NeuroImage: Clinical 22 (2019) 101772

Contents lists available at ScienceDirect

NeuroImage: Clinical

journal homepage: www.elsevier.com/locate/ynicl

Multi-atlas based detection and localization (MADL) for location-dependent quantification of white matter hyperintensities

Dan Wu[a,b,*], Marilyn Albert[c], Anja Soldan[c], Corinne Pettigrew[c], Kenichi Oishi[b], Yusuke Tomogane[b], Chenfei Ye[b], Ting Ma[b], Michael I. Miller[d,e,f], Susumu Mori[b,g]

[a] Key Laboratory for Biomedical Engineering of Ministry of Education, College of Biomedical Engineering & Instrument Science, Zhejiang University, Hangzhou, China
[b] Department of Radiology and Radiological Science, Johns Hopkins University School of Medicine, Baltimore, MD, USA
[c] Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, MD, USA
[d] Department of Biomedicine Engineering, Johns Hopkins University, Baltimore, MD, USA
[e] Center of Imaging Science, Johns Hopkins University, Baltimore, MD, USA
[f] Institute for Computational Medicine, Johns Hopkins University, Baltimore, MD, USA
[g] F.M. Kirby Research Center for Functional Brain Imaging, Kennedy Krieger Institute, Baltimore, MD, USA

ARTICLE INFO

Keywords:
White matter hyperintensities
FLAIR
Multi-atlas
Segmentation
Mild cognitive disorder

ABSTRACT

The extent and spatial location of white matter hyperintensities (WMH) on brain MRI may be relevant to the development of cognitive decline in older persons. Here, we introduce a new method, known as the Multi-atlas based Detection and Localization (MADL), to evaluate WMH on fluid-attenuated inversion recovery (FLAIR) data. This method simultaneously parcellates the whole brain into 143 structures and labels hyperintense areas within each WM structure. First, a multi-atlas library was established with FLAIR data of normal elderly brains; and then a multi-atlas fusion algorithm was developed by which voxels with locally abnormal intensities were detected as WMH. At the same time, brain segmentation maps were generated from the multi-atlas fusion process to determine the anatomical location of WMH. Areas identified using the MADL method agreed well with manual delineation, with an interclass correlation of 0.97 and similarity index (SI) between 0.55 and 0.72, depending on the total WMH load. Performance was compared to other state-of-the-art WMH detection methods, such as BIANCA and LST. MADL-based analyses of WMH in an older population revealed a significant association between age and WMH load in deep WM but not subcortical WM. The findings also suggested increased WMH load in selective brain regions in subjects with mild cognitive impairment compared to controls, including the inferior deep WM and occipital subcortical WM. The proposed MADL approach may facilitate location-dependent characterization of WMH in older individuals with memory impairment.

1. Introduction

White matter hyperintensities (WMH) that appear on T2-weighted or fluid attenuated inversion recovery (FLAIR) magnetic resonance imaging (MRI) are a common radiological feature. WMH are primarily thought to reflect the degree and distribution of small vessel disease (Wardlaw et al., 2013), and they are increasingly common with advancing age (Debette and Markus, 2010; Gorelick et al., 2011; Prins and Scheltens, 2015). Recent findings suggest they may also be one of the core features of Alzheimer's disease (AD) (Brickman, 2013; Lee et al., 2016), in addition to gray matter atrophy (Vemuri and Jack, 2010).

Several studies have suggested that it might be important to consider the spatial distribution of WMH when evaluating individuals with memory impairment. For example, periventricular WMH (PVWMH) are more strongly associated with cognitive performance than deep WMH (DWMH) (de Groot et al., 2002; DeCarli et al., 2005), and posterior WMH was shown to play an important role in the development of AD (Lee et al., 2016; Yoshita et al., 2006). Hypothesis-driven investigation of WMH in other brain regions were also reported (Biesbroek et al., 2013; Brickman et al., 2012; Brickman et al., 2015; Murray et al., 2010; Wu et al., 2006). These findings suggest the importance of developing tools that not only measure the total WMH load but also systematically evaluate the WMH distribution, e.g., WMH load in different lobular divisions and various subcortical structures.

Fully automated WMH detection algorithms have been developed over the past decade, including various forms of intensity-based thresholding methods (Admiraal-Behloul et al., 2005; Jack et al., 2001; Ji et al., 2013; Ong et al., 2012; Simoes et al., 2013; Yoo et al., 2014), clustering approaches (Ithapu et al., 2014; Lao et al., 2008; Seghier et al., 2008), outlier analysis methods (Maldjian et al., 2013; Ong et al., 2012; Van Leemput et al., 2001; Yang et al., 2010), morphological operations

* Corresponding author at: Zhejiang University, Room 525, Zhou Yiqin Building, Yuquan Campus, Hangzhou 310027, China.
E-mail address: danwu.bme@zju.edu.cn (D. Wu).

https://doi.org/10.1016/j.nicl.2019.101772
Received 18 November 2018; Received in revised form 5 February 2019; Accepted 10 March 2019
Available online 13 March 2019

(Beare et al., 2009; Shi et al., 2013), Bayesian approaches (Herskovits et al., 2008; Ithapu et al., 2014), and more recently, deep learning methods (Ghafoorian et al., 2016; Ghafoorian et al., 2017; Jin et al., 2018; Moeskops et al., 2018; Roa-Barco et al., 2018). Some of these methods use single FLAIR contrast, while others employ multiple modalities including T1, T2, proton density, and even diffusion tensor imaging data. The detection accuracies of these methods depend not only on the algorithms itself but also the data under investigation, e.g., small lesions are typically more difficult to detect compared to large lesions (Caligiuri et al., 2015). The aforementioned methods typically generate a measure of whole-brain WMH load, but few automatically provide region-specific measures for systematic evaluation of WMH distribution.

In this paper, we describe a new computational framework that simultaneously parcellates the brain and identifies WMH in individual parcels. The new method, known as Multi-atlas based Detection and Localization (MADL), uses a multi-atlas likelihood fusion approach to achieve brain segmentation and WMH detection, based on a FLAIR multi-atlas library. We first tested algorithm performance by comparing the detection results with manually defined WMH labels. We then investigated if location-specific WMH quantification would provide useful biological or clinical information by examining (i) the association between age and WMH load in different WM structures, and (ii) location-dependent WMH load in normal elderly subjects and individuals with mild cognitive impairment (MCI).

2. Methods and materials

2.1. Dataset

The participants are members of a cohort of older individuals who had been well characterized clinically and cognitively, as part of an ongoing project, known as the BIOCARD study (Albert et al., 2014; Soldan et al., 2016). The current analyses are cross-sectional, based on FLAIR data from 135 subjects scanned between 2015 and 2016. Image acquisition in this cohort is ongoing. All subjects provided informed consent in accordance with the requirements of the Johns Hopkins Medicine Institutional Review Board.

As part of the ongoing study, each participant receives a consensus clinical diagnosis for the visit at which the MRI is obtained, based on the clinical and cognitive assessments at that visit, using standard procedures. For the scans used in the current analyses, participants received a diagnosis of either cognitively normal (n = 113) or MCI (n = 22), based on the National Institute on Aging/Alzheimer's Association (NIA/AA) research diagnostic criteria (Albert et al., 2011). FLAIR data from 15 cognitively normal individuals were used as atlases, and the other 120 were used for algorithm evaluation and subsequent analyses. Basic demographic and clinical characteristics of the participants are summarized in Table 1.

MRI scans were obtained at the Kirby Imaging Center at the Kennedy Krieger Institute on a Philips Achieva 3.0 T scanner. FLAIR data were acquired with a multi-slice fast spin-echo sequence with inversion recovery pulses at inversion time (TI)/echo time (TE)/repetition time (TR) = 2800/100/11,000 ms, field-of-view (FOV) = 256 × 256 mm, in-plane resolution = 1 × 1 mm, 69 slices with slice-thickness of 2 mm. T1-weighted images were acquired with a 3D magnetization-prepared rapid gradient-echo (MPRAGE) sequence with TI/TE/TR = 800/3/7 ms, flip angle of 8°, FOV of 256 × 256 × 204 mm, and resolution of 1 × 1 × 1.2 mm.

Table 1
Demographic information of the BIOCARD participants used in this study, including those used as atlases and those used as testing data.

Participant characteristics at image acquisition	Data used as atlases	Data used in algorithm evaluation and diagnostic analysis	
Number of participants (N)	15 (all Normal)	Normal 98	MCI 22
Age in years (mean ± SD)	70.1 ± 8.3	70.0 ± 8.5	69.9 ± 8.7
Gender (% female)	73%	64%	59%
MMSE (mean ± SD)	29.6 ± 0.8	29.5 ± 0.8	27.6 ± 2.0*

* $p < .001$ difference in MMSE between normal and MCI groups.

2.2. FLAIR multi-atlas generation

A FLAIR multi-atlas library was created with FLAIR data from cognitively normal individuals who had minimal WMH (< 1.8 ml based on manual delineation). Seventeen images qualified for this criterion, but two had slight image artifacts, and therefore 15 images were chosen as atlases. The demographic and basic clinical information of the atlas data matched with the test data (Table 1), and the 15 atlases represented a range of anatomy from minimal to moderate degrees of brain atrophy (Fig. 1A).

The 15 FLAIR images were first registered to T1-weighted images of the same brains in Montreal Neurological Institute (MNI) coordinates, by maximizing the mutual information between the FLAIR and T1-weighted images using the SPM package in Matlab (mathworks.com). The T1-weighted images were then segmented into 283 regions of interest (ROIs) (Wu et al., 2016) using a multi-atlas segmentation pipeline established by the investigators (Tang et al., 2013)(Wu et al., 2016). This way, the coregistered FLAIR data were segmented into the same 283 ROIs. For the purposes of this study, we manually edited the segmentations on FLAIR images, and regrouped the finest level of ROIs based on their ontological relationships. In the end, 143 ROIs were defined in FLAIR atlases. Of these, there were 24 WM ROIs (Fig. 1B) were used in the following analyses, and the rest were ROIs for gray matter (GM), cerebrospinal fluid (CSF) and non-brain tissue. The WM ROIs were included in the analyses below (see Fig. 2A).

