A Semi-supervised Large Margin Algorithm for White Matter Hyperintensity Segmentation

Chen Qin[1], Ricardo Guerrero Moreno[1], Christopher Bowles[1], Christian Ledig[1], Philip Scheltens[2], Frederik Barkhof[2], Hanneke Rhodius-Meester[2], Betty Tijms[2], Afina W. Lemstra[2], Wiesje M. van der Flier[2], Ben Glocker[1], and Daniel Rueckert[1]

[1] Department of Computing, Imperial College London, UK
[2] Department of Neurology, VU University Medical Center, Amsterdam, the Netherlands

Abstract. Precise detection and quantification of white matter hyperintensities (WMH) is of great interest in studies of neurodegenerative diseases (NDs). In this work, we propose a novel semi-supervised large margin algorithm for the segmentation of WMH. The proposed algorithm optimizes a kernel based max-margin objective function which aims to maximize the margin averaged over inliers and outliers while exploiting a limited amount of available labelled data. We show that the learning problem can be formulated as a joint framework learning a classifier and a label assignment simultaneously, which can be solved efficiently by an iterative algorithm. We evaluate our method on a database of 280 brain Magnetic Resonance (MR) images from subjects that either suffered from subjective memory complaints or were diagnosed with NDs. The segmented WMH volumes correlate well with the standard clinical measurement (Fazekas score), and both the qualitative visualization results and quantitative correlation scores of the proposed algorithm outperform other well known methods for WMH segmentation.

1 Introduction

White matter hyperintensities (WMH) are areas of the brain in cerebral white matter (WM) that appear bright on T2-weighted fluid attenuated inversion recovery (FLAIR) magnetic resonance (MR) images due to localized, pathological changes in tissue composition [12]. WMH are commonly observed in elderly subjects and subjects with neurodegenerative diseases (NDs), such as vascular dementia (VaD), Alzheimer's disease (AD) and dementia with Lewy Bodies (DLB). Current research [2, 5] indicates that the WMH volume in subjects with dementia is significantly higher than that of a normal aging population, and the presence, severity and distribution of WMH also vary between different disorders. Clinically, the amount of WMH is usually characterized by the Fazekas score [3], which is useful in the assessment of subjects with possible dementia. However, such visual rating scales show poor sensitivity to clinical group differences and may also incur high intra- and inter-rater variability [2]. Thus, more reliable and precise methods for quantifying and analyzing WMH are still desirable.

Recently, several techniques that seek to automatically and precisely segment and quantify WMH have been put forward [2]. In the supervised setting, machine learning methods such as k-nearest neighbor (kNN) [1], support vector machines (SVM) [7] and random forests [5] have been applied to the problem of WMH segmentation. These approaches learn the characteristic features of lesions from training samples that have been manually segmented by an expert. Such supervised methods can achieve good performance, however, they rely on the manual segmentation which is costly, time consuming and inevitably contains some mislabelled training data. In contrast, unsupervised segmentation methods do not require labelled training data. Most approaches employ clustering techniques to group similar voxels, such as fuzzy C-means clustering [4] and EM-based algorithms [6]. A different type of approach considers lesions as outliers to normal tissues [11, 14]. Recently, a lesion growth algorithm [10] has been developed, which constructs a conservative lesion belief map with a pre-chosen threshold (τ), followed by the initial map being grown along voxels that appear hyperintense in the FLAIR image. However, such unsupervised approaches can not always produce satisfactory results in subjects with NDs, since WMH in those subjects are often small and irregular, and also heterogeneous within and across subjects [5].

In this work, we propose a semi-supervised large margin approach for WMH segmentation, which identifies WMH as outliers, i.e., patterns deviating from normal data. Specifically, our method optimizes a kernel based max-margin objective function formulated by both the limited labelled information and a large amount of unlabelled data. We show that the framework jointly learns a large margin classifier and a label assignment, which is solved by updating the classifier and the label indicator alternatingly. The main idea of the proposed approach is to tackle the uncertainty of unlabelled input data with the help of a small proportion of labelled ones, and to discover outliers by training a large margin classifier which maximizes the average margins of judged inliers and outliers. Instead of assuming that data is generated from a particular distribution as most of other outlier detection methods do [11, 14], which may not hold true for WMH segmentation, our method assumes that neighboring samples tend to have consistent classifications that are guided by available labelled data. Quantitative and qualitative results indicate that the proposed method outperforms the current well known methods on a large database of subjects with NDs.

