term [Eq. (3)] for NABS and MS lesions. Generalizing the method in Ref. 15, the structure memberships for GM, WM, and CSF were modeled by the corresponding confidence levels $\mathrm{CL}_k[y_j|\Theta_j(Y^I)]$ of the component estimates as

$$f_I(z_j = k|Y^I) = \log\{1 - \mathrm{CL}_k[y_j^I|\Theta_j(Y^I)]\},$$
$$k \in \{\mathrm{GM, WM, CSF}\}, \tag{5}$$

while the MS lesion class membership was determined by combining the fuzzy membership maps of the outlier scores and hyperintensity maps as

$$f_I(z_j = \mathrm{MS}|Y^I) = \log\{\mathrm{CL}_{\mathrm{MAP}}[y_j^I|\Theta_j(Y^I)] \underset{m'}{\wedge} W_{m'}[y_j^I|\Theta_j(Y^I)]\}, \tag{6}$$

where $\wedge$ is the fuzzy AND operator, and $m' \in \overline{1, M}$ are the MR sequences in, which lesions appear hyperintense compared to the NABS, e.g., T2 and/or FLAIR [see Eq. (7)]. The outlier scores are computed as voxel-wise CL of the MAP, $\mathrm{CL}_{\mathrm{MAP}}$, of components at that voxel following Ref. 13, which can also be interpreted as a fuzzy membership map since its values lie in interval [0, 1]. The CL for the k'th component of a GMM is computed from the cumulative density of χ_m^2 distribution of Mahalanobis distances $d_{\Lambda_k}^2(y_j|\mu_k) = (y_j - \mu_k)\Lambda_k^{-1}(y_j - \mu_k)^T$. The fuzzy membership maps $W_{m'}$ were defined as marginal Z-scores, i.e., $Z_{k',m'}(y_j) = (y_{j,m'}^I - \mu_{k',m'})/\sigma_{k',m'}$ of observations with respect to the $k' \in \{\mathrm{GM, WM}\}$ structures as

$$W_{j,m'} = \begin{cases} 0, & \text{if } Z_{k',m'}(y_j) < \mathrm{thr}_1(k',m') \\ 1, & \text{if } Z_{k',m'}(y_j) > \mathrm{thr}_2(k',m') \\ \frac{Z_{k',m'}(y_j) - \mathrm{thr}_1(k',m')}{\mathrm{thr}_2(k',m') - \mathrm{thr}_1(k',m')}, & \text{otherwise} \end{cases}, \tag{7}$$

where m' represents MR sequences $m' \in \overline{1, M}$, corresponding to T2w and/or FLAIR modalities. The thresholds $\mathrm{thr}_1(k', m')$ and $\mathrm{thr}_2(k', m')$ indicate the levels of Mahalanobis distances and define the corresponding hypo- or hyperintense MR intensity levels, respectively.

The unsupervised parameter estimation procedure that was performed to obtain the local intensity potential of a given MR image dataset is further detailed in Sec. 2.3.

2.2 *Location Potential Based on Probabilistic Atlas Coregistration*

The location potential in Eq. (3) was modeled as $f_L(z_j|Y^I, Y^L) = \log\{P[z_j|\mathcal{A}(Y^I, Y^L)]\}$, where $\mathcal{A} = \{\mathcal{A}_{\mathrm{GM}}, \mathcal{A}_{\mathrm{WM}}, \mathcal{A}_{\mathrm{CSF}}\}$ is the set of probabilistic atlases, which are spatially aligned to input MR images by intensity-based registration[16] of the average T1w atlas and the input T1w images. We used the MNI305 atlas[17] that provides the prior probabilities of GM, WM, or CSF classes, i.e., $P[z_j = k|\mathcal{A}(Y^I, Y^L)] = \mathcal{A}_k(Y^I, Y^L)$, $k \in \{\mathrm{GM, WM, CSF}\}$ while limiting the location of MS lesions to the regions with WM class label, i.e., $P[z_j = \mathrm{MS}|\mathcal{A}(Y^I, Y^L)] = \mathcal{A}_{\mathrm{WM}}(Y^I, Y^L)$.

2.3 *Robust Estimation of Local NABS Intensity Models*

Due to various MR acquisition artifacts and the presence of abnormal structures such as MS lesions, the unsupervised estimation of model parameters [Eq. (4)] has to be robust to outliers in MR intensity distributions. Let $\Theta(\tilde{Y}^I)$ denote an estimate of the parametric NABS intensity model from the intensity subsample $\tilde{Y}^I \subset Y^I$. Next, we consider that the local parameters Θ_j can be estimated on a subsample of observations in a close proximity to the voxel j. Thus, the subsample can be obtained from a local subvolume S_j, centered at corresponding voxel j, and the parameters of local model can then be estimated over the corresponding subvolumes, $\Theta_j(Y_{S_j}^I)$.

Under the assumption that the intensity bias field, structural nonuniformities, and noise nonstationarities vary smoothly across the image, the local models Θ_j can be approximated from parameters $\Theta_i(Y_{S_j}^I)$ estimated on a sparse lattice of voxels $i \in H^D \subset \mathcal{V}$, which is defined by the spacing $D = (d_1, d_2, d_3)$ and then interpolated to obtain the parameter estimates at remaining voxels.[8,10] Local subvolumes are centered at the lattice nodes, and subvolume size was chosen such that the neighboring subvolumes overlapped up to 50%, i.e., $S_i = [(y_i^L)_1 - d_1, L_1(y) + d_1] \times [L_2(y) - d_2, L_2(y) + d_2] \times [L_3(y) - d_3, L_3(y) + d_3]$ where $L(y)$ is the spatial location of the voxel i in 3-D. A relatively high overlap of 50% between the subvolumes is employed to obtain consistent estimates at neighboring vertices in the lattice.