2.3. MADL for WMH detection

To identify WMH, the MADL algorithm takes advantage of the fact that WMH voxels have abnormal intensities that do not comply with the local intensity profiles of the ROIs where they reside, resulting in low posterior probabilities in the multi-atlas fusion process. The algorithm flowchart is depicted in Fig. 2.

1) Global inhomogeneity correction on the target image (image to be segmented) was performed using N4 bias correction (Tustison et al., 2010), followed by histogram matching (Coltuc et al., 2006) between the target and atlas images. Atlas images were transformed to the target image first through affine registration and then non-linear transformation, using a method known as large deformation diffeomorphic metric mapping (LDDMM) (Christensen et al., 1996; Grenander and Miller, 1998).
2) Atlas-weighting and fusion were performed based on a multi-atlas likelihood fusion method, in which voxelwise posterior probabilities were derived through.

$$\hat{p}\,(l \mid x, I_T) = \sum_{i=1}^{N} w_i(l) \bullet p\,(l \mid x, I_A^i) \tag{1}$$

where $\hat{p}\,(l \mid x, I_T)$ is the posterior probability of voxel x in target image I_T being assigned to label l ($l \in [1, \cdots L]$); ${I_A}^i$ ($i \in [1, \cdots N]$, N being the number of atlases) are the warped atlas images; $w_i(l)$ is the atlas-weighting term determined iteratively by the spatial matching between the atlas label and the target label derived in the previous iteration; and $p(l|x, {I_A}^i)$ is the prior likelihood determined by the Gaussian probability density of voxel intensity at location x of the target image with respect to the intensity profile of label l in the atlas image ${I_A}^i$ (Tang et al., 2013). Voxels with abnormal intensities will give low prior likelihood with respect to the corresponding label.

3) Anatomical labels are obtained using Bayes maximum a posteriori (MAP) estimation:

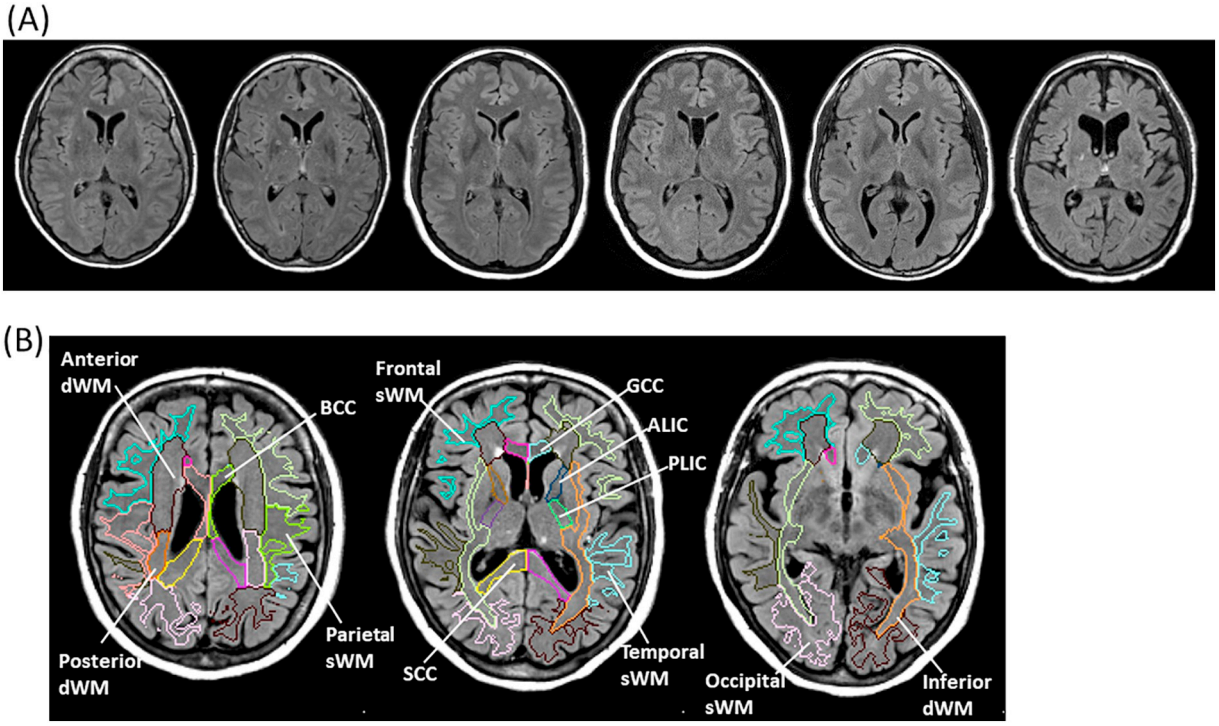

Fig. 1. The FLAIR multi-atlas library. (A) Images of 6 of the 15 FLAIR atlases that represented a range of anatomy from to minimal to moderate degrees of brain atrophy. (B) Major WM structures defined in the FLAIR atlases. Abbreviations: dWM—deep white matter; sWM—subcortical white matter; BCC—body of corpus callosum; GCC—genu of corpus callosum; SCC—splenium of corpus callosum; ALIC—anterior limb of internal capsule; PLIC—posterior limb of internal capsule.

$$L_T(x) = \underset{l\in[1,\ldots,L]}{\operatorname{argmax}} \hat{p}\,(l \mid x, I_T) \tag{2}$$

where L_T is the final label image. At the same time, the maximum posterior probability (*MPP*)— $\hat{p}\,(L_T \mid x, I_T)$ is obtained at each voxel.

4) WMH voxels are identified as voxels with low *MPP* values below a threshold. The threshold was empirically optimized based on comparison with manually delineated WMH labels. The effect of the choice of threshold was evaluated.

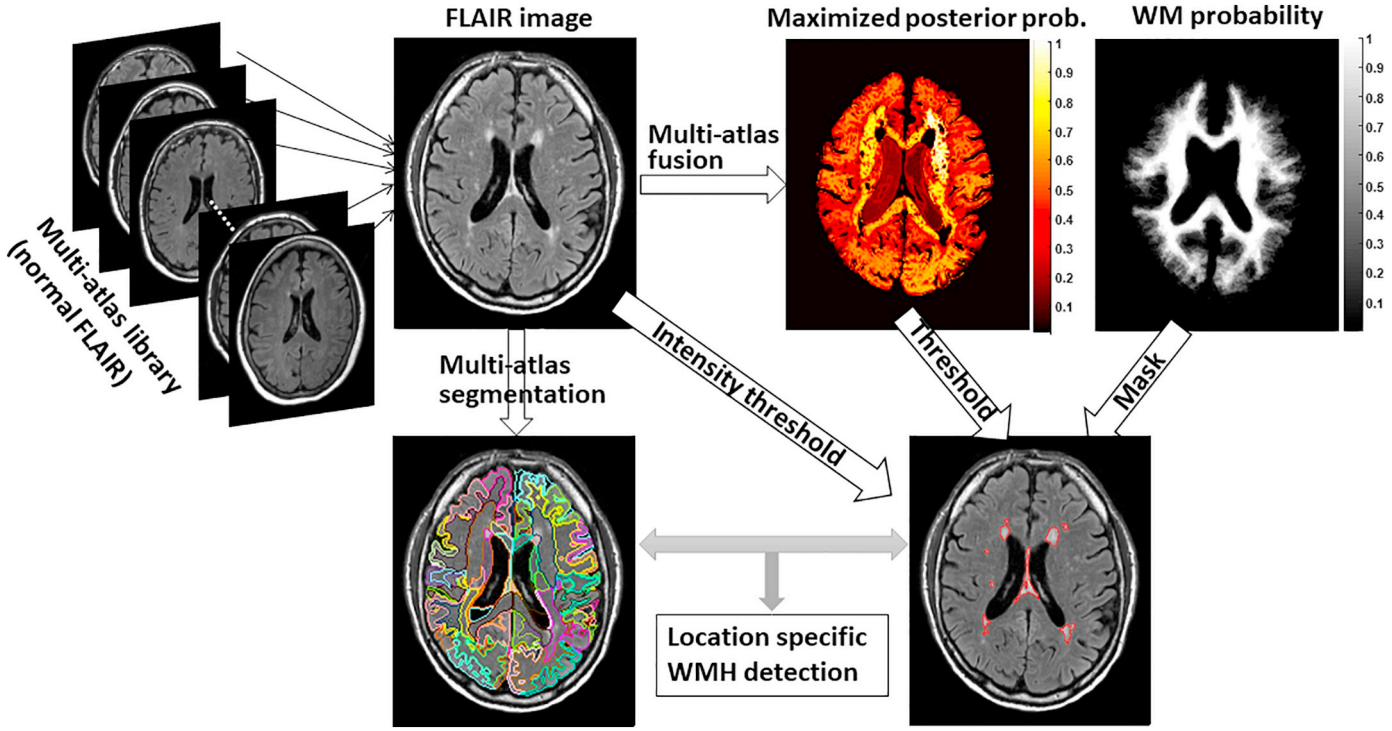

Fig. 2. Flowchart of the MADL pipeline. A multi-atlas library of FLAIR images is used to segment the brain and generate the posterior probability map based on a multi-atlas likelihood fusion algorithm. WMH voxels are detected if the maximum posterior probabilities are below a threshold, within a WM mask. Normalized image intensity was applied to exclude voxels with low intensities. Regional WMH load is obtained based on the simultaneously generated segmentation map and WMH label.

