Computerized Medical Imaging and Graphics 89 (2021) 101873

Contents lists available at ScienceDirect

Computerized Medical Imaging and Graphics

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

An anatomical knowledge-based MRI deep learning pipeline for white matter hyperintensity quantification associated with cognitive impairment

Li Liang [a,b], Pengzheng Zhou [b], Wanxin Lu [a], Xutao Guo [a,b], Chenfei Ye [b], Haiyan Lv [c], Wang Tong [a], Ting Ma [a,b,d,e,*]

[a] School of Electronic and Information Engineering, Harbin Institute of Technology at Shenzhen, Shenzhen, Guangdong, China
[b] Peng Cheng Laboratory, Shenzhen, Guangdong, China
[c] Mindsgo Life Science Shenzhen Ltd, Shenzhen, Guangdong, China
[d] National Clinical Research Center for Geriatric Disorders, Xuanwu Hospital Capital Medical University, Beijing, China
[e] Advanced Innovation Center for Human Brain Protection, Capital Medical University, Beijing, China

ARTICLE INFO

Keywords:
White matter hyperintensities
Anatomical knowledge
Segmentation
Deep learning
Cognitive impairment

ABSTRACT

Recent studies have confirmed that white matter hyperintensities (WMHs) accumulated in strategic brain regions can predict cognitive impairments associated with Alzheimer's disease (AD). The knowledge of white matter anatomy facilitates lesion-symptom mapping associated with cognition, and provides important spatial information for lesion segmentation algorithms. However, deep learning-based methods in the white matter hyperintensity (WMH) segmentation realm do not take full advantage of anatomical knowledge in decision-making and lesion localization processes. In this paper, we proposed an anatomical knowledge-based MRI deep learning pipeline (^{A}U-Net), handcrafted anatomical-based spatial features developed from brain atlas were integrated with a well-designed U-Net configuration to simultaneously segment and locate WMHs. Manually annotated data from WMH segmentation challenge were used for the evaluation. We then applied this pipeline to investigate the association between WMH burden and cognition within another publicly available database. The results showed that ^{A}U-Net significantly improved segmentation performance compared with methods that did not incorporate anatomical knowledge ($p < 0.05$), and achieved a 14–17% increase based on area under the curve (AUC) in the cohort with mild WMH burden. Different configurations for incorporating anatomical knowledge were evaluated, the proposed stage-wise ^{A}U-Net-two-step method achieved the best performance (Dice: 0.86, modified Hausdorff distance: 3.06 mm), which was comparable with the state-of-the-art method (Dice: 0.87, modified Hausdorff distance: 3.62 mm). We observed different WMH accumulation patterns associated with normal aging and cognitive impairments. Furthermore, the characteristics of individual WMH lesions located in strategic regions (frontal and parietal subcortical white matter, as well as corpus callosum) were significantly correlated with cognition after adjusting for total lesion volumes. Our findings suggest that ^{A}U-Net is a reliable method to segment and quantify brain WMHs in elderly cohorts, and is valuable in individual cognition evaluation.

1. Introduction

White matter hyperintensities (WMHs), defined as hyperintensities on T2-weighted magnetic resonance (MR) images, are located in cerebral white matter tissues and of varying sizes (Wardlaw et al., 2013b). Although the pathogenesis of WMHs are not well understood (Prins and Scheltens, 2015; Wallin et al., 2018), mounting evidence has confirmed that regional WMHs predict cognitive impairments associated with Alzheimer's disease (AD) (Brickman et al., 2018; Dadar et al., 2019; Damulina et al., 2019; Habes et al., 2018; Lee et al., 2016). Mild cognitive impairment (MCI) subjects with higher regional white matter hyperintensity (WMH) volumes have an increased chance of undergoing conversion to AD (Dadar et al., 2019; Li et al., 2016). Additionally, increased WMH burden in MCI cohorts has been detected in regions of inferior deep white matter and occipital subcortical white matter, when compared with cognitive normal controls, by a multi-atlas based method (Wu et al., 2019). Poor processing speed, executive functioning, and episodic memory are correlated with regional WMH lesions (Biesbroek

* Corresponding author at: School of Electronic and Information Engineering, Harbin Institute of Technology at Shenzhen, Shenzhen, Guangdong, China.
E-mail address: tma@hit.edu.cn (T. Ma).

https://doi.org/10.1016/j.compmedimag.2021.101873
Received 4 July 2020; Received in revised form 29 December 2020; Accepted 27 January 2021
Available online 3 February 2021

et al., 2017). As the brain is well structured for functioning, the localization of WMHs within strategic regions should be important for evaluating lesion-symptom mapping associated with cognition. However, the literature is heterogeneous and lacks standards in describing the regional accumulation of WMHs, which highlights the importance of developing atlas-based quantitative methods to characterize the anatomical distribution of WMHs.

Deep learning-based methods have become state-of-the-art in many brain lesion segmentation tasks based on MR imaging. Recently, a challenge was held aiming to compare methods for the automated segmentation of WMHs in Medical Image Computing and Computer-Assisted Intervention (MICCAI) 2017 (https://wmh.isi.uu.nl/). In this realm, most convolutional neural network (CNN) based methods use the intensities of T1 and T2-weighted fluid attenuated inversion recovery (T2-FLAIR) as the input features; however, the inclusion of spatial information has also been confirmed to improve the performance of segmentation algorithms (Ghafoorian et al., 2017; Griffanti et al., 2016; Rachmadi et al., 2018). The spatial coordinates in MNI space and voxel-based prior probability templates of WMH are commonly used approaches for incorporating spatial information. Although CNN-based methods have made considerable progress in the task of WMH segmentation, the literature still lacks enough studies to explore the strategies of 1) the incorporation of anatomical-based knowledge in the decision-making process of deep learning and 2) the identification of anatomical locations for detected individual lesions immediately after deep learning-based segmentation.