The use of local estimation of NABS model parameters was shown to improve MR intensity modeling and, consequently, the segmentation accuracy on brain MR images without abnormalities.[10] On MR images of the brains that contain abnormal structures, however, the unpredictable nature and amount of the MR intensity model outliers preclude a direct application of previous approaches for model estimation without resorting to the use of robust estimators. We propose to use the state-of-the-art robust trimmed-likelihood-based estimator,[13] which was shown to give accurate mixture parameters estimates for samples contaminated by up to 50% of outliers. The estimator

employs CL ordering to select the model outliers based on current model parameters Θ and determines the model parameters by solving a local optimization problem $\Theta_i(\tilde{Y}^I) = \arg\min_\Theta \prod_{i\in S_i'} p(y_i|\Theta)$. The subset of voxels $S_i' \subset S_i$ of size $[|S_i|(1-\alpha)]$ represents inliers found after excluding a fraction α of model outliers from the set S_i.

Once the parameter estimates of local models are obtained on the sparse grid, the voxel-wise parameters of the NABS model [Eq. (4)] are approximated by interpolation. The linear interpolation of the multidimensional parameters of Eq. (4) coincides with the stratified parameter estimation[18,19] and is computed as a linear combination of the related found multidimensional parameters

$$\begin{aligned}\pi_k^j &= \sum_{i\in H^D} w_i^j \pi_k^i,\\ \mu_k^j &= \sum_{i\in H^D} w_i^j \mu_k^i,\\ \Lambda_k^j &= \sum_{i\in H^D} w_i^j [\Lambda_k^i + (\mu_k^i - \mu_k^j)(\mu_k^i - \mu_k^j)^T],\end{aligned} \tag{8}$$

where w_i^j are the weights of the vertex i on the lattice regarding the voxel j. The linear interpolation [Eq. (8)] is performed by an interpolation function $\tilde{f}:\mathbb{R}\to\mathbb{R}$ on the 3-D lattice per vector component (with indices $m,n \in \overline{1,M}$) as follows:

$$\begin{aligned}[(\pi_k^j)_m]_{j\in\mathcal{V}} &= \tilde{f}\{[(\pi_k^i)_m]_{i\in H^D}\},\\ [(\mu_k^j)_m]_{j\in\mathcal{V}} &= \tilde{f}\{[(\mu_k^i)_m]_{i\in H^D}\},\\ [(\Lambda_k^j)_{m,n}]_{j\in\mathcal{V}} &= \tilde{f}\{[(\Lambda_k^i)_{m,n}]_{i\in H^D}\} + \tilde{f}\{[(\mu_k^i)_m \cdot (\mu_k^i)_n]_{i\in H^D}\}\\ &\quad - [(\mu_k^j)_m]_{j\in\mathcal{V}} \cdot \tilde{f}\{[(\mu_k^i)_n]_{i\in H^D}\}\\ &\quad - [(\mu_k^j)_n]_{j\in\mathcal{V}} \cdot \tilde{f}\{[(\mu_k^i)_m]_{i\in H^D}\}\\ &\quad + [(\mu_k^j)_m \cdot (\mu_k^j)_n]_{j\in\mathcal{V}},\end{aligned} \tag{9}$$

where $\cdot$ denotes the component-wise product.

3 Experiments and Results

3.1 *Datasets*

Brain images of 30 patients with MS were acquired on a 3T Siemens MR scanner. The database consisted of datasets of multislice T1- and T2-weighted and multislice FLAIR images. A sequence of preprocessing steps was performed on each dataset, including intrasubject registration of MR sequences,[16] brain mask extraction on T1w image,[20] and intensity inhomogeneity correction.[21] Datasets were resampled to a resolution of $1\times1\times3$ mm^3. The reference segmentation was obtained following the protocol defined in Ref. 22. The MS lesions were first independently segmented by two expert raters using mainly FLAIR sequences with occasional consideration of the coregistered T1- and T2-weighted images using the Segmentation Assistant tools of the publicly available BrainSeg3D software (available online at Ref. 23), which were previously shown to improve the consistency of expert annotations of MS lesions.[24] The raters revised the merged segmentation maps in several joint sessions so as to integrate and harmonize their expertise and finally reach a consensus on segmentation of lesions. The final consensus segmentations were used as the reference. Performance of the raters compared to the consensus was 0.78 ± 0.20 and 0.73 ± 0.19 in terms of mean Dice similarity coefficient (DSC).

The 30 datasets were classified according to the total lesion load (TLL), computed from the reference lesion segmentations, into three groups of subjects with mild (10 patients, TLL <5.5 cm^3), moderate (10 patients, 5.5 cm$^3 \le$ TLL ≤ 20 cm^3), and severe TLL (10 patients, 20 cm$^3 <$ TLL < 42 cm^3). On five datasets, the raters also delineated the normal brain structures, i.e., WM, GM, and CSF. The evaluation criteria computed between the reference segmentation and the segmentations obtained by the three tested automated methods were based on Dice similarity coefficient $\mathrm{DSC} = (2\times\mathrm{TP})/(\mathrm{FP}+\mathrm{FN}+2\times\mathrm{TP})$, where TP, FP, and FN represent the fractions of voxel labeled as true positive, false positive, and false negative, respectively.

3.2 *Parameters and Optimization*

In each subvolume, the robust GMM estimator[13] was executed with the trimming fraction set to $\alpha = 0.3$ while the initial parameters were obtained as sample estimates based on WB NABS segmentation of the T1w image with unsupervised nonparametric Otsu's thresholding method.[25] To exclude the unreliable estimates, the subvolumes that initially had the local fraction of outliers higher than the value of 0.5 were iteratively enlarged by scale factor of 1.5 until the aforementioned requirement was met. The hyperintensity maps [Eq. (7)] were computed for both methods using thresholds $\mathrm{thr}_1 = 2$, $\mathrm{thr}_2 = 3$, w.r.t. WM in T2w and FLAIR modalities and combined to the data term [Eq. (3)] with the location potential weighted by $\gamma = 1$. The smoothness in Eq. (2) was computed by the Potts model with six-voxel neighborhood $\mathcal{N}_j$ and the weight set to $\eta = 0.1$. The final segmentations were obtained by minimizing the energy function [Eq. (1)] by an accurate and efficient optimizer.[26]

3.3 *Evaluation and Comparison*

To make a fair comparison, we compare the proposed unsupervised segmentation method with other popular unsupervised methods implemented on our database. We tested four such methods: uVL2001 by Van Leemput et al.,[6] uGL2009 by García-Lorenzo et al.,[15] uGL2011 by García-Lorenzo et al.,[7] and LTOADS by Shiee et al.[27] The methods uVL2001, uGL2009, and uGL2011 were implemented using an improved NABS model,[19] whereas LTOADS was applied using its publicly available implementation (available online at Ref. 28). The proposed local MR modeling was first validated and compared to the WB stationary model on a database of images of higher resolution (1 mm^3), and the increase in subvolume size was found to improve the segmentation performance, with the best performance achieved on the smallest considered subvolume of 40 mm^3 (see Appendix A for details). For this experiment on datasets of a lower resolution, we tested the proposed method that used the locally adaptive NABS intensity model on a regular lattice with spacing of 30 mm and with local subvolumes of 60^3mm^3.