5) Post-processing. Several steps are taken to reduce the false positive detection.

i) A WM mask is generated using a simplified parcellation map (WM, GM, CSF, lateral ventricle, etc.) (Ma et al., 2015), using the same multi-atlas segmentation framework. Voxels outside the WM mask are removed.
ii) Because the MADL algorithm detects local intensity abnormalities, voxels with abnormally low intensities are also detected, e.g., CSF voxels (dark on FLAIR) that are encapsulated in cortical ROIs. Therefore, we applied an intensity threshold to exclude dark voxels lower than 1.5 standard deviations below the mean ROI intensity.
iii) After the above two steps, WMH clusters with volume below 50 mm^3 are removed.

2.4. Algorithm evaluation

1) WMH detection accuracy was evaluated based on manually delineated WMH labels by a board certified neurosurgeon (Y.T.), who is experienced in brain MRI analysis. The delineation was performed in ROIEditor (mristudio.org) using manually selected seeds, followed by region growing and manual editing. The following evaluation metrics were used.

i) The Dice similarity index (DSI, $\frac{2TP}{2TP+FP+FN}$), false-positive rate (FPR, $\frac{FP}{TP+FP}$), and false-negative rate (FNR, $\frac{FN}{TP+FN}$) (Griffanti et al., 2016) were calculated between the WMH labels detected with MADL and the manual labels in each subject. The DSI, FPR, and FNR were evaluated in three groups of individuals with low ($<$ 5 ml, $n = 75$), median (5–10 ml, $n = 34$), and high ($>$ 10 ml, $n = 15$) WMH load in the entire brain. Representative WMH maps of low, median, and high load brains are shown in Fig. 3C.
ii) WMH volume correspondence between the MADL and manual results was evaluated using interclass correlation (ICC) with the consistency agreement definition (McGraw and Wong, 1996).
iii) Receiver-operating characteristic (ROC) curves were calculated using voxelwise false-positive detection rate versus true-positive rate, based on which the area under the ROC curve (AUC) was calculated in each subject.

2) We compared the MADL outputs with two state-of-the-art WMH detection algorithms: (a) the Brain Intensity Abnormality Classification Algorithm (BIANCA) (Griffanti et al., 2016), which employs user-provided training data to classify abnormal intensities with a *k*-nearest neighbor algorithm and is implemented in FSL (fsl.fmrib.ox.ac.uk/fsl/fslwiki/BIANCA), and (b) the Lesion Segmentation Toolbox (LST, version 2.0.15), as implemented in SPM12 (https://www.applied-statistics.de/lst.html). We used the lesion prediction algorithm (LPA) (Schmidt, 2017) in LST, which was trained by a logistic regression model with internal training data.

For BIANCA, WMH probability maps were generated in a leave-one-out fashion (recursively, 119 of the 120 FLAIR data with manually delineated WMH were used as training data for the remaining test image). An empirical threshold of 0.9 and a cluster size of 10 was used to obtain WMH labels from the probability maps, as suggested by (Griffanti et al., 2016). For LST, WMH probability maps were generated with its internal training data, and we used a recommended probability threshold of 0.5 to obtain WMH labels. DSI, FPR, FNR, and ICC were used to evaluate the BIANCA and LST results.

2.5. Application of algorithm to clinical data

To examine the utility of MADL based WMH identification, we investigated the relationship between local WMH load and (i) the participant age and (ii) the clinical diagnosis. All statistical analyses were performed in R (www.r-project.org). Significance was detected at a 5% false discovery rate (FDR) after correcting for multiple comparisons (Benjamini and Hochberg, 1995). WMH ROIs used in this analysis are listed in Table 3. Note that the left and right sides of the corpus callosum (CC) were combined in the statistical analyses, resulting in three ROIs for the CC—the genu, body, and splenium parts (GCC, BCC, and SCC, respectively). In addition, the anterior and posterior limbs of the internal capsule were not included in the analysis since they did not show WMH in most subjects.

1) Relationships between age, WMH load, and WM volumes of individual ROIs were assessed with linear regressions, including (a) regressions between age and local WMH load (log-transformed), and (b) regressions between age and WM structural volumes. Regression analyses were adjusted for clinical diagnosis and sex, and corrected for multiple comparisons.
2) Group differences in local WMH load between the cognitively normal and MCI participants were assessed by analysis of covariance (ANCOVA), adjusted for age and sex and corrected for multiple comparison. Due to the unbalanced sample size between normal and MCI groups ($n = 98$ versus $n = 22$), we used type ANCOVA with type II sum of squares (Langsrud, 2003).

3. Results

3.1. Performance of the MADL pipeline

3.1.1. Effect of thresholding

We evaluated the detection accuracy at different *MPP* thresholds, and voxels with *MPP* below the threshold were detected as WMH. DSI and ICC were calculated in the low ($<$ 5 ml), median (5–10 ml), and high ($>$ 10 ml) groups based on the total WMH load (Fig. 3C). The effect of thresholding was relatively small in the range of 0.01–0.05 (Fig. 4A), except that the DSI of the low WMH group slightly decreased as the threshold increased above 0.03. An *MPP* threshold of 0.02 was used in the following analysis.

3.1.2. Comparison with manual detection

The whole-brain WMH load (in unit of ml) detected with the MADL pipeline showed a high level of agreement with manual delineation, with an ICC of 0.97 across 120 subjects (Fig. 3A). Majority of the participants in this study demonstrated a low-to-median amount of WMH, and the ICC was 0.89 in a sub-population with WMH $<$ 20 ml (dashed area in Fig. 3A). ROC curve of the voxelwise detection accuracy is shown in Fig. 3B, with an overall AUC of 0.89 $\pm$ 0.05. AUCs in the low, median, and high WMH groups were 0.89 $\pm$ 0.05, 0.89 $\pm$ 0.03, and 0.85 $\pm$ 0.04, respectively. DSI, FPR, and FNR measurements in low, median, and high WMH groups are reported in Table 2. DSI increased and FPR/FNR decreased as the total WMH load increased. The overall DSI, FPR, and FNR of the entire study population were 0.62 $\pm$ 0.09, 0.35 $\pm$ 0.14, and 0.37 $\pm$ 0.12, respectively, given that 88% of the study population were in the low and median groups.

3.1.3. Comparison with other algorithms

We compared the performance of the MADL, BIANCA, and LST, based on DSI, FPR, and FNR in the low, median, and high WMH groups and ICC of the entire population. Table 2 shows that the highest ICC was obtained using MADL, and the DSI were similar among the three methods, using pairwise *t*-tests (all $p > .05$). BIANCA showed the highest FPR and lowest FNR, and LST showed the lowest FPR and highest FNR, among the three methods ($p < .01$ by paired *t*-tests between MADL and BIANCA and between MADL and LST).

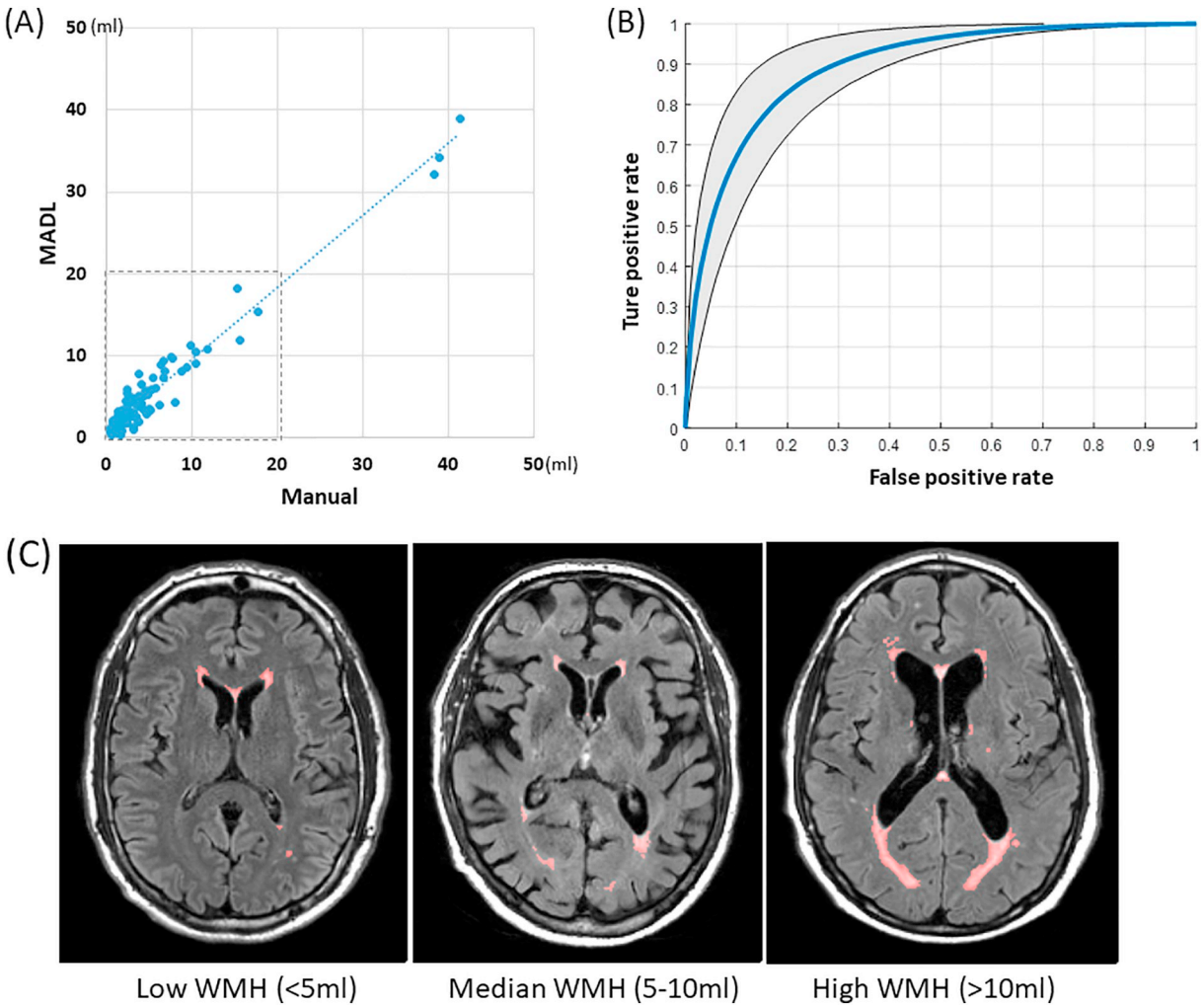

Fig. 3. Evaluation of the MADL method. (A) Whole-brain WMH load detected by MADL correlated well with manually delineated results, with an interclass correlation of 0.97. The dashed black box includes a subset of data that had total WMH load < 20 ml and an ICC of 0.89. (B) Voxelwise ROC analysis was performed for each subject, and the mean and standard deviation is presented. (C) Representative FLAIR images with low, median, and high amounts of WMH load from the study population. Highlighted areas show the WMH labels detected by MADL.