2 Unsupervised One-Class Learning

Let $\mathcal{X} = \left\{\mathbf{x}_i \in \mathrm{R}^d\right\}_{i=1}^n$ denote a set of n unlabelled input samples, and y_i represent the corresponding soft label that assigns normal samples a positive value (c^+) while a negative value (c^-) for outliers. Additionally, let $\mathcal{H}$ be a reproducing kernel Hilbert space (RKHS) of the function: $f(\mathbf{x}) = \sum_{i=1}^n \kappa(\mathbf{x}, \mathbf{x}_i)\alpha_i$, with associated kernel κ as the functional base and the expansion coefficient α. The unsupervised one-class learning (UOCL) proposed by Liu et al. [9] is an unsupervised algorithm that uses a self-guided labelling procedure to discover potential

outliers in the data. This method aims to separate the positive samples from outliers by training a large margin classifier, which is obtained from minimizing the following objective function:

$$\begin{aligned} \min_{f\in\mathcal{H},\{y_i\}} & \sum_{i=1}^{n} \left(f(\mathbf{x}_i)-y_i\right)^2+\gamma_1\left\|f\right\|_{\mathcal{M}}^2-\frac{2\gamma_2}{n_+}\sum_{i,y_i>0} f(\mathbf{x}_i) \\ s.t. \quad & y_i\in\{c^+,c^-\},\ \forall i\in[1:n],\ 0<n_+=|\{i|y_i>0\}|<n. \end{aligned} \tag{1}$$

Here, the first term in function (1) uses the squared loss to make the classification function consistent with the label assignment. The second term is a manifold regularizer, which endows f with the smoothness along the intrinsic manifold structure $\mathcal{M}$ underlying the data. Here $\|f\|_{\mathcal{M}}^2$ can be formulated as $\mathbf{f}^{\mathrm{T}}\mathbf{L}\mathbf{f}$, in which $\mathbf{f}=\left[f(\mathbf{x}_1),\cdots,f(\mathbf{x}_n)\right]^{\mathrm{T}}$, and $\mathbf{L}$ is the graph Laplacian matrix. The last term represents the margin averaged over the judged positive samples, which aims to push the majority of the inliers as far away as possible from the decision boundary $f(\mathbf{x})=0$. The importance of all three terms are balanced by the trade-off parameters γ_1 and γ_2. For more details, please refer to [9].

This minimization problem is solved by an alternating optimization scheme, with the continuous function f and discrete label assignment y_i being optimized iteratively. The method has been shown to be robust to high outlier proportion, which is a highly desirable trait in WMH segmentation of subjects with NDs.

3 Semi-supervised Large Margin Algorithm

When it comes to the WMH segmentation, the classification results of UOCL are not always satisfactory. Since outliers originate from low-density samples and are later separated from high-density regions without guidance from labelled information, the UOCL method can produce many false positives when segmenting WMH, such as identifying edges and partial volume as outliers. To address this problem, we extend the UOCL method to a semi-supervised large margin algorithm (SSLM). A limited amount of labelled data is introduced to provide some guidance for unlabelled samples, with the aim of improving its performance over unsupervised methods as well as reducing the need of expensive labelled data required in fully supervised learning.