In this paper, we described an anatomical knowledge-based MRI deep learning pipeline (^{A}U-Net) to simultaneously segment and locate WMHs. Anatomical knowledge was defined in brain atlas that was specifically developed for WMH segmentation and quantification. Two configurations were designed for ^{A}U-Net to combine anatomical knowledge with intensity information in the segmentation procedure. We evaluated the performance of ^{A}U-Net by comparing it to other methods that did not integrate anatomical knowledge, including the state-of-the-art method in MICCAI 2017. Finally, we used this pipeline to investigate the associations among WMH burden, normal aging and cognitive impairments.

2. Materials and methods

2.1. Data preparation

Two publicly available datasets were employed in this study. The first one is from WMH segmentation challenge in MICCAI 2017. Sixty cases were enrolled in publicly available training set. For each subject, a 3D T1 weighted image and a 2D T2-FLAIR image were provided. WMHs on T2-FLAIR were annotated manually according to Standards for Reporting Vascular changes on nEuroimaging (STRIVE) (Wardlaw et al., 2013b), these annotations were golden standard for evaluation of WMH segmentation in this study. The second is from Alzheimer's Disease Neuroimaging Initiative database (ADNI) (http://adni.loni.usc.edu). We enrolled 111 MCI, 80 AD and 94 cognitive normal (CN) subjects from ADNI database in this study. Inclusion criteria are as follow: 1) acquisitions from ADNI2 project with good quality of T1 and T2-FLAIR, 2) age between 60–85 years, 3) neurological assessment batteries are available, including: Alzheimer's Disease Assessment Scale (ADAS), Mini Mental State Examination (MMSE) and Functional Assessment Questionnaire (FAQ), and 4) neurological assessments should be in the same year as MRI acquisitions. The demographic statistics are listed in Table 1. All participants in ADNI had provided informed written consents before recruitment and filled out questionnaires approved by the respective Institutional Review Board (IRB).

For preprocessing, we first performed intra-subject coregistration between T2-FLAIR and T1 using FSL FLIRT affine transformation (Jenkinson et al., 2002; Jenkinson and Smith, 2001). After coregistration, global inhomogeneity corrections of T1 and T2-FLAIR were performed using N4 bias field correction (Tustison et al., 2010), which is built in advanced normalization tools (ANTs). Gaussian normalization was employed to normalize the voxel intensities of each subject with mean zero and standard deviation of one. Intracranial brain tissue masks were generated using FSL BET function (Smith, 2002).

Table 1
Demographic statistics for ADNI subjects enrolled in this study.

	CN	MCI	AD	P-value
Number of subjects	94	111	80	–
Gender (Male / Female)	43 / 51	64 / 47	41 / 39	0.23 *
Age (years)	74.00 ± 7.11	74.25 ± 7.51	75.10 ± 7.66	0.57 #
Education (years)	16.32 ± 2.46	16.16 ± 2.66	16.34 ± 2.68	0.85 #
ADAS	9.00 ± 4.58	18.87 ± 7.31	31.03 ± 7.64	**< 0.0001** #
MMSE	28.78 ± 1.42	27.19 ± 1.82	22.70 ± 2.22	**< 0.0001** #
FAQ	0.19 ± 1.02	4.25 ± 4.55	14.22 ± 7.08	**< 0.0001** #

Data are number or mean ± standard deviation.
* : Chi-square test was used.
# : Non-parametric Kruskal-Wallis test was used. Bold font indicates significance after correcting for multiple comparisons. Abbreviations: ADAS: Alzheimer's Disease Assessment Scale, MMSE: Mini Mental State Examination, FAQ: Functional Assessment Questionnaire.

2.2. Atlas generation

A T1 image in Montreal Neurological Institute (MNI) space was segmented using Brain Label (http://brainlabel.org), which provides automated cloud service for brain parcellation of T1 image based on multiple atlas likelihood fusion algorithm and pre-selection strategy (Tang et al., 2013; Wu et al., 2016). Brain Label segments the whole brain into different anatomical structures, with the finest segmentation including 283 regions.

We regrouped the finest 283 ROIs and defined 29 brain regions for intracranial brain tissue, thus generated a brain atlas in MNI space specifically designed for the segmentation and localization of WMHs. The hierarchical definition of 29 brain regions are shown in Fig. 1. Intracranial brain tissue was first divided into 5 regions, which were cortex, nuclei, white matter, ventricle and brain stem. The reason of distinguishing cortex, nuclei, ventricle and brain stem from white matter is that hyperintensities located in these regions are generally not included in the category of WMHs as suggested by STRIVE (Wardlaw et al., 2013b). We further segmented white matter into 25 sub-regions based on their ontological relationships (Glasser et al., 2016; Puelles et al., 2013; Wu et al., 2019, 2016) and the location of major white matter tracts (Hua et al., 2008; Zhang et al., 2010).

2.3. ^{A}U-Net for WMH segmentation

The pipeline of ^{A}U-Net framework is shown in Fig. 2. For a target subject, the brain atlas was transformed to the native space by coregistration between T1 images using symmetric diffeomorphic image registration algorithm (Avants et al., 2008), which is built in Advanced Normalized Tools (ANTs). Since intra-subject coregistration had been performed between T1 and T2-FLAIR, all 29 ROIs were immediately defined on T2-FLAIR.