Segmentation performance in terms of DSC over the three groups of TLL and over all datasets is shown in Fig. 2. The proposed method outperformed the other four methods on the whole database, as well as in groups of moderate and severe TLL and performed second best after LTOADS in mild TLL group. The statistical significance of the difference between

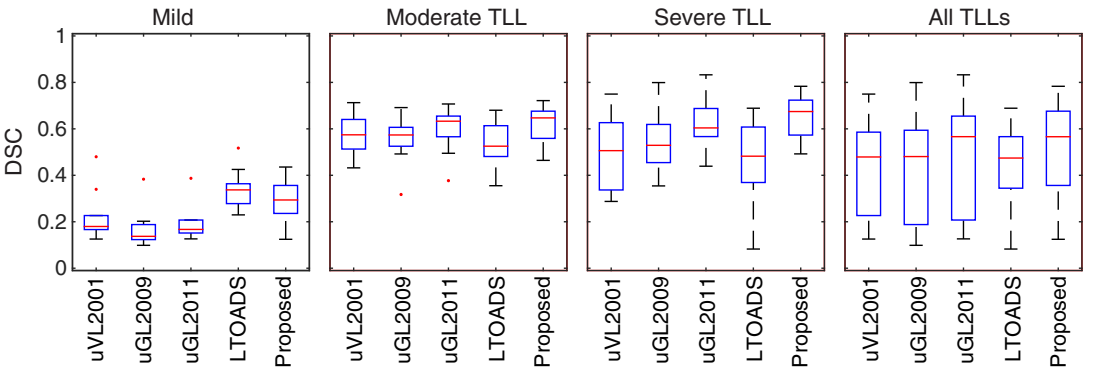

Fig. 2 Performance of the automated lesion segmentation methods in terms of DSC on datasets grouped according to TLL and on the whole dataset.

the proposed and baseline methods was tested using Wilcoxon's signed rank test. The differences were found significant ($p < 0.05$) for 10 out of 16 comparisons (see Table 1). No significant difference was found for the only decreased performance w.r.t. LTOADS in mild TLL group, for increased performance w.r.t. uGL2011 in severe TLL group, and for the increased performance across all the methods in moderate TLL group. On the whole database, the proposed method statistically significantly outperformed all the comparators. The consistency of lesion segmentation was validated by analysis of the TLL as estimated from the reference and the automated segmentations in terms of Pearson's correlation coefficient (CC) value. In this regard, the proposed method performed second best after LTOADS; however, its resulting estimates were closer to the reference segmentation TLLs (Fig. 3).

Three of the methods that perform NABS segmentation, i.e., uVL2001, LTOADS, and the proposed method, were compared on five datasets with manual delineations of CSF, GM, and WM. The resulting DSC values as shown in Fig. 4 indicate that the proposed method in general outperformed uVL2001 but was slightly worse than LTOADS. However, the stability of the proposed method is clearly better as the variation between different datasets is the lowest among the tested methods.

Example segmentations by the tested segmentation methods for datasets of the three different TLL groups (Fig. 5) are shown in Fig. 6. Note that the segmentations obtained by the proposed method were more specific (less false positives) than those obtained by uVL2001, uGL2009, and uGL2011, and also more sensitive (more true positives) than LTOADS.

Table 1 Median DSC differences of the proposed method and the four baseline segmentation methods. In parentheses are the *p*-values of Wilcoxon's signed rank test.

Datasets	uVL2001	uGL2009	uGL2011	LTOADS
Mild TLL	0.07 ($p = 0.049$)	0.14 ($p = 0.004$)	0.11 ($p = 0.014$)	−0.06 ($p = 0.105$)
Moderate TLL	0.01 ($p = 0.322$)	0.08 ($p = 0.083$)	0.02 ($p = 0.625$)	0.09 ($p = 0.064$)
Severe TLL	0.14 ($p = 0.001$)	0.10 ($p = 0.010$)	0.04 ($p = 0.193$)	0.14 ($p = 0.002$)
All	0.07 ($p = 0.000$)	0.11 ($p = 0.000$)	0.04 ($p = 0.006$)	0.09 ($p = 0.020$)

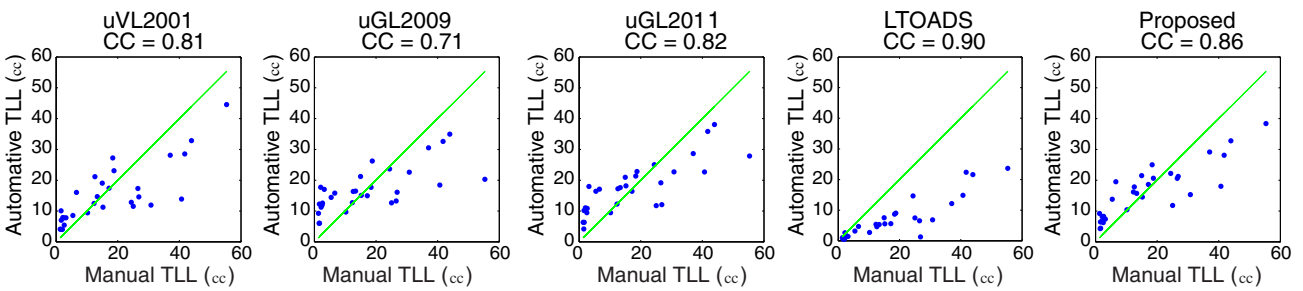

Fig. 3 Correlations between the automated segmentation methods and the reference lesion segmentation (in terms of Pearson's CC).