3.1 Learning Model

Following the notations defined in Section 2, we define L as a labelled data set and U as the unlabelled data set, which, in WMH segmentation case, represent sets of voxels with known and unknown labels respectively. The objective function of our proposed model is formulated as:

$$\begin{aligned} \min_{f\in\mathcal{H},\{y_i\}} & \sum_{\mathbf{x}_i\in U}\left(f(\mathbf{x}_i)-y_i\right)^2+\lambda\sum_{\mathbf{x}_j\in L}\left(f(\mathbf{x}_j)-y_j\right)^2+\gamma_1\left\|f\right\|_{\mathcal{M}}^2 \\ & -\frac{2\gamma_2}{n_+}\sum_{k,y_k>0} f(\mathbf{x}_k)+\frac{2\gamma_3}{n_-}\sum_{k,y_k<0} f(\mathbf{x}_k), \\ s.t. \quad & y_i\in\{c^+,c^-\},\ n_+=|\{i|y_i>0\}|,\ n_-=|\{i|y_i<0\}|, \end{aligned} \tag{2}$$

where $\mathbf{x}_k \in L \cup U$, $\lambda, \gamma_1, \gamma_2$ and γ_3 are trade-off parameters controlling the model, and n_+ and n_- are numbers of positive and negative samples respectively during the learning. In this model, we introduce a new term $\sum_{\mathbf{x}_j \in L} \left(f(\mathbf{x}_j) - y_j\right)^2$ that represents squared loss for labelled data. This enables the classification to be informed by the available labels, thereby allowing it to better discriminate between inliers and outliers. Additionally, motivated by [13], we also introduce a new term $\sum_{k,y_k<0} f(\mathbf{x}_k)/n_-$ into the objective function (2), which aims to maximize the average margin between the judged outliers and the decision boundary. The last two terms in objective function (2) work together to push the positive samples and outliers far away from the decision boundary, thus enabling these two groups of data to be far away from each other.

For a more concise notation, we further define the vectorial kernel mapping $\mathbf{k}(\mathbf{x}) = \left[\kappa(\mathbf{x}_i, \mathbf{x}), \cdots, \kappa(\mathbf{x}_n, \mathbf{x})\right]^{\mathrm{T}}$, and the kernel matrix $\mathbf{K} = \left[\kappa(\mathbf{x}_i, \mathbf{x}_j)\right]_{1 \le i,j \le n}$, so the target function can be expressed as $f(\mathbf{x}) = \alpha^{\mathrm{T}}\mathbf{k}(\mathbf{x})$ and $\mathbf{f} = \mathbf{K}\alpha$, in which $\alpha = \left[\alpha_1, \cdots, \alpha_n\right]^{\mathrm{T}} \in \mathrm{R}^n$. Then the objective function can be rewritten as

$$\begin{aligned} &\min_{\alpha,\mathbf{y}} \ \alpha^{\mathrm{T}}\mathbf{K}(\mathbf{\Lambda} + \gamma_1\mathbf{L})\mathbf{K}\alpha - 2\alpha^{\mathrm{T}}\mathbf{K}\mathbf{\Lambda}\mathbf{y} + \mathbf{y}^{\mathrm{T}}\mathbf{\Lambda}\mathbf{y} - 2\alpha^{\mathrm{T}}\mathbf{K}\widetilde{\mathbf{y}} \\ &s.t. \quad \mathbf{y} \in \{c^+, c^-\}^{n\times 1}, \ \mathbf{\Lambda} = diag(1, \ldots 1, \underbrace{\lambda, \ldots \lambda}_{\mathbf{x}_j \in L}, 1, \ldots 1), \\ &\qquad \widetilde{y_i} = \begin{cases} \frac{\gamma_2}{\|\mathbf{y}\|_+}, & y_i = c^+, \\ -\frac{\gamma_3}{\|\mathbf{y}\|_-}, & y_i = c^-, \end{cases} \end{aligned} \tag{3}$$

in which $\|\mathbf{y}\|_+ = n_+$ and $\|\mathbf{y}\|_- = n_-$, standing for the number of positive elements and negative elements in vector $\mathbf{y}$ respectively. In our method, we adopt the same soft label assignment for (c^+, c^-) as in [9], i.e. $(\sqrt{\frac{n_-}{n_+}}, -\sqrt{\frac{n_+}{n_-}})$.

3.2 Algorithm

Similar to the UOCL method, solving the proposed model involves a mixed optimization of a continuous variable α and a discrete variable $\mathbf{y}$. One key observation is that if one of the two components is fixed, the optimization problem becomes easy to solve. Here we propose a procedure that alternatingly optimizes α and $\mathbf{y}$ similar to the EM framework by updating α and $\mathbf{y}$ iteratively.