For the calculation of anatomical-based spatial features, we merged bilateral white matter regions as one ROI, resulted in 14 symmetric white matter regions. Modified prior probabilities of WMH occurrence (prob_m) were calculated and assigned to 18 brain regions (anterior deep white matter, posterior deep white matter, inferior deep white matter, body of corpus callosum, splenium of corpus callosum, genu of corpus callosum, anterior limb of internal capsule, posterior limb of internal capsule, frontal subcortical white matter, temporal subcortical

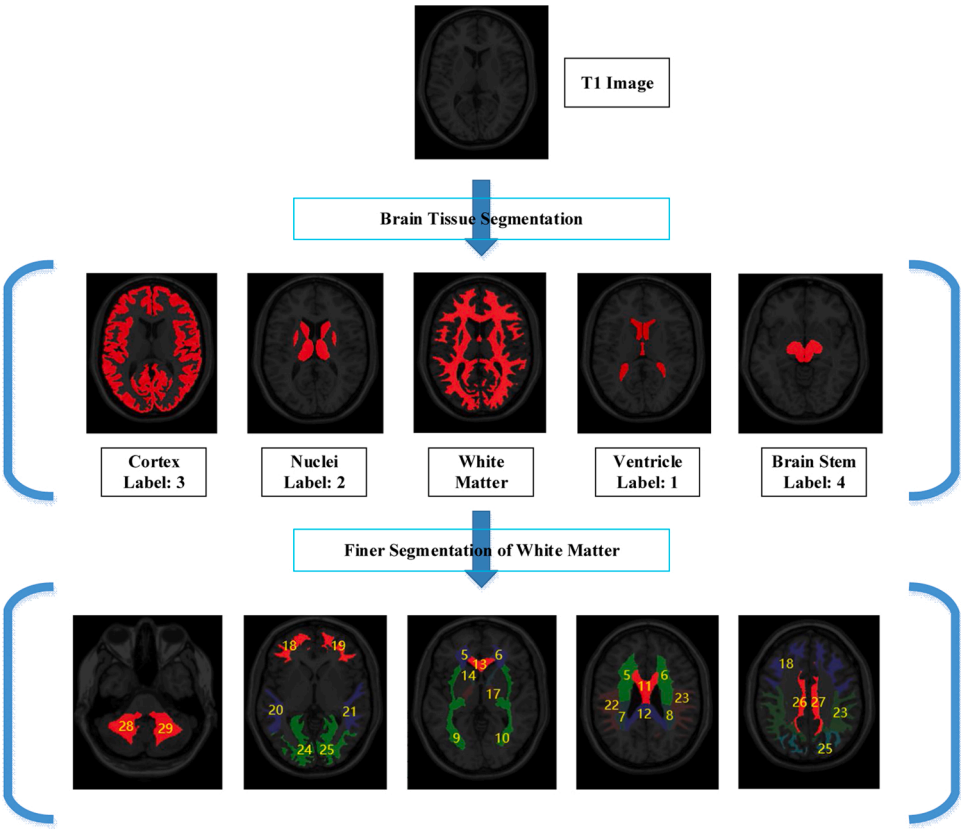

Fig. 1. Hierarchical definition of brain regions in atlas. Labels (1-29) were assigned to each brain regions for identification. 1: ventricle, 2: nuclei, 3: cortex, 4: brain stem. Hyperintensities located in regions 1-4 are not considered as WMHs suggested by (Wardlaw et al., 2013b). 5-6: Anterior dWM, 7-8: posterior dWM, 9-10: inferior dWM, 11: body of corpus callosum, 12: splenium of corpus callosum, 13: genu of corpus callosum, 14-15: anterior limb of internal capsule, 16-17: posterior limb of internal capsule, 18-19: frontal sWM, 20-21: temporal sWM, 22-23: parietal sWM, 24-25: occipital sWM, 26-27: limbic white matter, 28-29: cerebellum WM. Abbreviations: dWM: deep white matter, sWM: subcortical white matter.

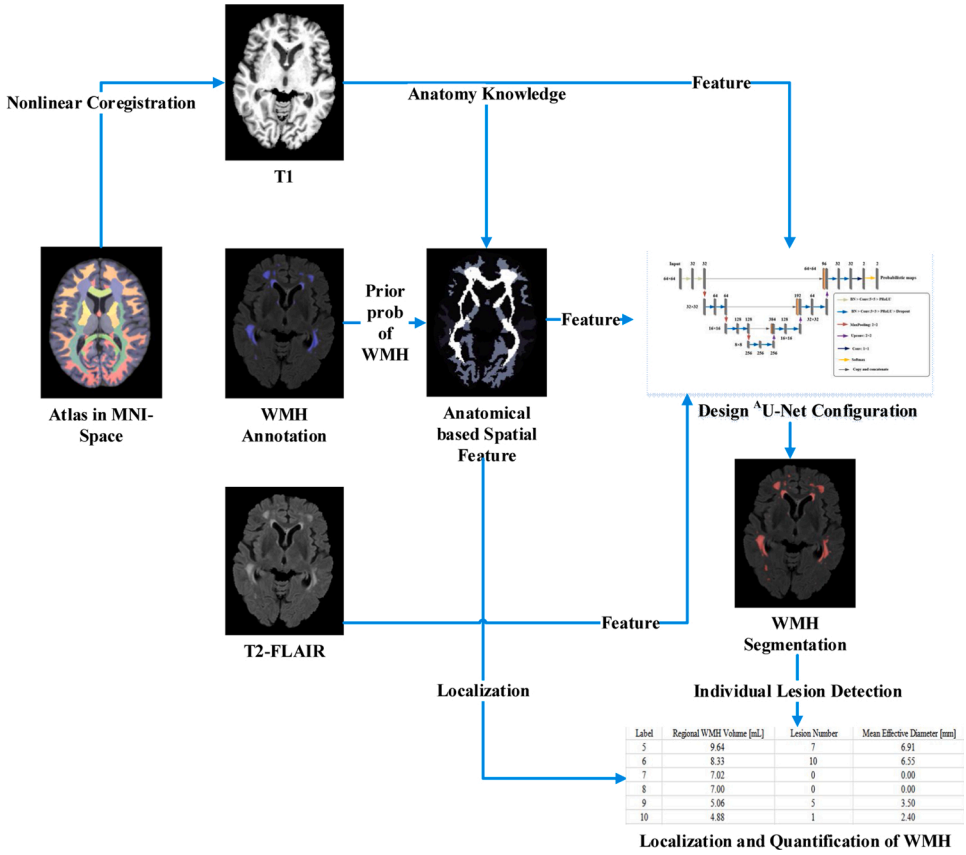

Fig. 2. Flowchart of anatomical knowledge-based MRI deep learning pipeline (AU-Net).