3.2. Relationship between location-dependent WMH and age

Whole-brain WMH load (log-scaled) was significantly correlated with age ($r = 0.36$, $p < .01$). The association between age and WMH load was significant in most of the deep WM (dWM) ROIs, including the bilateral posterior dWM, bilateral occipital dWM, left frontal dWM, and SCC ($r = 0.24$–0.43, $p < .05$), but not in the subcortical WM (sWM), after FDR multiple comparison correction (Table 3 and Fig. 5A). The r value maps in Fig. 5C depict a central-to-peripheral pattern for the accumulation of WMH with age.

WM volumes were negatively correlated with several regions, including the bilateral parietal sWM, bilateral temporal sWM, right frontal sWM, bilateral anterior dWM, left posterior dWM, and right inferior dWM ($p < .05$ after FDR correction) (Table 2 and Fig. 5B). Fig. 5D demonstrates that age-dependent volume loss was more prominent in peripheral WM structures, which is opposite to the pattern of age-WMH relationships observed in Fig. 5C.

3.3. Relationship between location-dependent WMH and diagnosis

We also compared local WMH load in participants who were cognitively normal ($n = 98$) versus participants who had a diagnosis of MCI ($n = 22$). Significant group differences were found in the right inferior dWM, and left occipital sWM ($p < .05$), based on ANCOVA, adjusted for age and sex and corrected for multiple comparisons (Fig. 6A). No statistical group difference was found in the whole-brain WMH loads ($p = .06$). An example of the WMH distribution in an MCI subject is shown in Fig. 6C. Visual inspection indicated that WMH (red arrows in Fig. 6C) often crossed the inferior dWM and occipital sWM. Therefore, we further combined these two ROIs, and found significant differences in both left and right inferior-occipital WM regions between the normal and MCI groups ($p < .05$). The WM volumetric analysis revealed significant group differences in bilateral temporal sWM, after adjusting for age and sex and correcting for multiple comparison (Fig. 6B).

4. Discussion

4.1. The MADL framework

The new MADL algorithm seamlessly integrates WMH detection and image segmentation to quantify the WMH distribution. The method was built on a multi-atlas fusion algorithm and FLAIR multi-atlas library. While multi-atlas algorithms are traditionally used for segmenting images, we utilized the multi-atlas fusion process for WMH detection purpose. The MADL method can be considered as a special case of outlier detection. While previous outlier detection methods (Maldjian et al., 2013; Ong et al., 2012; Van Leemput et al., 2001; Yang et al., 2010) utilized the intensity profiles within the patient images; whereas MADL identifies outliers based

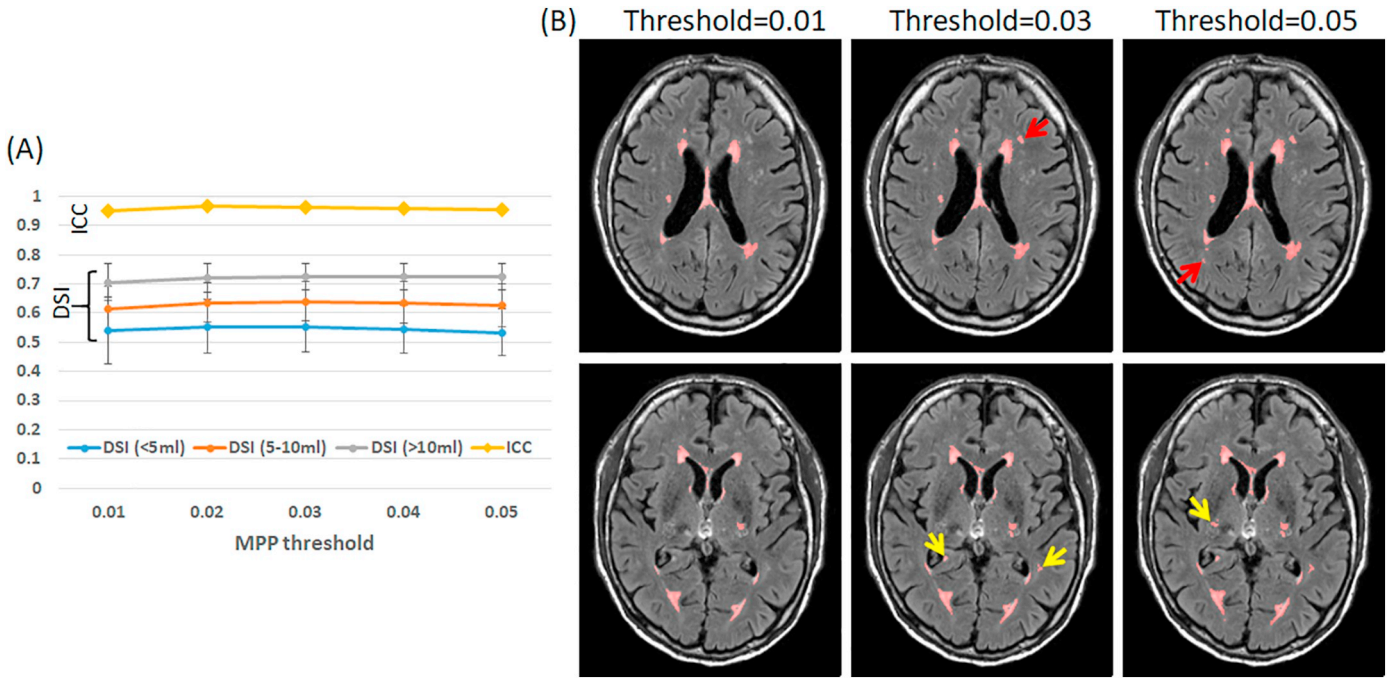

Fig. 4. Effects of varying thresholds (Maximized Posterior Probability, *MPP*) from 0.01 to 0.05 on the detection accuracy. (A) Algorithm evaluations based on the interclass correlation (ICC) and the Dice Similarity Index (DSI), in three groups (whole-brain WMH load of 0–5 ml, 5–10 ml, and > 10 ml). (B) MADL results at *MPP* threshold of 0.01, 0.03, and 0.05 in a brain at two transverse locations. Red arrows point to the increase of true-positive detection as the threshold increased; and yellow arrows point to the increase of false-positive detection as the threshold increased.

Table 2
Performance of MADL, BIANCA, and LST in detecting WMH, using manually delineated WMH as the gold standard.

Method	Load	DSI	FPR	FNR	ICC
MADL	0–5 ml	0.55 ± 0.09	0.42 ± 0.17	0.37 ± 0.16	0.97
	5–10 ml	0.64 ± 0.07	0.29 ± 0.11	0.33 ± 0.09	
	> 10 ml	0.72 ± 0.05	0.19 ± 0.09	0.31 ± 0.08	
BIANCA	0–5 ml	0.55 ± 0.14	0.55 ± 0.18* (↑)	0.22 ± 0.17* (↓)	0.95
	5–10 ml	0.67 ± 0.12	0.39 ± 0.14* (↑)	0.21 ± 0.15*(↓)	
	> 10 ml	0.74 ± 0.05	0.26 ± 0.07* (↑)	0.22 ± 0.08*(↓)	
LST	0–5 ml	0.53 ± 0.19	0.27 ± 0.21* (↓)	0.52 ± 0.24* (↑)	0.94
	5–10 ml	0.66 ± 0.12	0.22 ± 0.17* (↓)	0.41 ± 0.21* (↑)	
	> 10 ml	0.73 ± 0.09	0.17 ± 0.09	0.32 ± 0.13	

The dice similarity index (DSI), false positive rate (FPR), and false negative rate (FNR), were evaluated in low, median, and high WMH groups. Intra-class correlation (ICC) was calculated based on the detected and the manual delineated total WMH load in all subjects (n = 120), using the three methods.

* $p < .01$ by paired Student t-test between MADL and BIANCA and between MADL and LST.

on local intensity mismatches between the hyperintense voxels and normal appearing FLAIR images (atlases). Atlas-based approaches have also been proposed for lesion detection, e.g., creation of WMH probability atlases from patient groups and use it as a priori in the detection algorithms (Bricq et al., 2008; Yoshita et al., 2006). Alternatively, Shiee et al. (2010) used a single population-based atlas from normal subjects to identify intensity deviations in patient images, which is similar to the MADL idea. Strength of MADL is that it utilizes a multi-atlas fusion method that potentially offers a more flexible platform in that it selects atlases matched for the anatomical feature of the subject, such as the degree of atrophy. More importantly, if the atlases contain predefined segmentation, the anatomical locations of WMH can be automatically identified as done in the our method.