First, for a given label indicator $\mathbf{y}$, computing the optimal α is equivalent to minimization of the following sub-problem:

$$\min_{\alpha} \ Q(\alpha) := \alpha^{\mathrm{T}}\mathbf{K}(\mathbf{\Lambda} + \gamma_1\mathbf{L})\mathbf{K}\alpha - 2\alpha^{\mathrm{T}}\mathbf{K}\mathbf{\Lambda}\mathbf{y} - 2\alpha^{\mathrm{T}}\mathbf{K}\widetilde{\mathbf{y}}. \tag{4}$$

The gradient of the objective function (4) with respect to α is $\delta Q/\delta\alpha = 2\{[\mathbf{K}(\mathbf{\Lambda} + \gamma_1\mathbf{L})\mathbf{K}]\alpha - \mathbf{K}\mathbf{\Lambda}\mathbf{y} - \mathbf{K}\widetilde{\mathbf{y}}\}$. By using the gradient, problem (4) can be efficiently solved by the conjugate gradient descent method.

When α is fixed, we need to deal with the $\mathbf{y}$-subproblem, that is

$$\begin{aligned} &\max_{\mathbf{y}} \ H(\mathbf{y}) := \ 2\alpha^{\mathrm{T}}\mathbf{K}(\mathbf{\Lambda}\mathbf{y} + \widetilde{\mathbf{y}}) - \mathbf{y}^{\mathrm{T}}\mathbf{\Lambda}\mathbf{y} \\ &s.t. \quad \mathbf{y} \in \{c^+, c^-\}^{n\times 1}, \ \widetilde{y_i} = \begin{cases} \frac{\gamma_2}{\|\mathbf{y}\|_+}, & y_i = c^+, \\ -\frac{\gamma_3}{\|\mathbf{y}\|_-}, & y_i = c^-. \end{cases} \end{aligned} \tag{5}$$

Algorithm 1 SSLM

Input: The kernel and graph Laplacian matrices $\mathbf{K}$, $\mathbf{L}$, model parameters $\lambda, \gamma_1, \gamma_2, \gamma_3 > 0$, $\mathbf{\Lambda}$ and *maxiter*

Initialization

$\alpha_0 = \frac{1}{\sqrt{n}}$, $m_0 = \arg\max_m H(q(\mathbf{K}\alpha_0, m))$, $\mathbf{y}_0 = q(\mathbf{K}\alpha_0, m_0)$, $\widetilde{\mathbf{y}_0} = h(m_0, \mathbf{y}_0)$, $t = 0$;

repeat

Update α_{t+1} by optimizing function (4) using conjugate gradient descent method;

Update m_{t+1}: $m_{t+1} = \arg\max_m H(q(\mathbf{K}\alpha_{t+1}, m))$;

Update $\mathbf{y}_{t+1}$ and $\widetilde{\mathbf{y}}_{t+1}$: $\mathbf{y}_{t+1} = q(\mathbf{K}\alpha_{t+1}, m_{t+1})$, $\widetilde{\mathbf{y}}_{t+1} = h(m_{t+1}, \mathbf{y}_{t+1})$;

$t = t + 1$;

until convergence or $t > maxiter$

Output: expansion coefficient $\alpha^* = \alpha_t$ and the soft label assignment $\mathbf{y}^* = \mathbf{y}_t$.

Here a simpler case is shown to solve this discrete optimization problem. If an integer $m = \|\mathbf{y}\|_+$ is given, then $\mathbf{y}^{\mathrm{T}}\mathbf{\Lambda}\mathbf{y}$ and the soft label assignment for labelled data remain the same regardless of the label assignment for unlabelled data. Thus this problem reduces to the same one as in UOCL, i.e., to maximize $(\mathbf{K}\alpha)^{\mathrm{T}}(\mathbf{y}+\widetilde{\mathbf{y}})$ in the unlabelled data set. It has been shown in [9] that an optimal solution satisfies $y_i > 0$ if and only if f_i is among m largest elements of $\mathbf{f}$.