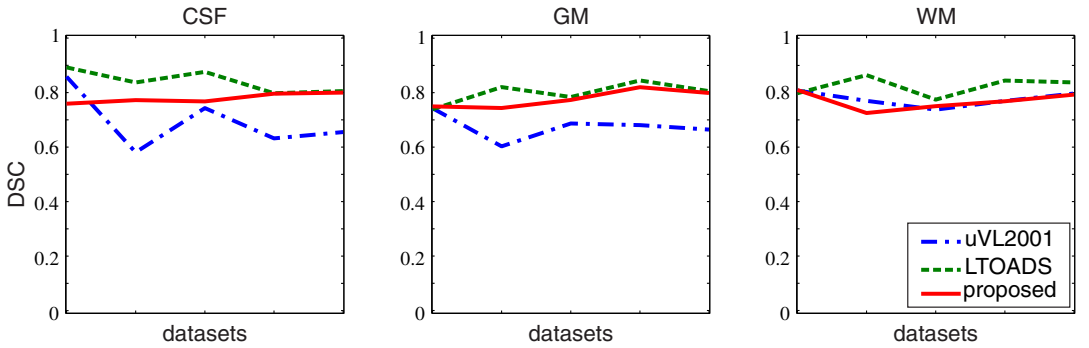

Fig. 4 Performance of automated segmentation of the CSF, normal-appearing GM, and WM in terms of DSC on five datasets with manual segmentation of the structures.

4 Discussion

We proposed a method for the automated unsupervised segmentation of normal-appearing and abnormal structures in brain MR images that effectively incorporates locally adaptive NABS intensity model, robustly estimates its parameters, and then uses the NABS model to perform the MRF-based segmentation. The proposed method was tested on datasets of MR images of 30 patients with MS and, compared to four other unsupervised segmentation methods, showed improved segmentation of MS lesions.

The main component of the proposed segmentation method is the locally adaptive NABS model, which aims to address the local variations of MR intensities in brain images. The within-structure intensity variations, including NABS and the MS lesions, are of the main concern,[5] which depend on the anatomical location within the brain. One possible solution is to model each major structure of the NABS as a mixture model.[5,19] Conversely, in our approach, we use a simple three-component GMM as the NABS model, but to account for the aforementioned anatomical variations then split the MR volume into small subvolumes, in which these variations have a lesser effect. Another source of the spatial intensity variabilities is the MR acquisition imperfections, such as nonstationarity of the noise, which can be addressed by noise suppression,[4] and intensity bias field, which can be addressed by the bias field correction methods.[3] The noise suppression and bias field correction are typically executed in a preprocessing step. Local adaptation of the NABS intensity model implicitly addresses these spatially varying MR acquisition imperfections or might even reduce these artifacts if they are still present after the preprocessing step was carried out.

Locally adaptive NABS models were previously considered in several methods for the segmentation of brain MR images of healthy subjects.[8–10] However, when the abnormal structures are also present in the MR images, the estimation of such models on local subvolumes becomes very difficult due to several reasons: (1) the risk of overfitting the NABS model when there is a high fraction of volume containing abnormalities w.r.t. the volume of normal structures and (2) difficulty of predicting in advance the volume or fraction of volume occupied by abnormal structures. We effectively solved these issues by the use of a robust GMM estimator,[13] which is designed such that the estimates are consistent up to high fractions of outliers (e.g., volume fraction of abnormalities can be up to 50%) and, therefore, an accurate assessment of the fraction of outliers is not required in advance.

Recent challenge summary in Ref. 29 reports mean DSC values for 14 methods ranging from 0.44 to 0.64, while on our dataset, the values were from 0.43 to 0.47 for the four baseline methods and 0.53 for the proposed method. Although these values are comparable, due to the image and annotation quality and protocol variability in different datasets, performance of the methods often varies considerably, and the fair comparison is only possible when all the methods are applied on the same database as was done in this work. Variability within the same database can also be caused by the inhomogeneity of the MS lesion load across the subjects. Similar to the work in Ref. 7, we additionally studied segmentation performance in three groups defined by the TLL (datasets with mild, moderate, and severe TLL).

The proposed automated segmentation method is unsupervised, thus does not require any training data to learn the parameters and instead adapts to the given data by robust estimation of the NABS parameters. A well-known disadvantage of the unsupervised methods is that on the same dataset, they typically provide accuracy lower than the supervised methods that are trained on the portion of that dataset.[30] Nevertheless, their advantage is in increased generalizing ability compared to supervised methods, which cannot be straightforwardly applied to datasets that are significantly different from their training datasets. In addition, the unsupervised methods do not require the laborious and time-consuming manual annotation for training, which gives them more potential for translation to practice.

A limitation of this study is the relatively small database of 30 MS patient images, which were obtained following the same imaging and annotation protocols. The future work will include multisite validation on datasets from different vendors and annotations with different protocols as well as application of the methodology to different neurological conditions.

5 Conclusion

We proposed a method for robust fully automated unsupervised segmentation of normal-appearing and abnormal structures in brain MR images that is based on the unsupervised estimation of local intensity models, which effectively compensate the spatial variability of the structure intensity and the MR intensity bias field and, therefore, accurately model the intensity distributions of NABS. The method was validated on datasets of brain MR images of 30 patients with MS and compared to four other unsupervised methods.[6,7,15,27] The experiments indicate that the locally adapted modeling gives a statistically significant improvement over the WB modeling approach and the tested methods.