white matter, parietal subcortical white matter, occipital subcortical white matter, limbic white matter, cerebellum white mater, ventricle, nuclei, cortex and brain stem) using Eq. (1).

$$\text{prob\_}m = \frac{I(V_{WMH} > threshold)}{\text{NS}} \tag{1}$$

In this equation, $\mathbf{V}_{\text{WMH}}$ is the normalized WMH volume in each of the 18 brain regions, which was calculated by Eq. (2). Indication function **I** (.) returns one if $\mathbf{V}_{\text{WMH}}$ is larger than the threshold. **NS** is the number of subjects used in the training set. We used leave-one-out cross-validation in the evaluation process, so for each iteration, 59 subjects were used to calculate the modified prior probabilities (**NS** = 59). In the scenario of final modal training, the whole MICCAI 2017 data set was split into training set and validation set, 50 subjects were used in training process, so **NS** should be set to 50.

$$\text{V}_{\text{WMH}} = \frac{\text{volume of WMH in specific brain region}}{\text{volume of this specific brain region}} \tag{2}$$

In our study, the modified prior probabilities were used for spatial encoding of 18 brain regions. In order to separate different regions as far as possible through encoding, we should maximize the code distance denoted by **D** in Eq. (3). For optimization, we traversed through set of possible threshold values range from 0 to 1 with step: 0.0001 to find the maximum **D**.

$$\text{D} = \frac{\sum_{i}\sum_{j\neq i}\sqrt{(\text{prob\_}m_i - prob\_m_j)^2}}{\text{C}_N^2} \tag{3}$$

In Eq. (3), $\mathbf{prob\_m_i}$, $_j$denote the modified prior probabilities associated with brain region i and j. The denominator is to calculate how many combinations we could have if we chose 2 elements from a set with size N (N represents the total number of predefined brain regions: 18), as shown in Eq. (4), where N, M are both integers and N should be larger than M.

$$\text{C}_{\text{N}}^{\text{M}} = \frac{\text{N!}}{\text{M!(N} - \text{M)!}} \tag{4}$$

Anatomical based spatial features were generated after modified prior probabilities were assigned to 18 brain regions, then we saved them as images with the same dimension as T2-FLAIR and T1 after coregistration (as shown in Fig. 2).

The basic U-Net architectures (Ronneberger et al., 2015) used for segmentation are shown in supplementary Fig. S1 along with detail parameters for training. Two configurations for $^{\text{A}}$U-Net were evaluated to investigate the optimal approach of including anatomical-based spatial features. The first approach is to use two connective U-Net (represents as $^{\text{A}}$U-Net-two-step) with T1 and T2-FLAIR intensities fed into the first U-Net for rough segmentation, then the output lesion probabilistic maps were concatenated with anatomical-based spatial features and fed into the second U-Net for refine segmentation. For the second approach, we simply concatenated anatomical-based spatial features with intensities and fed them into the first U-Net for lesion segmentation (represents by $^{\text{A}}$U-Net-Spatial). Probabilistic thresholds for generating binary lesion maps were determined using leave-one-out cross-validation by optimizing Dice similarity coefficient (DSC).

2.4. Evaluation, ablation study and final model training

Two independent human observers with more than 2 years of research experience in the realm of WMH segmentation manually annotated 60 subjects in MICCAI 2017 dataset. The lesions were fully manually traced using the interactive software package Display (https://github.com/BIC-MNI), which was developed at the McConnell Brain Imaging Center of the Montreal Neurological Institute.

We evaluated the following methods for WMH segmentation on MICCAI 2017 dataset: 1) two independent human observers, 2) $^{\text{A}}$U-Net-two-step, and 3) the top ranked method (sysu_media) in WMH segmentation challenge MICCAI 2017 (state-of-the-art) (Kuijf et al., 2019; Li et al., 2018). Leave-one-out cross-validation was used based on following criteria: 1) DSC, 2) Modified Hausdorff distance (95th percentile) (H95), 3) Sensitivity for detecting individual lesions (recall), 4) F1-score for detecting individual lesions (F1), 5) precision-recall curve and area under the curve (AUC). Calculation of these criteria were based on definitions in (Kuijf et al., 2019). Paired *t*-test was used to detect significant differences between performances of WMH segmentation methods. False discovery rate (FDR) correction was used to correct for multiple comparison.

We performed an ablation study to show the importance of each elements in our $^{\text{A}}$U-Net pipeline. Leave-one-out cross-validation was used. Four configurations were evaluated: 1) typical U-Net configuration (Fig. S1 A)) with T1 and T2-FLAIR intensities as input features, denoted by U-Net-Intens, 2) $^{\text{A}}$U-Net-Spatial, 3) Stage-wise configurations: $^{\text{A}}$U-Net-two-step, and 4) the same network architecture as $^{\text{A}}$U-Net-two-step, but spatial features were concatenated with intensities in the input layer of the first U-Net, denoted by $^{\text{A}}$U-Net-Spatial + refine.

The training dataset was further divided into three groups: small WMH group with total WMH volume <5 mL (mL) (22 subjects), medium WMH group with total WMH volume between 5–15 mL (8 subjects), and large WMH group with total WMH volume larger than 15 mL (30 subjects). Methods were also evaluated in groups with different WMH burden. Computational efficiency of $^{\text{A}}$U-Net-two-step, $^{\text{A}}$U-Net and U-Net-Intens configurations were evaluated by time cost for preprocessing, training an epoch and segmenting a new subject.

The whole MICCAI 2017 dataset consists of 60 subjects, which was split into training set and validation set for final model training. Ten subjects were randomly chosen to form a validation set (small WMH group: 4 subjects, medium WMH group: 2 subjects and large WMH group: 4 subjects). The loss function is defined in Eq. (5), both training loss and validation loss were monitored during final model training procedure.

$$\text{Loss} = 1 - \frac{2\sum_{\text{i}=1}^{\text{N}}|\text{p}_\text{i}\circ g_i| + s}{\sum_{\text{i}=1}^{N}(|p_i| + |g_i|) + s} \tag{5}$$

Where G = {g_1, …, g_N} are the ground-truth lesion maps (golden standard), and P = {p_1, …, p_N} are the predicted probabilistic maps over N patches, ° represents the element-wise product of two matrices, |.| represents the sum of every elements in the matrix. The s term is used to ensure the stability of loss function by avoiding division by 0, and it is set to 1e-6 in our models.