One optimal solution to the equation (5) can be simply obtained by sorting $\mathbf{f}$ for unlabelled data in a descending order. Then $y_i > 0$ is assigned to samples before and including the m_U-th element, while $y_i < 0$ to those after the m_U-th element. Here $m_U = m - m_L$, in which m_U and m_L stand for the number of positive samples in the unlabelled and labelled data sets respectively, with m_L a fixed number. Therefore, the solution to the subproblem (5) can be expressed as $\mathbf{y}^*(\alpha) = q(\mathbf{K}\alpha, m^*(\alpha))$, in which $m^*(\alpha) = \arg\max_m H(q(\mathbf{K}\alpha, m))$. Note that the known labels are kept unchanged when learning. For simplicity, we further define $\widetilde{\mathbf{y}}$ as a function of m and $\mathbf{y}$, i.e., $\widetilde{\mathbf{y}} = h(m, \mathbf{y})$. The summarization of this method is shown in Algorithm 1.

4 Results

Data used in the preparation of this work consisted of T1 and FLAIR MR images from 280 subjects acquired on a 3T MR scanner. The cohort included 53 subjects with subjective memory complaints (SMC), 155 subjects with probable AD, 34 subjects with fronto-temporal lobe dementia (FTD), 10 subjects with VaD, and 28 subjects with DLB. All images have been rated by an expert in terms of WMH using the Fazekas score.

Here, multichannel information (T1 and FLAIR MR images) is used to identify WMH. To do this, we first applied an automated brain segmentation tool [8] to the T1 scan to remove non-brain tissue and to extract a WM tissue probability map. All T1 and white matter tissue maps were registered to the FLAIR space. Additionally, bias correction and intensity normalization were also applied to both T1 and FLAIR images. The WMH segmentation was then performed for

voxels with WM probability larger than 0.1. For each voxel of interest, a feature vector was constructed with intensities of a 3×3 neighborhood from both FLAIR and coregistered T1 images. Here we used 2D patches as FLAIR MR images commonly have slices with low resolution in the through-plane direction.

We have evaluated the performance of the proposed method against the lesion growth algorithm (LGA) [10] and lesion predict algorithm (LPA) as implemented in Lesion Segmentation Toolbox (LST), which is a widely used tool for WM lesion segmentation. LPA is a supervised method which was trained by a logistic regression model with the data of 53 MS patients, and the pre-chosen threshold τ in LGA was set to 0.3 as suggested by [10]. Preliminary experiments showed that UOCL method failed on images with fewer lesions and thus its results were omitted. For the proposed method, labelled data was automatically and conservatively determined based on the distribution of the WM intensities on each subject, which took up a proportion of around 25%. Note that the number of labelled voxels from normal tissues is much more than that of labelled WMH. Gaussian kernel $\kappa(\mathbf{x}, \mathbf{x}') = \exp(-\left\|\mathbf{x} - \mathbf{x}'\right\|^2 / 2\sigma^2)$ was used in the classification function in which $\sigma^2 = \sum_{i,j=1}^{n} \left\|\mathbf{x}_i - \mathbf{x}_j\right\|^2 / n^2$, and the model parameters $\lambda, \gamma_1, \gamma_2$ and γ_3 were determined empirically based on a subset of the data. A comparison of the WMH segmentation visualization results is shown in Fig.1.

For a more quantitative assessment of the performance of our method, we have computed the correlation between the segmented WMH volumes and their Fazekas scores. The results are shown in Table 1. Here *Corr* denotes the correlation coefficient between the Fazekas score and the percentage of WMH volume relative to the WM volume, and $Corr_{(1,2,3)}$ is the same measure but excluding subjects with Fazekas score of 0. From Table 1, it can be seen that our approach achieved higher correlation score than LGA and LPA on both *Corr* and $Corr_{(1,2,3)}$, which indicates that the segmentation results of the proposed model are more consistent with the the standard clinical measurement. Furthermore, the proposed method can better discriminate WMH voxels from non-WMH voxels on subjects with relatively higher volume of lesions and thus can give a higher correlation value when excluding subjects with lower Fazekas scores. This can be further explained by Fig.2, which shows the distribution of the segmentation results with Fazekas scores. It can also be seen that the proposed method is able to better classify subjects according to Fazekas score. Overall, from both the visualization results and the correlation score, it can be concluded that the proposed method outperforms both LGA and LPA and has promising results.