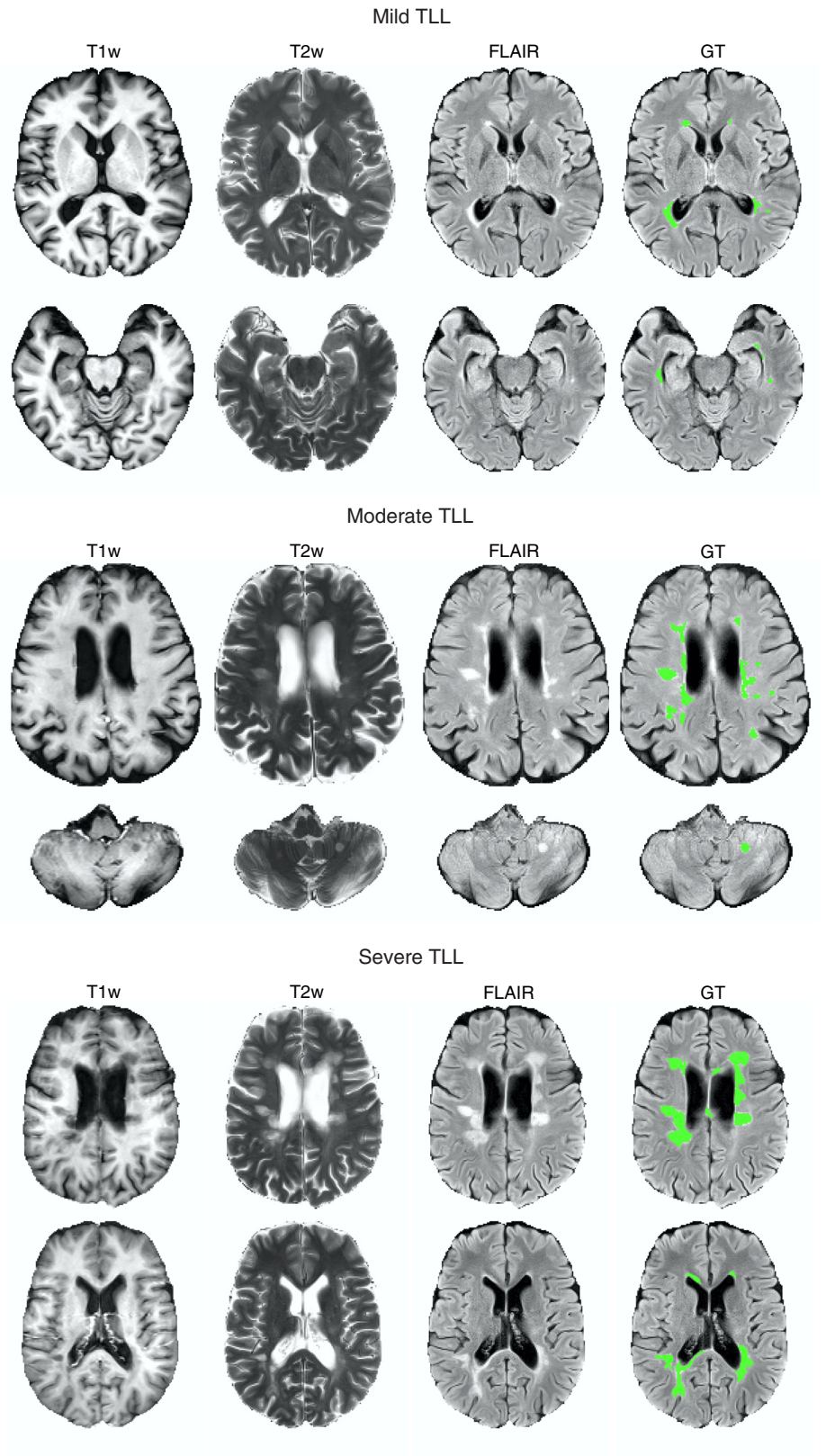

Fig. 5 MR images and the reference segmentations of three patients with mild, moderate, and severe TLLs.