2.5. Investigate relationships between WMH burden and cognition

We segmented WMHs in ADNI dataset using the optimal method evaluated by cross-validation. Then we detected connected components found in the binary lesion map. For each individual WMH lesion, voxel by voxel labels were identified, the statistical mode of voxel-based labels was used to denote the location of this lesion. Following measurements were calculated for each subject: number of individual WMH lesions, effective diameters of individual lesions (calculated by cubic root of individual lesion volumes, in mm), and WMH volumes (in mL). Given that measurements for WMHs tend to have a positive skewed distribution (Habes et al., 2016; Kuijf et al., 2019; Wu et al., 2019), we performed log transformation to achieve approximate normal distribution.

We used partial correlation to evaluate:

1) Associations between age and regional WMH volumes (log-transformed), adjusted for gender and education years.

2) Associations between regional WMH volumes (log-transformed) and neuropsychological scores (MMSE, FAQ and ADAS), adjusted for

gender, education years and age.

3) Associations between number of regional individual WMH lesions and neuropsychological scores (MMSE, FAQ and ADAS), adjusted for gender, education years, age and total WMH volume.

4) Associations between mean effective diameters of regional individual WMH lesions and neuropsychological scores (MMSE, FAQ and ADAS), adjusted for gender, education years, age and total WMH volume.

$p < 0.05$ was considered significant after correcting for multiple comparison (FDR correction). All statistical analysis was performed using R version 3.6.2.

3. Results

3.1. Anatomical-based spatial features

In the final model training procedure, as shown in Fig. 3 A), we traversed through set of possible threshold values (Eq. (1)) range from 0 to 1 with step: 0.0001 to find the maximum distance D in Eq. (3). The threshold values in Fig. 3 A) are log transformed, maximum distance D was found when threshold equals to $10^{-2.51}$ (0.0031). Modified prior probabilities for 18 brain regions are shown in Fig. 3 B), anterior deep white matter and splenium of corpus callosum suffered from the highest WMH burdens and were both assigned to number of 0.92. Modified prior probabilities for brain stem, ventricle and cortex are rather small (0.02, 0.03 and 0.02), which is consistent with the definition of WMHs suggested by STRIVE (Wardlaw et al., 2013b). Modified prior probability for cerebellum white matter was zero, we assigned 0.01 to this region instead. The relatively large modified prior probabilities associated with nuclei (0.53) might be due to mismatch in coregistration or large WMH burdens within periventricular areas.

3.2. ^{A}U-Net segmentation

3.2.1. Comparison with human observers and state-of-the-art

Table 2 shows the evaluation of WMH segmentation using leave-one-out cross-validation. The performance of state-of-the-art method (sysu_media) was directly quoted from supplementary materials of (Li et al., 2018). Our method (^{A}U-Net-two-step) achieved best performance in terms of H95, and comparable good performance when compared with state-of-the-art method based on DSC and F1. No significant difference exists between our proposed method and sysu_media detected by paired *t*-test. Both our method and sysu_media outperformed independent human observers based on DSC, H95 and F1. State-of-the-art method and human observer 1 achieved best performance in terms of recall for detecting individual WMH lesions.

Table 2
Performance of our CNN models, compare with MICCAI-2017 top ranked method and independent human observers.

Algorithm	DSC	H95 (mm)	Recall	F1
Sysu_media	**0.87 ± 0.10**	3.62 ± 5.88	**0.85 ± 0.08**	**0.83 ± 0.09**
^{A}U-Net-two-step	0.86 ± 0.09	**3.06 ± 4.11**	0.83 ± 0.10	0.82 ± 0.11
Human Observer 1	0.82 ± 0.08 *	4.17 ± 3.18	**0.85 ± 0.12**	0.81 ± 0.13
Human Observer 2	0.82 ± 0.12 *	4.25 ± 4.07	0.83 ± 0.09	0.81 ± 0.10

*: $p < 0.05$, compared with sysu_media. Boldface indicates the best performance.

3.2.2. Ablation study

As shown in Table 3, the incorporation of anatomical-based spatial features largely increased the segmentation performance (^{A}U-Net-Spatial compared with U-Net-Intens), significant improvements were detected based on DSC and F1. The performance of ^{A}U-Net-Spatial + refine achieved better performance compared with U-Net-Intens based on DSC, H95 and F1, however, no significant difference was detected. ^{A}U-Net-Spatial + refine had worse performance compared with ^{A}U-Net-Spatial, which might be due to overfitting problem. At last, ^{A}U-Net-two-step achieved best performance among all configurations,

Table 3
Ablation study for ^{A}U-Net structure.

Algorithm	DSC	H95 (mm)	Recall	F1
^{A}U-Net-two-step	**0.86 ± 0.09** *,^	**3.06 ± 4.11** *,^	**0.83 ± 0.10** *,^	**0.82 ± 0.11** *,^
^{A}U-Net-Spatial + refine	0.81 ± 0.14	4.70 ± 5.32	0.75 ± 0.16	0.77 ± 0.17
^{A}U-Net-Spatial	0.84 ± 0.08 *	4.26 ± 5.23	0.80 ± 0.12	0.79 ± 0.13 *
U-Net-Intens	0.78 ± 0.16	5.57 ± 6.67	0.76 ± 0.15	0.73 ± 0.19

*: $p < 0.05$, compared with U-Net-Intens.
^: $p < 0.05$, compared with ^{A}U-Net-Spatial + refine. Boldface indicates the best performance.