5 Conclusion

In this work, we proposed a novel semi-supervised large margin algorithm. The proposed model can better discover suspicious outliers under the supervision of a limited amount of available labelled data. We have shown that the framework jointly learns a large margin classifier and a label assignment, which can be solved effectively by an iterative algorithm. Experiments for WMH segmentation were implemented on a database of 280 MR images from subjects with SMC or NDs.

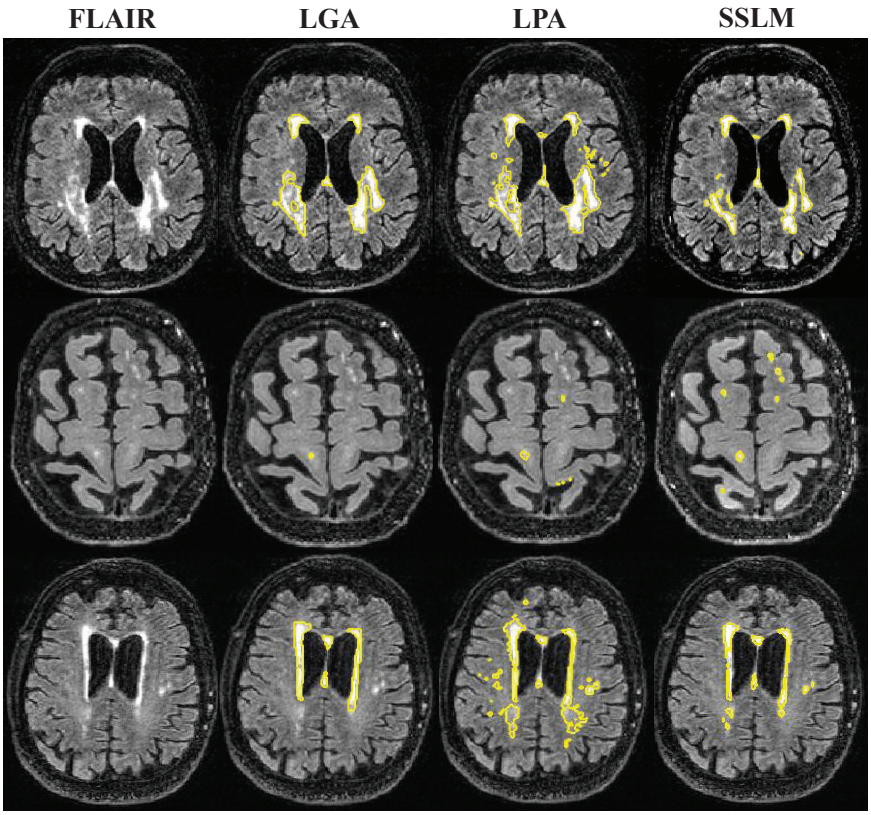

Fig. 1. Example WMH segmentation results compared with LGA and LPA on three different subjects with Fazekas score 1, 2 and 3 (from bottom to top) respectively.

Table 1. Correlation between WMH segmentation results and Fazekas scores

Method	$Corr$	$Corr_{(1,2,3)}$
LGA	0.5902	0.6532
LPA	0.6977	0.7661
SSLM	**0.7540**	**0.8333**

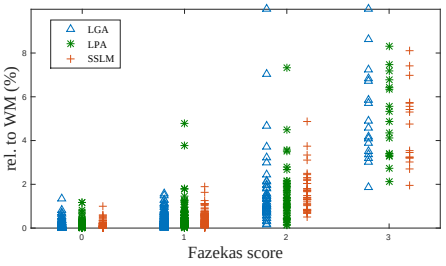

Fig. 2. Results of (WMH volume/WM volume) corresponding to Fazekas scores.