4.2. Algorithm performance

MADL algorithm showed a comparable detection accuracy compared to the state-of-the-art methods, and its performance was robust with respect to the choice of *MPP* thresholds. The accuracy of WMH detection algorithms varied depending on the amount of WMH in the brain. Using single-modality FLAIR contrast, DSI was between 0.51 and 0.71 in the subjects with < 5 ml WMH load, and overall accuracy was about 0.68 in previous reported fully-automated methods (Gibson et al., 2010; Ji et al., 2013; Khademi et al., 2012; Schmidt et al., 2012; Simoes et al., 2013; Yoo et al., 2014). Higher detection accuracy can be achieved with semi-automated methods (Iorio et al., 2013; Itti et al., 2001; Kawata et al., 2010; Ramirez et al., 2011). When compared directly using the same dataset, MADL, BIANCA, and LST showed similar DSI. Interestingly, the FPR was the highest in BIANCA and lowest in LST, and the FNR was the highest in LST and lowest in BIANCA, while MADL was intermediate (Table 2). There are several reasons why the outcomes of these algorithms may differ. BIANCA is a supervised learning method based on k-nearest neighborhood clustering, with user-defined options for spatial weighting, local intensity averaging, and choice of training points (Griffanti et al., 2016). In addition, the empirically defined threshold on the lesion probability map played an important role in balancing the FPR and FNR. We used a suggested threshold of 0.9 and default values for other parameters, which might not be optimal for our study population with relatively low WMH load. By comparison, the LPA algorithm (Schmidt, 2017) in LST was trained by a logistic regression model based on internal training data from multiple sclerosis patients with severe lesions. Given the intrinsic differences in lesion volume and pattern between the LPA training data and our testing data, it is possible that the algorithm is not sensitive enough to capture the small and subtle lesions in our data, leading to the high FNR.

Detection accuracy in the low load group was not ideal with all three methods. Small lesions with subtle abnormalities are known to be challenging for fully automated algorithms. Even human readers show considerable disagreement and inconsistency on detection of small lesions (Boutet et al., 2016). It is, therefore, recommended that visual inspection, parameter tuning, and manual correction are performed after automated detection. Another source of detection error resides in the ambiguity of WMH definition. WMH are commonly used to examine

Table 3
Correlation coefficients (*r*) between age and log-scale WMH load and those between age and WM volumes in major WM structures, as well as the total WMH load (in the unit of ml) in normal elderly and MCI patients in these structures.

White matter structure	*r* between age and WMH load	*r* between age and volumes	WMH (ml) in normal elderly	WMH (ml) in MCI patients
Anterior dWM (left)	0.26*	−0.37**	0.28 ± 0.41	0.35 ± 0.47
Anterior dWM (right)	0.09	−0.37**	0.20 ± 0.29	0.28 ± 0.40
Posterior dWM (left)	0.34**	−0.28*	0.06 ± 0.14	0.10 ± 0.27
Posterior dWM (right)	0.43**	0.15	0.04 ± 0.09	0.12 ± 0.28
Inferior dWM (left)	0.39**	−0.17	0.14 ± 0.23	0.24 ± 0.37
Inferior dWM (right)	0.24*	−0.23*	0.10 ± 0.16	0.25 ± 0.40
Frontal sWM (left)	0.15	−0.14	0.02 ± 0.03	0.01 ± 0.02
Frontal sWM (right)	0.06	−0.19*	0.03 ± 0.08	0.01 ± 0.03
Parietal sWM (left)	0.06	−0.32**	0.03 ± 0.05	0.03 ± 0.07
Parietal sWM (right)	0.26	−0.43**	0.02 ± 0.04	0.02 ± 0.05
Temporal sWM (left)	0.01	−0.28**	0.00 ± 0.01	0.00 ± 0.00
Temporal sWM (right)	0.09	−0.25*	0.01 ± 0.03	0.02 ± 0.04
Occipital sWM (left)	0.19	−0.13	0.08 ± 0.09	0.15 ± 0.20
Occipital sWM (right)	0.11	−0.16	0.08 ± 0.10	0.13 ± 0.19
Genu of CC	0.02	−0.12	0.08 ± 0.05	0.09 ± 0.06
Body of CC	0.17	0.17	0.14 ± 0.09	0.18 ± 0.12
Splenium of CC	0.28**	0.01	0.14 ± 0.19	0.21 ± 0.22

*$p < .05$ and **$p < .01$ after adjustment for gender and diagnosis and correction for multiple comparison. Abbreviations: dWM—deep white matter, sWM—subcortical white matter, CC—corpus callosum.

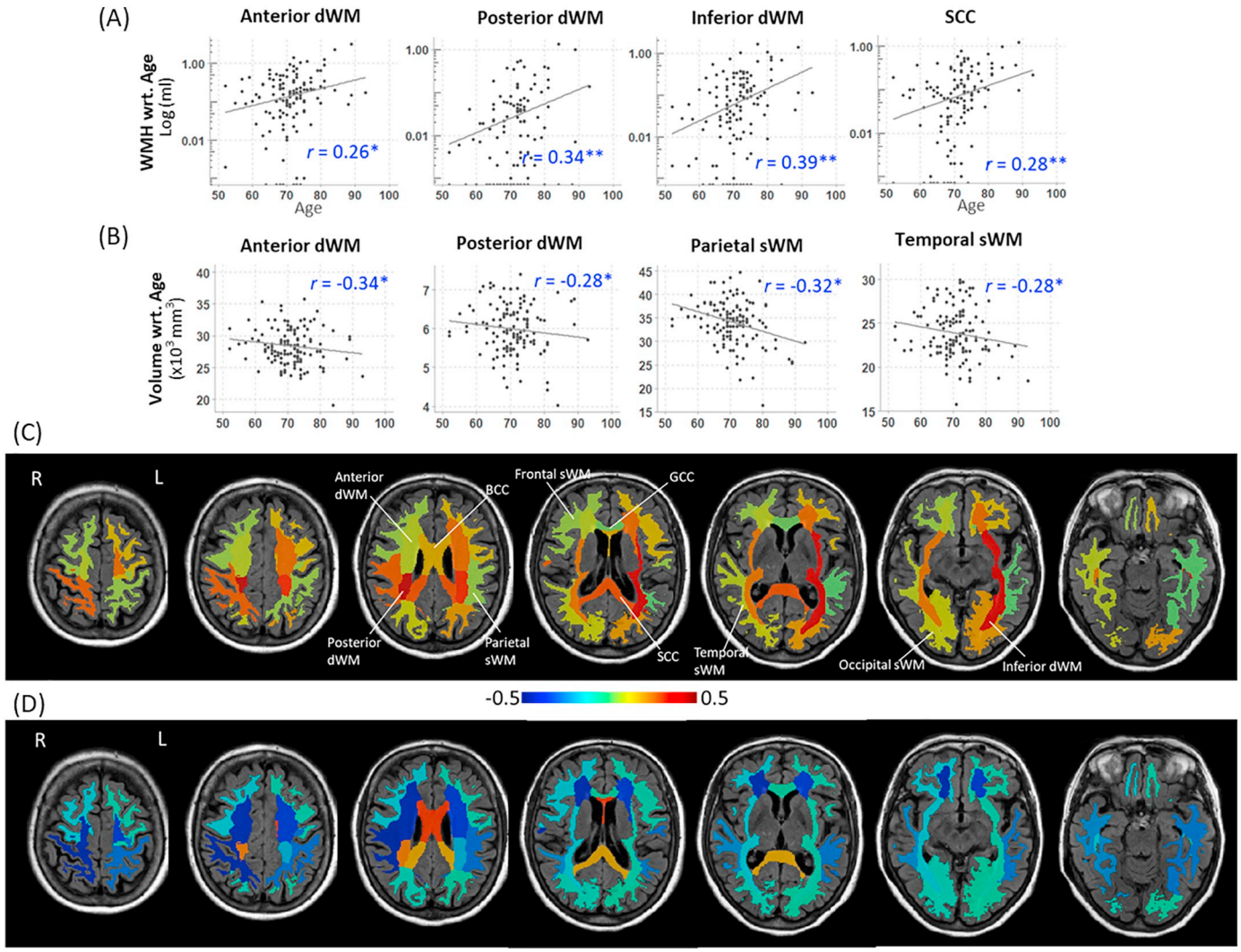

Fig. 5. Relations between WMH load, structural volume, and age. Linear regressions in structures that had significant correlations between their WMH and age (A) and between their volumes and age (B) are shown. Three-dimensional maps of the Pearson's correlation coefficients (*r*) between local WMH load and age (C), and between WM volumes and age (D) are also shown. Abbreviations: dWM, deep white matter; GCC, genu of the corpus callosum; BCC, body of the corpus callosum; SCC, splenium of the corpus callosum.; sWM, subcortical white matter.