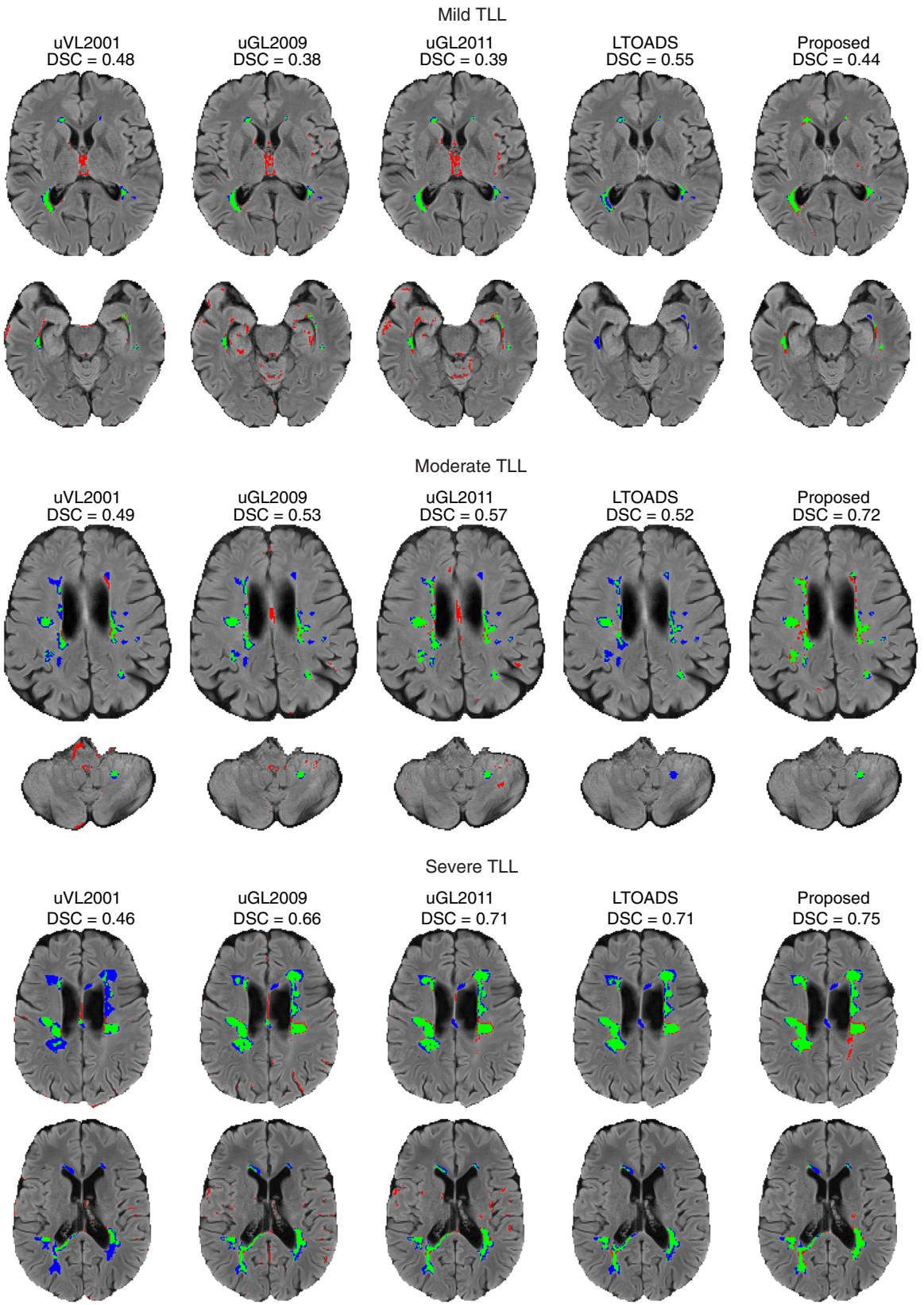

Fig. 6 Lesion segmentations by five automated methods labeled as true positives (in green), false negatives (in blue), or false positives (in red).

Appendix

In this section, we provide results of additional experiments for validation of the proposed locally adaptive MR intensity model.

A1 Datasets

For this evaluation experiment, we used a database different from the one used for the experiments described in the main paper. Specifically, the FLAIR sequence was of higher resolution, which leads to more variability in intensities across the image subvolumes. Brain images of 30 patients with MS were acquired on a 3T Siemens MR scanner. The database consisted of datasets of multislice T1- and T2-weighted and 3-D FLAIR. A sequence of preprocessing steps was performed on each dataset, including intrasubject registration of MR sequences,[16] brain mask extraction on T1w image,[20] intensity inhomogeneity correction.[21] Datasets were resampled to an isotropic resolution of $1 \times 1 \times 1$ mm^3.

The 30 datasets were classified according to the TLL, computed from the reference lesion segmentations, into three groups of subjects with mild (10 patients, TLL < 5.5 cm^3), moderate (10 patients, 5.5 cm$^3 \leq$ TLL ≤ 20 cm^3), and severe TLL (10 patients, 20 cm$^3 <$ TLL < 42 cm^3). On five datasets, the raters also delineated the normal brain structures, i.e., WM, GM, and CSF. The evaluation criteria computed between the reference segmentation and the segmentations obtained by the three tested automated methods were based on Dice similarity coefficient $\mathrm{DSC} = (2 \times \mathrm{TP})/(\mathrm{FP} + \mathrm{FN} + 2 \times \mathrm{TP})$, true positive rate $\mathrm{TPR} = \mathrm{TP}/(\mathrm{TP} + \mathrm{FN})$, and false discovery rate $\mathrm{FDR} = \mathrm{FP}/(\mathrm{FP} + \mathrm{TP})$, where TP, FP, and FN represent the fractions of voxel labeled as true positive, false positive, and false negative, respectively.

A2 Evaluation of the Model

Here, we analyze the performance of the proposed segmentation method with WB and the locally adaptive NABS intensity models with different lattice sampling across the brain volume on our database with 3-D FLAIR images. The lattice sampling was varied in four step sizes of $d_{1,2,3} \in \{20,30,40,50\}$ mm and resulted in subvolumes of $|S_{i \in H^D}| \in \{40^3, 60^3, 80^3, 100^3\}$ mm^3, respectively. On these lattices of different sampling steps, the locally adaptive NABS models were estimated.

Figure 7 shows the evaluation criteria computed over three different TLL groups and over the whole database of the 30