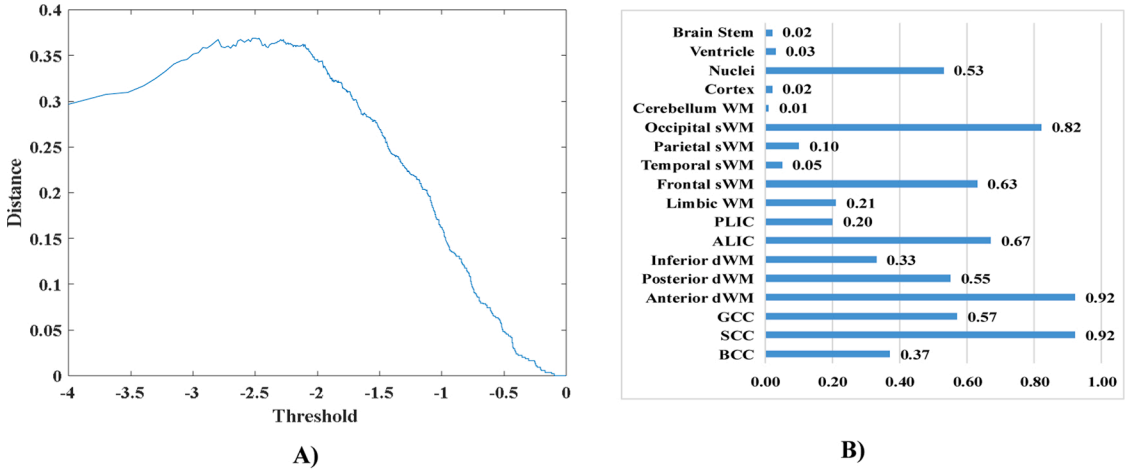

Fig. 3. Generate anatomical-based spatial features in final model training. A) Relationship between distance and the threshold of normalized WMH volume (log-transformed). B) Modified prior probabilities of WMH occurrence for 18 brain regions. Abbreviations: sWM: subcortcal whit matter, dWM: deep white matter, ALIC: anterior limb of internal capsule, PLIC: posterior limb of internal capsule, BCC: body of corpus callosum, SCC: splenium of corpus callosum, and GCC: genu of corpus callosum.

significant improvements were detected based on DSC, H95, recall and F1, when compared with U-Net-Intens and ^{A}U-Net-Spatial + refine.

3.2.3. Performance in cohorts with different WMH burden

We calculated precision-recall curves of ^{A}U-Net-two-step, ^{A}U-Net-Spatial and U-Net-Intens during cross-validation procedure. Fig. 4 A) shows the results on small and medium WMH group, AUC are 0.82, 0.80, and 0.70 respectively. Fig. 4 B) shows results on large WMH group, AUC are 0.88, 0.89, and 0.80 respectively. ^{A}U-Net largely improved WMH segmentation compared to U-Net-Intens with 14 (^{A}U-Net-Spatial compared with U-Net-Intens) to 17% (^{A}U-Net-two-step compared with U-Net-Intens) increase based on AUC in small and medium WMH burden group. ^{A}U-Net-two-step achieved the best performance in small and medium WMH burden group, whereas ^{A}U-Net-Spatial outperformed others in cohort with large WMH burden.

Fig. 5 shows some examples of WMH segmentation using proposed ^{A}U-Net-two-step and U-Net-Intens configuration, compared with golden standard. Hyperintensities located in cerebellum gray matter are not considered as WMHs in manually delineation as shown in the first row, proposed ^{A}U-Net-two-step successfully excluded them from WMHs whereas U-Net-Intens over-segmented these regions. The second row in Fig. 5 shows a slice with small WMH burden. Proposed ^{A}U-Net-two-step detected all individual WMH lesions with one false positive happened in the region of left parietal subcortical white matter (identified by region 1), however, U-Net-Intens had more trouble detecting small individual lesions, such as false negatives in regions of left parietal subcortical white matter (region 2) and left anterior deep white matter (region 3). The third row shows a slice with medium WMH burden. ^{A}U-Net-two-step achieved almost perfect segmentation. U-net-Intens architecture successfully detected individual lesion in region 1, however, false positives appeared in regions of insular cortex (region 2) and caudate nucleus (region 3).

3.2.4. Computational efficiency

Preprocessing, training, and testing were performed on workstation in a Linux server with 32 Intel Core i7−6700 @ 3.4 GHz and NVIDIA GeForce RTX 2080 Ti 11GB. In the training set, over 100,000 patches with size 64 * 64 were generated. Time cost of training each epoch for ^{A}U-Net-two-step were calculated by adding the time duration in rough segmentation and refine segmentation. As shown in Table 4, the incorporation of anatomical-based spatial features caused only 2.31% increase for training one epoch (^{A}U-Net-Spatial compared with U-Net-Intens). However, using stage-wise CNN to include spatial features largely increased the mean time cost for training one epoch, our proposed model ^{A}U-Net-two-step spent 61.76% more time in training process compared with U-Net-Intens configuration. When segmenting a new instance, ^{A}U-Net-two-step took slightly longer time compared with U-Net-Intens (increased from 7.61 s to 9.18 s). In order to generate anatomical-based spatial features for new instance, we used symmetric diffeomorphic image registration algorithm (Avants et al., 2008), which is built in Advanced Normalized Tools (ANTs). For an image with size of: (240, 240, 83), the coregistration process took about 78.16 s.

3.2.5. Training final CNN model

Proposed ^{A}U-Net-two-step method achieved the best performance in cross-validation procedure, and was selected as our final model in this study. Loss on training set and validation set were calculated associated with each epoch. As shown in Fig. 6 A), after 40 epoches, we achieved minimal loss on validation set for rough segmentation. Then the output probabilistic maps were fed into a second U-Net combined with anatomical-based spatial features for refine segmentation. As shown in panel B), parameters in the 42th epoch were chosen.

The final performance of ^{A}U-Net-two-step on validation set were: DSC: 0.83 ± 0.18, H95: 4.16 ± 4.11, Recall: 0.81 ± 0.15, and F1: 0.79 ± 0.12.