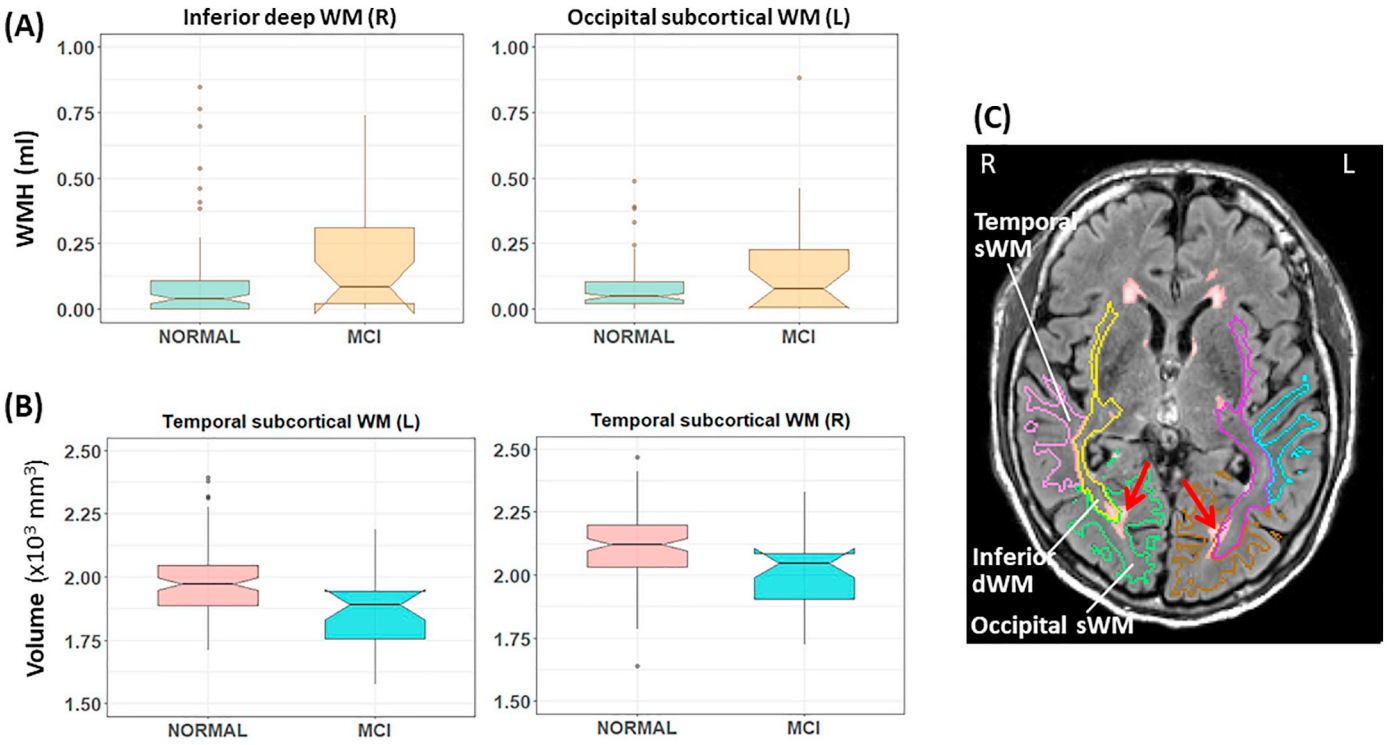

Fig. 6. (A) WM regions that show significantly increased WMH load in the mild cognitive impairment (MCI) subjects compared to normal controls ($p < .05$ after FDR correction). (B) WM regions that show significantly reduced volumes (normalized to intracranial volume) in the MCI subjects compared to the normal controls ($p < .05$ after FDR correction). (C) A representative FLAIR image of a MCI subject that shows WMH in the inferior deep WM and occipital subcortical WM (red arrows).

brain changes associated with small vessel disease, where most hyperintense voxels reside in WM, but lesions can also appear in GM (Wardlaw et al., 2013). The GM lesions could also be important but whether or not they should be detected by algorithms designed for WMH is controversial. In MADL, we only characterized lesions within a WM mask. However, our manual WMH delineation included lesions in both GM and WM. If we only use manual delineation within the WM mask as the gold standard, agreement between the MADL and manual results were higher, e.g., DSI can be improved to 0.59 ± 0.08, 0.66 ± 0.08, and 0.71 ± 0.04 in the low, median, and high load groups.

It should be noted that in the current MADL pipeline, we selected 15 FLAIR data for the multi-atlas library, which were from cognitively normal subjects with normal-appearing images and minimal WMH. Previous studies have reported that performance of T1-weighted multi-atlas segmentation improved with the number of atlases, but segmentation accuracy became relatively stable between 15 and 25 atlases, depending on the structures of interest (Aljabar et al., 2009). We examined the effect of atlas number by expanding the original 15 atlases to 20 atlases with brains that had WMH < 2 ml. We compared algorithm performance in a randomly selected subset of subjects ($n = 22$). The detection accuracy was not affected by atlas number: DSI was 0.65 ± 0.08 using 20 atlases, and 0.66 ± 0.08 using 15 atlases ($p > .05$ by paired t-test).

4.3. Significance of location-dependent WMH analysis

To examine the clinical use of the location-dependent WMH analysis by MADL, we first investigated WMH accumulation with advancing age. The location-dependent WMH analysis demonstrated that local WMH was significantly correlated with age in the dWM regions, while age-volume correlations were more prominent in the sWM. We also examined the location dependency of WMH distribution in normal elderly subjects compared to those with MCI. The association between WMH and AD has been reported in a number of studies (Barber et al., 1999; Burns et al., 2005; Prins et al., 2004; Vermeer et al., 2003), and MCI subjects tend to have an intermediate WMH burden (Yoshita et al., 2006). In our study, whole-brain WMH load was marginally different between cognitively normal and MCI participants, but we were able to identify significant local WMH increases in the inferior dWM and occipital sWM in MCI subjects, which is congruent with previous reports (Brickman et al., 2012; Yoshita et al., 2006). It remains to be demonstrated whether WMH load and WM atrophy can be used synergistically in identification of MCI. Although the WMH remains an incremental and non-specific feature in many neurodegenerative diseases (Wardlaw et al., 2013), our results suggested the value of examining the locations of WMH in subjects with memory impairment. The current study did not investigate associations between WMH and cognitive performances or other clinical factors, which will be an interesting clinical research topic to explore but outside the scope of the current study.

4.4. Limitations and future directions

The present study represents the initial phase of WMH lesion-detection techniques based on multi-atlas approaches; many interesting challenges remain. For example, ROI definitions in the FLAIR atlases were inherited from T1-weighted anatomical definitions, which could be optimized according to the characteristics of WMH distribution, e.g., by merging some of the ROIs, to facilitate clinical interpretation. In the current pipeline, histogram matching between subjects and atlases was performed based on the whole brain intensity profiles, which might be affected by the abnormal intensities in WMH regions. While this was not a particular concern for our study population who had low-to-median WMH load, for patients with extensive WMH, such as vascular dementia patients, it might be necessary to first exclude the WMH regions in histogram matching to ensure the accuracy of image registration, which would, in turn, improve the segmentation and detection of large WMH.

A few more strategies may also be attempted to further improve the performance of MADL. For example, one may explicitly use WMH probability information obtained from existing subjects as a *prior* to refine the MADL detection results, similar to (Bricq et al., 2008; Yoshita et al., 2006). Another natural extension of the technology is to combine multiple MR contrasts (e.g., T1-weighted images), into a multi-contrast multi-atlas approach to further improve detection and segmentation accuracy. Within the FLIAR image contrast, additional features beyond the intensity may be utilized, including the first and second order features from texture analysis (Haralick et al., 1973), Haar-like features that enhance edge and shape information (Lienhart and Maydt, 2002), context information (Torralba et al., 2003), or even features learned from radiomics (Gillies et al., 2016) or machine-learning. These high-order image features may assist the detection of small lesions with subtle intensity abnormalities. Also, the boundaries of hyperintense regions, which were subject to partial volume effects, may be better captured by incorporating graph-cut (Shi and Malik, 1997) or level-set (Chan and Vese, 2001) algorithms.

In addition, recent advances in deep learning have opened a new avenue for medical image analysis, including WMH detection (Ghafoorian et al., 2016; Ghafoorian et al., 2017; Jin et al., 2018; Moeskops et al., 2018; Roa-Barco et al., 2018). For example, using a fully-connected convolutional neural network (CNN) with multiple branches, Moeskops et al. showed that WMH can be segmented jointed with GM/WM/CSF with DSI around 0.54 using three image contrasts and DSI of 0.51 using only FLAIR (Moeskops et al., 2018). Rachmadi and colleagues reported that the segmentation accuracy of CNN also depended on the WMH load, e.g., for small-to-median size WMH (1.5–13 ml), DSI was between 0.46 and 0.55, and it was increased to 0.72 for very large WMH ($>$ 24 ml). In the 2017 MICCAI Challenge of WMH Segmentation (http://wmh.isi.uu.nl/), deep learning based methods were among the top performers, which achieved DSI over 0.7 using T1-weighted and FLAIR images in patients with presumed vascular disease. Potential integration of deep learning and traditional image processing algorithms may strengthen the advantages of both and fundamentally improve the performance of image segmentation and lesion detection.

5. Conclusion

The multi-atlas based method provided a one-stop-shop solution for simultaneous detection and localization of WMH on FLAIR data. This method demonstrated good detection accuracy compared with manual delineation and other existing methods. The location-dependent WMH analysis suggested a higher association between deep WMH and age compared to subcortical WMH. Our findings also suggested that WMH may differentially accumulate in the inferior and occipital WM during the MCI phase of AD.

Acknowledgements

This work is supported by NIH grants R01 NS107417, R21 NS098018, R03 AG060340, P41 EB015909, R01 NS084957, R01 NS086888, P50 AG005146, and U19 AG033655, Natural Sicence Fundation of China (No. 61801424), Major Scientific Project of Zhejiang Lab (No. 2018DG0ZX01) (China), and the Fundamental Research Funds for the Central Universities (China).

None of the authors have a conflict of interest to declare.

References

Admiraal-Behloul, F., van den Heuvel, D.M.J., Olofsen, H., van Osch, M.J.P., van der Grond, J., van Buchem, M.A., Relber, J.H.C., 2005. Fully automatic segmentation of white matter hyperintensities in MR images of the elderly. Neuroimage 28, 607–617.

Albert, M.S., DeKosky, S.T., Dickson, D., Dubois, B., Feldman, H.H., Fox, N.C., Gamst, A., Holtzman, D.M., Jagust, W.J., Petersen, R.C., Snyder, P.J., Carrillo, M.C., Thies, B., Phelps, C.H., 2011. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement. J. Alzheimers Assoc. 7, 270–279.

Albert, M., Soldan, A., Gottesman, R., McKhann, G., Sacktor, N., Farrington, L., Grega, M., Turner, R., Lu, Y., Li, S.S., Wang, M.C., Selnes, O., 2014. Cognitive changes preceding clinical symptom onset of mild cognitive impairment and relationship to ApoE genotype. Curr. Alzheimer Res. 11, 773–784.

Aljabar, P., Heckemann, R.A., Hammers, A., Hajnal, J.V., Rueckert, D., 2009. Multi-atlas based segmentation of brain images: atlas selection and its effect on accuracy. Neuroimage 46, 726–738.