Encouraging experimental results were obtained on the qualitative visualization results and the quantitative correlation scores, showing the effectiveness and competitiveness of the proposed model against other methods.

References

1. Anbeek, P., Vincken, K.L., van Osch, M.J., Bisschops, R.H., van der Grond, J.: Probabilistic segmentation of white matter lesions in MR imaging. NeuroImage 21(3), 1037–1044 (2004)
2. Caligiuri, M.E., Perrotta, P., Augimeri, A., Rocca, F., Quattrone, A., Cherubini, A.: Automatic detection of white matter hyperintensities in healthy aging and pathology using magnetic resonance imaging: A review. Neuroinformatics 13(3), 261–276 (2015)
3. Fazekas, F., Chawluk, J.B., Alavi, A., Hurtig, H.I., Zimmerman, R.A.: MR signal abnormalities at 1.5 T in Alzheimer's dementia and normal aging. American Journal of Neuroradiology 8(3), 421–426 (1987)
4. Gibson, E., Gao, F., Black, S.E., Lobaugh, N.J.: Automatic segmentation of white matter hyperintensities in the elderly using FLAIR images at 3T. Journal of Magnetic Resonance Imaging 31(6), 1311–1322 (2010)
5. Ithapu, V., Singh, V., Lindner, C., Austin, B.P., Hinrichs, C., Carlsson, C.M., Bendlin, B.B., Johnson, S.C.: Extracting and summarizing white matter hyperintensities using supervised segmentation methods in Alzheimer's disease risk and aging studies. Human brain mapping 35(8), 4219–4235 (2014)
6. Kikinis, R., Guttmann, C.R., Metcalf, D., Wells, W.M., Ettinger, G.J., Weiner, H.L., Jolesz, F.A.: Quantitative follow-up of patients with multiple sclerosis using MRI: technical aspects. Journal of Magnetic Resonance Imaging 9(4), 519–530 (1999)
7. Lao, Z., Shen, D., Liu, D., Jawad, A.F., Melhem, E.R., Launer, L.J., Bryan, R.N., Davatzikos, C.: Computer-assisted segmentation of white matter lesions in 3D MR images using support vector machine. Academic radiology 15(3), 300–313 (2008)
8. Ledig, C., Heckemann, R.A., Hammers, A., Lopez, J.C., Newcombe, V.F., Makropoulos, A., Lötjönen, J., Menon, D.K., Rueckert, D.: Robust whole-brain segmentation: application to traumatic brain injury. Medical image analysis 21(1), 40–58 (2015)
9. Liu, W., Hua, G., Smith, J.: Unsupervised one-class learning for automatic outlier removal. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 3826–3833 (2014)
10. Schmidt, P., Gaser, C., Arsic, M., Buck, D., Förschler, A., Berthele, A., Hoshi, M., Ilg, R., Schmid, V.J., Zimmer, C., et al.: An automated tool for detection of FLAIR-hyperintense white-matter lesions in multiple sclerosis. Neuroimage 59(4), 3774–3783 (2012)
11. Van Leemput, K., Maes, F., Vandermeulen, D., Colchester, A., Suetens, P.: Automated segmentation of multiple sclerosis lesions by model outlier detection. IEEE Transactions on Medical Imaging 20(8), 677–688 (2001)
12. Wardlaw, J.M., Smith, E.E., Biessels, G.J., Cordonnier, C., Fazekas, F., Frayne, R., Lindley, R.I., T O'Brien, J., Barkhof, F., Benavente, O.R., et al.: Neuroimaging standards for research into small vessel disease and its contribution to aging and neurodegeneration. The Lancet Neurology 12(8), 822–838 (2013)
13. Wu, M., Ye, J.: A small sphere and large margin approach for novelty detection using training data with outliers. IEEE Transactions on Pattern Analysis and Machine Intelligence 31(11), 2088–2092 (2009)
14. Yang, F., Shan, Z.Y., Kruggel, F.: White matter lesion segmentation based on feature joint occurrence probability and $\chi 2$ random field theory from magnetic resonance (MR) images. Pattern Recognition Letters 31(9), 781–790 (2010)