3.3. Patterns of WMH burden associated with cognition

As shown in Table 1, no group difference existed based on gender, age and education years among CN, MCI and AD. Significant group differences were detected based on neuropsychological sores (MMSE, ADAS and FAQ scales) ($p < 0.0001$).

As shown in Table 5, different WMH accumulating patterns (measured by WMH volumes) were detected associated with normal aging and cognitive impairments. Age-associated WMH volumes were located in deep white matter regions (splenium of corpus callosum, left anterior and posterior deep white matter). Whereas, regional WMH volumes significantly correlated with cognitive impairments measured by ADAS, MMSE or FAQ were detected in: bilateral anterior deep white matter, left posterior deep white matter, bilateral inferior deep white matter, splenium of corpus callosum, bilateral parietal subcortical white matter, and bilateral occipital subcortical white matter.

As shown in Fig. 7 A)-C), after adjustment for gender, age, education and total WMH volume, number of individual WMH lesions located in right frontal subcortical white matter, right parietal subcortical white matter, and corpus callosum were positively correlated with ADAS and FAQ. Lesion number in right frontal subcortical white matter was negatively correlated with MMSE.

Mean effective diameters of individual lesions located in right parietal subcortical white matter and genu of corpus callosum presented significant positive correlation with cognitive impairments measured by ADAS or FAQ (adjusted for gender, age, education and total WMH

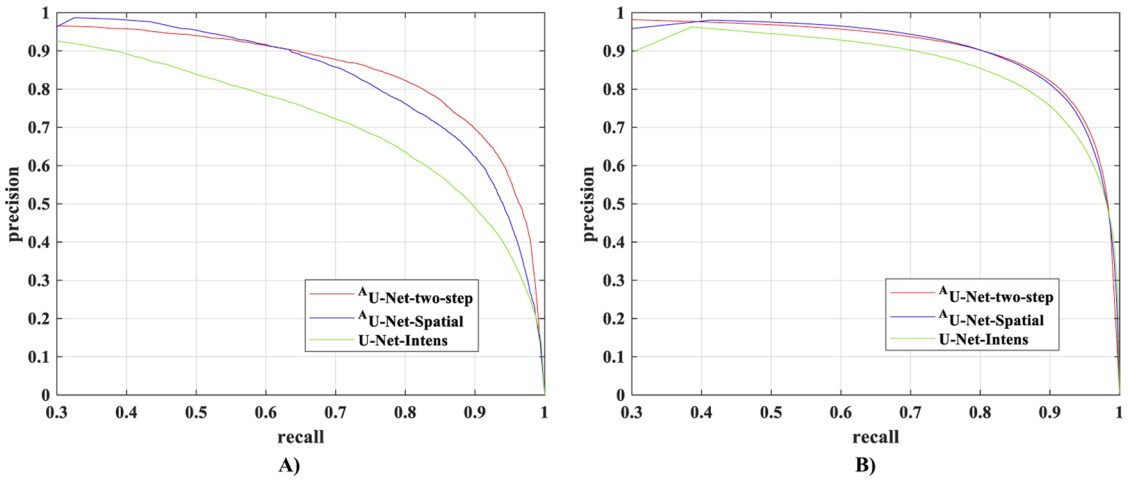

Fig. 4. Precision-recall curves of WMH segmentation methods evaluated on A) small and medium WMH group, and B) large WMH group.

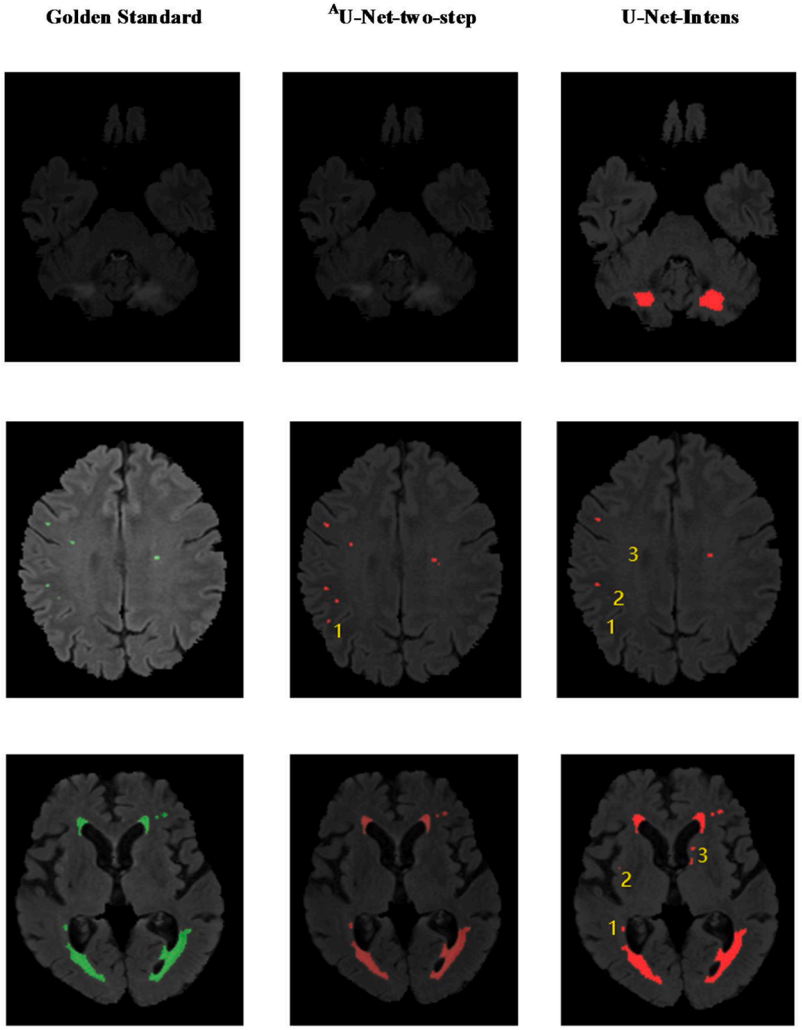

Fig. 5. Examples of WMH segmentation results. Each row shows the same slice. The first row: slice with none WMH burden, the second row: slice with small WMH burden, and the third row: slice with medium WMH burden. Left column shows manual delineation of WMHs overlaid on T2-FLAIR, middle column shows results of proposed ^{A}U-Net-two-step configuration, and right column shows results of U-Net-Intens.