Barber, R., Scheltens, F., Gholkar, A., Ballard, C., McKeith, I., Ince, P., Perry, R., O'Brien, J., 1999. White matter lesions on magnetic resonance imaging in dementia with Lewy bodies, Alzheimer's disease, vascular dementia, and normal aging. J. Neurol. Neurosurg. Psychiatr. 67, 66–72.

Beare, R., Srikanth, V., Chen, J., Phan, T.G., Stapleton, J., Lipshut, R., Reutens, D., 2009. Development and validation of morphological segmentation of age-related cerebral white matter hyperintensities. Neuroimage 47, 199–203.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate – a practical and powerful approach to multiple testing. J. R. Stat. Soc. B Methodol. 57, 289–300.

Biesbroek, J.M., Kuijf, H.J., van der Graaf, Y., Vincken, K.L., Postma, A., Mali, W.P.T.M., Biessels, G.J., Geerlings, M.I., Grp, S.S., 2013. Association between subcortical vascular lesion location and cognition: a voxel-based and tract-based lesion-symptom mapping study. The SMART-MR study. PLoS One 8.

Boutet, C., Rouffiange-Leclair, L., Schneider, F., Camdessanche, J.P., Antoine, J.C., Barral, F.G., 2016. Visual assessment of age-related white matter Hyperintensities using FLAIR images at 3 T: inter- and intra-rater agreement. Neurodegener. Dis. 16, 279–283.

Brickman, A.M., 2013. Contemplating Alzheimer's disease and the contribution of white matter hyperintensities. Curr. Neurol. Neurosci. Rep. 13, 415.

Brickman, A.M., Provenzano, F.A., Muraskin, J., Manly, J.J., Blum, S., Apa, Z., Stern, Y., Brown, T.R., Luchsinger, J.A., Mayeux, R., 2012. Regional white matter hyperintensity volume, not hippocampal atrophy, predicts incident Alzheimer disease in the community. Arch. Neurol. 69, 1621–1627.

Brickman, A.M., Zahodne, L.B., Guzman, V.A., Narkhede, A., Meier, I.B., Griffith, E.Y., Provenzano, F.A., Schupf, N., Manly, J.J., Stern, Y., Luchsinger, J.A., Mayeux, R., 2015. Reconsidering harbingers of dementia: progression of parietal lobe white matter hyperintensities predicts Alzheimer's disease incidence. Neurobiol. Aging 36, 27–32.

Bricq, S., Collet, C., Armspach, J.P., 2008. Lesions detection on 3D brain MRI using trimmmed likelihood estimator and probabilistic atlas. 2008 IEEE Int. Symp. Biomed. Imaging 1-4, 93.

Burns, J.M., Church, J.A., Johnson, D.K., Xiong, C., Marcus, D., Fotenos, A.F., Snyder, A.Z., Morris, J.C., Buckner, R.L., 2005. White matter lesions are prevalent but differentially related with cognition in aging and early Alzheimer disease. Arch. Neurol. 62, 1870–1876.

Caligiuri, M.E., Perrotta, P., Augimeri, A., Rocca, F., Quattrone, A., Cherubini, A., 2015. Automatic detection of white matter Hyperintensities in healthy aging and pathology using magnetic resonance imaging: a review. Neuroinformatics 13, 261–276.

Chan, T.F., Vese, L.A., 2001. Active contours without edges. IEEE Trans. Image Process. 10, 266–277.

Christensen, G.E., Rabbitt, R.D., Miller, M.I., 1996. Deformable templates using large deformation kinematics. IEEE Trans. Image Process. 5, 1435–1447.

Coltuc, D., Bolon, P., Chassery, J.M., 2006. Exact histogram specification. IEEE Trans. Image Process. 15, 1143–1152.

de Groot, J.C., de Leeuw, F.E., Oudkerk, M., van Gijn, J., Hofman, A., Jolles, J., Breteler, M.M.B., 2002. Periventricular cerebral white matter lesions predict rate of cognitive decline. Ann. Neurol. 52, 335–341.

Debette, S., Markus, H.S., 2010. The clinical importance of white matter hyperintensities on brain magnetic resonance imaging: systematic review and meta-analysis. BMJ 341, c3666.

DeCarli, C., Fletcher, E., Ramey, V., Harvey, D., Jagust, W.J., 2005. Anatomical mapping of white matter hyperintensities (WMH): exploring the relationships between periventricular WMH, deep WMH, and total WMH burden. Stroke 36, 50–55.

Ghafoorian, M., Karssemeijer, N., Heskes, T., Uder, I.W.M.v., Leeuw, F.E.d., Marchiori, E., Ginneken, B.v., Platel, B., 2016. Non-uniform patch sampling with deep convolutional neural networks for white matter hyperintensity segmentation. In: 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), pp. 1414–1417.

Ghafoorian, M., Karssemeijer, N., Heskes, T., van Uden, I.W.M., Sanchez, C.I., Litjens, G., de Leeuw, F.E., van Ginneken, B., Marchiori, E., Platel, B., 2017. Location sensitive deep convolutional neural networks for segmentation of white matter Hyperintensities. Sci. Rep. 7.

Gibson, E., Gao, F.Q., Black, S.E., Lobaugh, N.J., 2010. Automatic segmentation of white matter hyperintensities in the elderly using FLAIR images at 3T. J. Magn. Reson. Imaging 31, 1311–1322.

Gillies, R.J., Kinahan, P.E., Hricak, H., 2016. Radiomics: images are more than pictures, they are data. Radiology 278, 563–577.

Gorelick, P.B., Scuteri, A., Black, S.E., Decarli, C., Greenberg, S.M., Iadecola, C., Launer, L.J., Laurent, S., Lopez, O.L., Nyenhuis, D., Petersen, R.C., Schneider, J.A., Tzourio, C., Arnett, D.K., Bennett, D.A., Chui, H.C., Higashida, R.T., Lindquist, R., Nilsson, P.M., Roman, G.C., Sellke, F.W., Seshadri, S., American Heart Association Stroke Council, C.o.E., Prevention, C.o.C.N.C.o.C.R., Intervention, Council on Cardiovascular, S., Anesthesia, 2011. Vascular contributions to cognitive impairment and dementia: a statement for healthcare professionals from the american heart association/american stroke association. Stroke 42, 2672–2713.

Grenander, U., Miller, M.I., 1998. Computational anatomy: an emerging discipline. Q. Appl. Math. 56, 617–694.