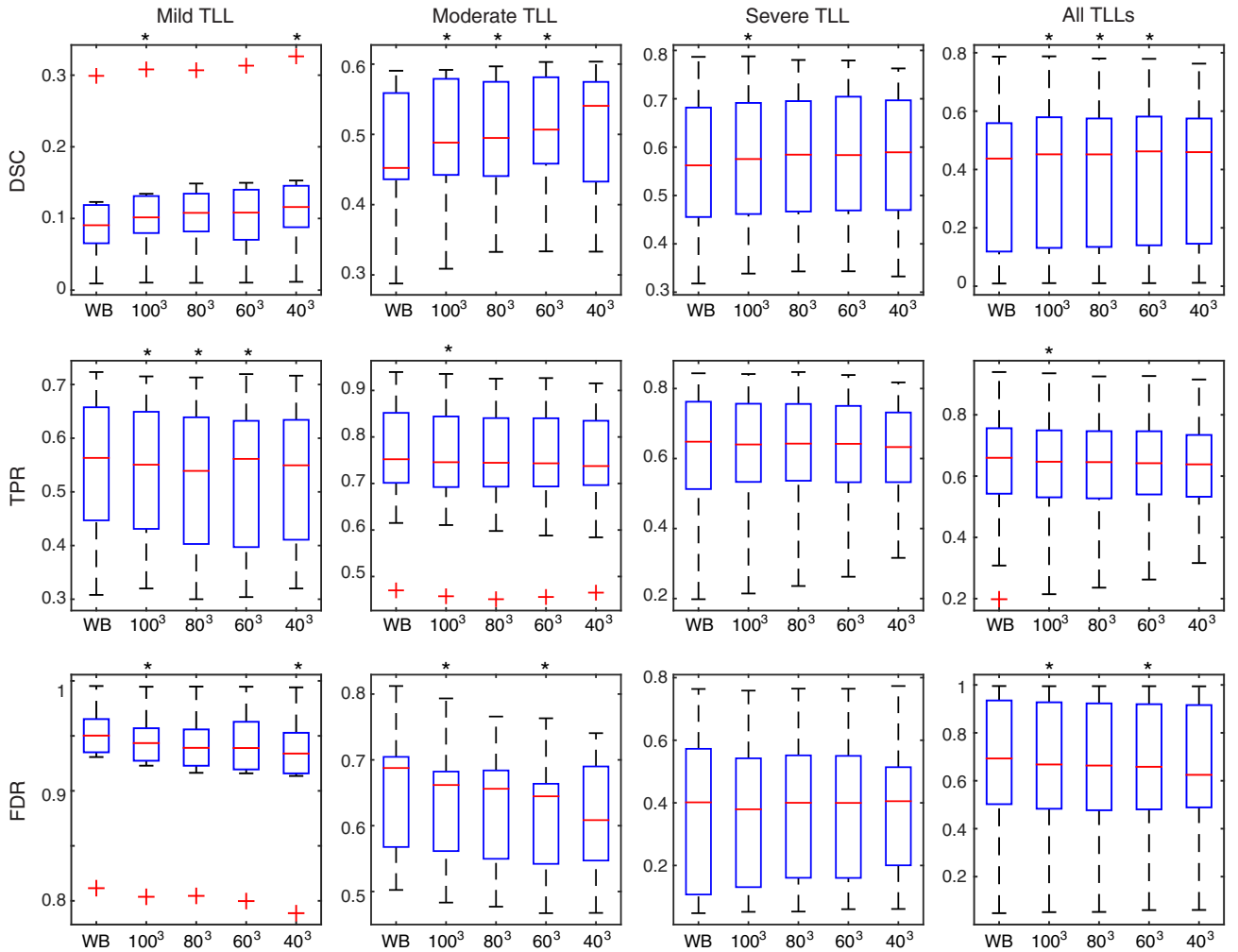

Fig. 7 Box-whiskers diagrams of DSC, TPR, and FDR for the proposed segmentation method with varying sizes of the subvolumes used for the intensity model estimation: the WB volume and subvolumes of sizes 100^3, 80^3, 60^3, and 40^3 mm^3. The methods were tested on datasets in three groups of different TLLs and over the whole dataset. The asterisks (*) indicate significance of the differences compared to the closest larger subvolume tested using Wilcoxon's signed rank test (at significance level of 0.05).

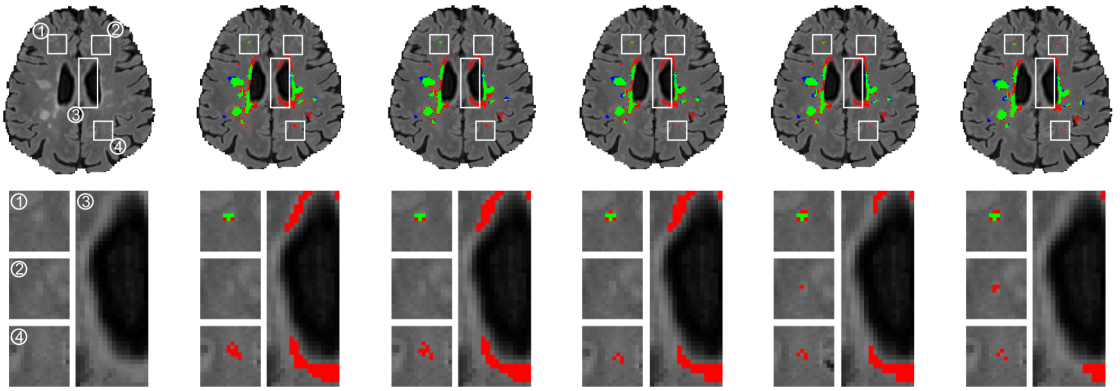

Fig. 8 Segmentation example by the proposed segmentation method with varying sizes of the subvolumes used for the intensity model estimation. From left to right: the FLAIR image at an axial slice, and the segmentation results using the WB volume, and subvolumes of sizes 100^3, 80^3, 60^3, and 40^3 mm^3. The lesion segmentations were labeled as true positives (in green), false negatives (in blue), or false positives (in red). Note, how the segmentation performance changes with decrease of the subvolume size: a small lesion in the region 1 remains preserved, while a lesion missed by the raters in the region 2 is detected, and false positives are gradually eliminated in the regions 3 and 4.

image datasets of MS patients. The lattice with a denser sampling step, that is more locally adaptive NABS model estimates, consistently improved the segmentation by WB NABS model in terms of DSC and false discovery rate (FDR) while the true positive rate (TPR) was not affected as much to deteriorate the performance. The main improvements in the segmentation performance obtained by the local robust estimation of GMMs are due to the improved MS lesion membership maps, which affect the overall segmentations of both NABS and MS lesions. The most prominent changes appear in the groups of mild and moderate TLL, which are considered as more difficult cases for automated segmentation compared to cases with severe TLL. Relatively low values in the group of the mild TLL are due to a known bias of voxel-wise performance measurements, such as the DSC, since the effect of a misclassified voxel is higher in such cases.

The impact of a denser lattice sampling step and thus the use of more locally adaptive NABS models is shown in Fig. 8, which indicates that increasingly local modeling improves MS lesion segmentation primarily by elimination of FPs and preservation of TPs.

Disclosures

Authors have no conflicts of interest to disclose.

Acknowledgments

This research was supported by Slovenian Research Agency under Grant Nos. P2-0232, J2-5473, J7-6781, J2-7211, J2-7118, and J2-7183.