Griffanti, L., Zamboni, G., Khan, A., Li, L., Bonifacio, G., Sundaresan, V., Schulz, U.G., Kuker, W., Battaglini, M., Rothwell, P.M., Jenkinson, M., 2016. BIANCA (Brain intensity AbNormality Classification Algorithm): a new tool for automated segmentation of white matter hyperintensities. Neuroimage 141, 191–205.
Haralick, R.M., Shanmugam, K., Dinstein, I., 1973. Textural features for image classification. In: IEEE Transactions on Systems, Man, and Cybernetics SMC-3, pp. 610–621.
Herskovits, E.H., Bryan, R.N., Yang, F., 2008. Automated Bayesian segmentation of microvascular white-matter lesions in the ACCORD-MIND study. Adv. Med. Sci. 53, 182–190.
Iorio, M., Spalletta, G., Chiapponi, C., Luccichenti, G., Cacciari, C., Orfei, M.D., Caltagirone, C., Piras, F., 2013. White matter hyperintensities segmentation: a new semi-automated method. Front. Aging Neurosci. 5.
Ithapu, V., Singh, V., Lindner, C., Austin, B.P., Hinrichs, C., Carlsson, C.M., Bendlin, B.B., Johnson, S.C., 2014. Extracting and summarizing white matter Hyperintensities using supervised segmentation methods in Alzheimer's disease risk and aging studies. Hum. Brain Mapp. 35, 4219–4235.
Itti, L., Chang, L., Ernst, T., 2001. Segmentation of progressive multifocal leukoencephalopathy lesions in fluid-attenuated inversion recovery magnetic. J. Neuroimaging 11, 412–417.
Jack, C.R., O'Brien, P.C., Rettman, D.W., Shiung, M.M., Xu, Y.C., Muthupillai, R., Manduca, A., Avula, R., Erickson, B.J., 2001. FLAIR histogram segmentation for measurement of leukoaraiosis volume. J. Magn. Reson. Imaging 14, 668–676.
Ji, S., Ye, C., Li, F., Sun, W., Zhang, J., Huang, Y., Fang, J., 2013. Automatic segmentation of white matter hyperintensities by an extended FitzHugh & Nagumo reaction diffusion model. J. Magn. Reson. Imaging 37, 343–350.
Jin, D., Xu, Z., Harrison, A.P., Mollura, D.J., 2018. White matter hyperintensity segmentation from T1 and FLAIR images using fully convolutional neural networks enhanced with residual connections. In: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), pp. 1060–1064.
Kawata, Y., Arimura, H., Yamashita, Y., Magome, T., Ohki, M., Toyofuku, F., Higashida, Y., Tsuchiya, K., 2010. Computer-aided evaluation method of white matter hyperintensities related to subcortical vascular dementia based on magnetic resonance imaging. Comput. Med. Imaging Graph. 34, 370–376.
Khademi, A., Venetsanopoulos, A., Moody, A.R., 2012. Robust white matter lesion segmentation in FLAIR MRI. IEEE Trans. Biomed. Eng. 59, 860–871.
Langsrud, Y., 2003. ANOVA for unbalanced data: use type II instead of type III sums of squares. Stat. Comput. 13, 163–167.
Lao, Z.Q., Shen, D.G., Liu, D.F., Jawad, A.F., Melhern, E.R., Launer, L.J., Bryan, R.N., Davatzikos, C., 2008. Computer-assisted segmentation of white matter lesions in 3D MR images using support vector machine. Acad. Radiol. 15, 300–313.
Lee, S., Viqar, F., Zimmerman, M.E., Narkhede, A., Tosto, G., Benzinger, T.L., Marcus, D.S., Fagan, A.M., Goate, A., Fox, N.C., Cairns, N.J., Holtzman, D.M., Buckles, V., Ghetti, B., McDade, E., Martins, R.N., Saykin, A.J., Masters, C.L., Ringman, J.M., Ryan, N.S., Forster, S., Laske, C., Schofield, P.R., Sperling, R.A., Salloway, S., Correia, S., Jack Jr., C., Weiner, M., Bateman, R.J., Morris, J.C., Mayeux, R., Brickman, A.M., Dominantly Inherited Alzheimer, N., 2016. White matter hyperintensities are a core feature of Alzheimer's disease: evidence from the dominantly inherited Alzheimer network. Ann. Neurol. 79, 929–939.
Lienhart, R., Maydt, J., 2002. An extended set of haar-like features for rapid object detection. In: 2002 International Conference on Image Processing, Vol I, Proceedings, pp. 900–903.
Ma, J.B., Ma, H.T., Li, H.T., Ye, C.F., Wu, D., Tang, X.Y., Miller, M., Mori, S., 2015. A fast atlas pre-selection procedure for multi-atlas based brain segmentation. In: 2015 37th Annual International Conference of the Ieee Engineering in Medicine and Biology Society (Embc), pp. 3053–3056.
Maldjian, J.A., Whitlow, C.T., Saha, B.N., Kota, G., Vandergriff, C., Davenport, E.M., Divers, J., Freedman, B.I., Bowden, D.W., 2013. Automated white matter total lesion volume segmentation in diabetes. Am. J. Neuroradiol. 34, 2265–2270.
McGraw, K.O., Wong, S.P., 1996. Forming inferences about some intraclass correlations coefficients (vol 1, pg 30, 1996). Psychol. Methods 1, 390.
Moeskops, P., de Bresser, J., Kuijf, H.J., Mendrik, A.M., Biessels, G.J., Pluim, J.P.W., Isgum, I., 2018. Evaluation of a deep learning approach for the segmentation of brain tissues and white matter hyperintensities of presumed vascular origin in MRI. Neuroimage Clin. 17, 251–262.
Murray, M.E., Senjem, M.L., Petersen, R.C., Hollman, J.H., Preboske, G.M., Weigand, S.D., Knopman, D.S., Ferman, T.J., Dickson, D.W., Jack, C.R., 2010. Functional impact of white matter hyperintensities in cognitively normal elderly subjects. Arch. Neurol. 67, 1379–1385.
Ong, K.H., Ramachandram, D., Mandava, R., Shuaib, I.L., 2012. Automatic white matter lesion segmentation using an adaptive outlier detection method. Magn. Reson. Imaging 30, 807–823.
Prins, N.D., Scheltens, P., 2015. White matter hyperintensities, cognitive impairment and dementia: an update. Nat. Rev. Neurol. 11.
Prins, N.D., van Dijk, E.J., den Heijer, T., Vermeer, S.E., Koudstaal, P.J., Oudkerk, M., Hofman, A., Breteler, M.M., 2004. Cerebral white matter lesions and the risk of dementia. Arch. Neurol. 61, 1531–1534.
Ramirez, J., Gibson, E., Quddus, A., Lobaugh, N.J., Feinstein, A., Levine, B., Scott, C.J.M., Levy-Cooperman, N., Gao, F.Q., Black, S.E., 2011. Lesion explorer: a comprehensive segmentation and parcellation package to obtain regional volumetrics for subcortical hyperintensities and intracranial tissue. Neuroimage 54, 963–973.
Roa-Barco, L., Serradilla-Casado, O., de Velasco-Vázquez, M., López-Zorrilla, A., Graña, M., Chyzhyk, D., Price, C., 2018. A 2D/3D Convolutional Neural Network for Brain White Matter Lesion Detection in Multimodal MRI. Springer International Publishing, Cham, pp. 377–385.
Schmidt, P., 2017. Bayesian Inference for Structured Additive Regression Models for Large-Scale Problems with Applications to Medical Imaging. Ludwig-Maximilians University, München.
Schmidt, P., Gaser, C., Arsic, M., Buck, D., Forschler, A., Berthele, A., Hoshi, M., Ilg, R., Schmid, V.J., Zimmer, C., Hemmer, B., Muhlau, M., 2012. An automated tool for detection of FLAIR-hyperintense white-matter lesions in multiple sclerosis. Neuroimage 59, 3774–3783.
Seghier, M.L., Ramlackhansingh, A., Crinion, J., Leff, A.P., Price, C.J., 2008. Lesion identification using unified segmentation-normalisation models and fuzzy clustering. Neuroimage 41, 1253–1266.
Shi, J., Malik, J., 1997. Normalized cuts and image segmentation. In: Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition (CVPR '97). IEEE Computer Society, pp. 731.
Shi, L., Wang, D.F., Liu, S.P., Pu, Y.H., Wang, Y.L., Chu, W.C.W., Ahuja, A.T., Wang, Y.J., 2013. Automated quantification of white matter lesion in magnetic resonance imaging of patients with acute infarction. J. Neurosci. Methods 213, 138–146.
Shiee, N., Bazin, P.L., Ozturk, A., Reich, D.S., Calabresi, P.A., Pham, D.L., 2010. A topology-preserving approach to the segmentation of brain images with multiple sclerosis lesions. Neuroimage 49, 1524–1535.
Simoes, R., Monninghoff, C., Dlugaj, M., Weimar, C., Wanke, I., van Walsum, A.M.V., Slump, C., 2013. Automatic segmentation of cerebral white matter hyperintensities using only 3D FLAIR images. Magn. Reson. Imaging 31, 1182–1189.
Soldan, A., Pettigrew, C., Cai, Q., Wang, M.C., Moghekar, A.R., O'Brien, R.J., Selnes, O.A., Albert, M.S., Team, B.R., 2016. Hypothetical preclinical Alzheimer disease groups and longitudinal cognitive change. Jama Neurol. 73, 698–705.
Tang, X.Y., Oishi, K., Faria, A.V., Hillis, A.E., Albert, M.S., Mori, S., Miller, M.I., 2013. Bayesian parameter estimation and segmentation in the multi-atlas random orbit model. PLoS One 8.
Torralba, A., Murphy, K.P., Freeman, W.T., Rubin, M.A., 2003. Context-based vision system for place and object recognition. In: Ninth Ieee International Conference on Computer Vision, Vols I and II, Proceedings, pp. 273–280.
Tustison, N.J., Avants, B.B., Cook, P.A., Zheng, Y.J., Egan, A., Yushkevich, P.A., Gee, J.C., 2010. N4ITK: improved N3 Bias correction. IEEE Trans. Med. Imaging 29, 1310–1320.
Van Leemput, K., Maes, F., Vandermeulen, D., Colchester, A., Suetens, P., 2001. Automated segmentation of multiple sclerosis lesions by model outlier detection. IEEE Trans. Med. Imaging 20, 677–688.
Vemuri, P., Jack, C.R., 2010. Role of structural MRI in Alzheimer's disease. Alzheimers Res. Ther. 2, 23.
Vermeer, S.E., Prins, N.D., den Heijer, T., Hofman, A., Koudstaal, P.J., Breteler, M.M., 2003. Silent brain infarcts and the risk of dementia and cognitive decline. N. Engl. J. Med. 348, 1215–1222.
Wardlaw, J.M., Smith, E.E., Biessels, G.J., Cordonnier, C., Fazekas, F., Frayne, R., Lindley, R.I., O'Brien, J.T., Barkhof, F., Benavente, O.R., Black, S.E., Brayne, C., Breteler, M., Chabriat, H., DeCarli, C., de Leeuw, F.E., Doubal, F., Duering, M., Fox, N.C., Greenberg, S., Hachinski, V., Kilimann, I., Mok, V., van Oostenbrugge, R., Pantoni, L., Speck, O., Stephan, B.C.M., Teipel, S., Viswanathan, A., Werring, D., Chen, C., Smith, C., van Buchem, M., Norrving, B., Gorelick, P.B., Dichgans, M., Changes, S.R.V., 2013. Neuroimaging standards for research into small vessel disease and its contribution to ageing and neurodegeneration. Lancet Neurol. 12, 822–838.
Wu, M.J., Rosano, C., Butters, M., Whyte, E., Nable, M., Crooks, R., Meltzer, C.C., Reynolds, C.F., Aizenstein, H.J., 2006. A fully automated method for quantifying and localizing white matter hyperintensities on MR images. Psychiatr. Res. Neuroimaging 148, 133–142.
Wu, D., Ma, T., Ceritoglu, C., Li, Y., Chotiyanonta, J., Hou, Z., Hsu, J., Xu, X., Brown, T., Miller, M.I., Mori, S., 2016. Resource atlases for multi-atlas brain segmentations with multiple ontology levels based on T1-weighted MRI. Neuroimage 125, 120–130.
Yang, F.G., Shan, Z.Y., Kruggel, F., 2010. White matter lesion segmentation based on feature joint occurrence probability and chi(2) random field theory from magnetic resonance (MR) images. Pattern Recogn. Lett. 31, 781–790.
Yoo, B.I., Lee, J.J., Han, J.W., Oh, S.Y., Lee, E.Y., MacFall, J.R., Payne, M.E., Kim, T.H., Kim, J.H., Kim, K.W., 2014. Application of variable threshold intensity to segmentation for white matter hyperintensities in fluid attenuated inversion recovery magnetic resonance images. Neuroradiology 56, 265–281.
Yoshita, M., Fletcher, E., Harvey, D., Ortega, M., Martinez, O., Mungas, D.M., Reed, B.R., DeCarli, C.S., 2006. Extent and distribution of white matter hyperintensities in normal aging, MCI, and AD. Neurology 67, 2192–2198.