References

1. C. H. Polman et al., "Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria," *Ann. Neurol.* **69**, 292–302 (2011).
2. H. Vrenken et al., "Recommendations to improve imaging and analysis of brain lesion load and atrophy in longitudinal studies of multiple sclerosis," *J. Neurol.* **260**, 2458–2471 (2013).
3. U. Vovk, F. Pernus, and B. Likar, "A review of methods for correction of intensity inhomogeneity in MRI," *IEEE Trans. Med. Imaging* **26**, 405–421 (2007).
4. J. V. Manjón et al., "Adaptive non-local means denoising of MR images with spatially varying noise levels," *J. Magn. Reson. Imaging* **31**, 192–203 (2010).
5. Y. Xiao et al., "Optimal Gaussian mixture models of tissue intensities in brain MRI of patients with multiple-sclerosis," *Lect. Notes Comput. Sci.* **6357**, 165–173 (2010).
6. K. Van Leemput et al., "Automated segmentation of multiple sclerosis lesions by model outlier detection," *IEEE Trans. Med. Imaging* **20**, 677–688 (2001).
7. D. García-Lorenzo et al., "Trimmed-likelihood estimation for focal lesions and tissue segmentation in multisequence MRI for multiple sclerosis," *IEEE Trans. Med. Imaging* **30**(8), 1455–1467 (2011).
8. B. Scherrer et al., "Distributed local MRF models for tissue and structure brain segmentation," *IEEE Trans. Med. Imaging* **28**, 1278–1295 (2009).
9. J. Rajapakse, J. Giedd, and J. Rapoport, "Statistical approach to segmentation of single-channel cerebral MR images," *IEEE Trans. Med. Imaging* **16**, 176–186 (1997).
10. D. W. Shattuck et al., "Magnetic resonance image tissue classification using a partial volume model," *NeuroImage* **13**, 856–876 (2001).
11. R. Harmouche et al., "Probabilistic multiple sclerosis lesion classification based on modeling regional intensity variability and local neighborhood information," *IEEE Trans. Biomed. Eng.* **62**, 1281–1292 (2015).
12. D. Biediger, C. Collet, and J.-P. Armspach, "Multiple sclerosis lesion detection with local multimodal Markovian analysis and cellular automata 'GrowCut'," *J. Comput. Surg.* **1**, 3–15 (2014).
13. A. Galimzianova et al., "Robust estimation of unbalanced mixture models on samples with outliers," *IEEE Trans. Pattern Anal. Mach. Intell.* **37**(11), 2273–2285 (2015).
14. S. Z. Li, *Markov Random Field Modeling in Image Analysis*, 3rd ed., Springer Publishing Company, Incorporated, London (2009).
15. D. García-Lorenzo et al., "Multiple sclerosis lesion segmentation using an automatic multimodal graph cuts," *Lect. Notes Comput. Sci.* **5762**, 584–591 (2009).
16. S. Klein et al., "elastix: a toolbox for intensity-based medical image registration," *IEEE Trans. Med. Imaging* **29**, 196–205 (2010).
17. A. Evans et al., "3D statistical neuroanatomical models from 305 MRI volumes," in *Proc. Nuclear Science Symp. and Medical Imaging Conf.*, Vol. **3**, pp. 1813–1817 (1993).
18. K. Wakimoto, "Stratified random sampling (II) estimation of the population covariance," *Ann. Inst. Stat. Math.* **23**, 327–337 (1971).
19. A. Galimzianova et al., "Stratified mixture modeling for segmentation of white-matter lesions in brain MR images," *NeuroImage* **124**, 1031–1043 (2016).
20. J. Iglesias et al., "Robust brain extraction across datasets and comparison with publicly available methods," *IEEE Trans. Med. Imaging* **30**, 1617–1634 (2011).

21. N. Tustison et al., "N4ITK: improved N3 bias correction," *IEEE Trans. Med. Imaging* **29**, 1310–1320 (2010).
22. Z. Lesjak et al., "A novel public MR image dataset of multiple sclerosis patients with lesion segmentations based on multi-rater consensus," *Neuroinformatics* (in press).
23. Z. Lesjak, A. Galimzianova, and Z. Spiclin, "BrainSeg3D - a free 3D image viewer and segmentation tool," Laboratory of Imaging Technologies, 2015, http://lit.fe.uni-lj.si/tools (22 October 2017).
24. Z. Lesjak et al., "Increased accuracy and reproducibility of MS lesion volume quantification by using publicly available BrainSeg3D image analysis software," *Mult. Scler. J.* **21**, 500–501 (2015).
25. N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst. Man Cybern.* **9**(1), 62–66 (1979).
26. N. Komodakis, G. Tziritas, and N. Paragios, "Performance vs computational efficiency for optimizing single and dynamic MRFs: setting the state of the art with primal-dual strategies," *Comput. Vis. Image Understanding* **112**, 14–29 (2008).
27. N. Shiee et al., "A topology-preserving approach to the segmentation of brain images with multiple sclerosis lesions," *NeuroImage* **49**, 1524–1535 (2010).
28. A. Carass et al., "TOADS-CRUISE Brain Segmentation Tools," NITRC, 2014, https://www.nitrc.org/projects/toads-cruise/ (22 October 2017).
29. A. Carass et al., "Longitudinal multiple sclerosis lesion segmentation: resource and challenge," *NeuroImage* **148**, 77–102 (2017).
30. T. Jerman et al., "Combining unsupervised and supervised methods for lesion segmentation," in *Int. Workshop on Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, pp. 45–56, Springer, Cham (2015).

Alfiia Galimzianova is a postdoctoral scholar at the Stanford University School of Medicine. She received her diploma degree in computer science from Kazan Federal University, Russia, in 2010 and her PhD in electrical engineering from the University of Ljubljana, Slovenia, in 2015. Her current research interests include autosegmentation, computer-aided diagnosis, and predictive models of diseases.

Žiga Lesjak is a research fellow at the Faculty of Electrical Engineering, the University of Ljubljana, Slovenia. He received his MS and PhD degrees in electrical engineering from the University of Ljubljana, Slovenia, in 2012, and 2017, respectively. His current interests include research in the field of magnetic resonance image analysis with focus on automatic change detection and integration of this research into commercial applications.

Daniel L. Rubin is an associate professor of biomedical data science, radiology, and medicine (Biomedical Informatics Research) at Stanford University. His NIH-funded research program focuses on quantitative imaging, integrating imaging with clinical and molecular data, and mining these big data to discover imaging phenotypes that can predict disease biology, define disease subtypes, and personalize treatment.

Boštjan Likar is a professor at the Faculty of Electrical Engineering, the University of Ljubljana, Slovenia. He is a member of the Laboratory of Imaging Technologies and his research interests include visual quality inspection, computer and machine vision systems, biomedical image processing, and hyperspectral imaging. He is a cofounder of a high-tech company Sensum, which supplies machine vision solutions for the pharmaceutical industry.

Franjo Pernuš is a professor at the Faculty of Electrical Engineering, the University of Ljubljana, Slovenia. He is the head of the Laboratory of Imaging Technologies and his research interests involve biomedical image processing and analysis, computer vision, and their application to various biomedical and industrial problems. He is a cofounder of a high-tech company Sensum, which supplies machine vision solutions for the pharmaceutical industry.

Žiga Špiclin is an assistant professor at the Faculty of Electrical Engineering, the University of Ljubljana, Slovenia. His research interests include the development of image registration, restoration, and reconstruction techniques for computer vision and biomedical applications.