Table 4
Computational efficiency of WMH segmentation methods.

Algorithm	Nonlinear coregistration [seconds]	Training for one epoch [seconds]	Segmenting new instance [seconds]
^{A}U-Net-two-step	78.16	585.03 ± 7.50	9.18
^{A}U-Net-Spatial	78.16	370.03 ± 8.20	7.90
U-Net-Intens	–	361.67 ± 4.40	7.61

volume), as shown in Fig. 7 D)-E).

4. Discussion

4.1. Performance of ^{A}U-Net

In this paper, we proposed an anatomical knowledge-based MRI deep learning pipeline (^{A}U-Net) by integrating handcrafted anatomical-based features with well-designed U-Net configurations to simultaneously segment and locate WMHs. ^{A}U-Net largely improved WMH segmentation (compared with the U-Net-Intens configuration), with a 14–17% increase based on AUC in small and medium WMH burden group. This is preferable because WMHs at early stage of neurodegeneration tend to have sparse distributions and small volumes (Habes et al., 2018; Rachmadi et al., 2018; Wardlaw et al., 2013a). Our method achieved the best performance in terms of H95, and a comparable good performance based on DSC and F1-score, when compared with state-of-the-art method in MICCAI 2017 (Kuijf et al., 2019; Li et al., 2018). We used paired *t*-test to detect significant differences between performances of WMH segmentation methods. No significant difference existed between our proposed method and state-of-the-art method. However, our method was able to simultaneously locate WMHs within predefined anatomical brain regions after segmentation, whereas most methods in MICCAI-2017 are not capable of quantification for anatomical-based lesion spatial distributions (Kuijf et al., 2019). Our method outperformed independent human observers based on DSC, H95 and F1 score.Areas for improvement should be focused on the recall metrics for detecting individual lesions, of which human observers and state-of-the-art method had better performance. Moreover, the proposed ^{A}U-Net pipeline can be applied to other deep learning-based methods with the expectation of improving WMH segmentation performance.

4.2. Anatomical-based spatial features

The inclusion of spatial features in WMH segmentation tasks has been previously proposed by several studies (Ghafoorian et al., 2017;

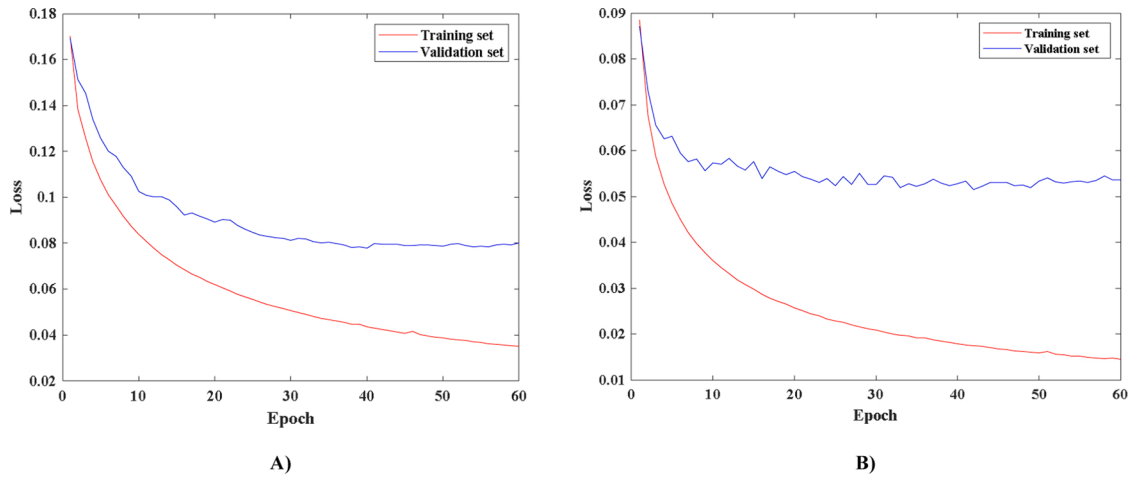

Fig. 6. Loss associated with each epoch in the final model training procedure: A) for the first step: rough segmentation and B) for the second step: refine segmentation.

Table 5
Correlation coefficients (r) between regional WMH volumes (in mL), age and neuropsychological sores, only regions with significant correlation are listed.

White matter regions	r between WMH volumes and age	r between WMH volumes and ADAS	r between WMH volumes and MMSE	r between WMH volumes and FAQ
Anterior dWM (L)	0.26*	0.18*	−0.18**	0.16*
Anterior dWM (R)	0.17	0.20*	−0.18**	0.21**
Posterior dWM (L)	0.21*	0.14	−0.17*	0.16*
Inferior dWM (L)	0.15	0.20*	−0.23**	0.17*
Inferior dWM (R)	0.20	0.15*	−0.16*	0.13
SCC	0.22*	0.19*	−0.23**	0.21**
Parietal sWM (L)	0.18	0.13	−0.19**	0.16*
Parietal sWM (R)	0.06	0.11	−0.16*	0.18*
Occipital sWM (L)	0.16	0.17*	−0.21**	0.19*
Occipital sWM (R)	0.23	0.15*	−0.20**	0.16

Abbreviations: dWM: deep white matter, sWM: subcortical white matter, SCC: splenium of corpus callosum, ADAS: Alzheimer's Disease Assessment Scale, MMSE: Mini Mental State Examination, FAQ: Functional Assessment Questionnaire.
*: $p < 0.05$.
**: $p < 0.01$.

Griffanti et al., 2016; Rachmadi et al., 2018). Spatial coordinates in the MNI space and voxel-based prior probabilistic templates of WMHs are commonly used approaches for incorporating spatial features. Several studies have conformed that significant improvement would achieve in segmenting WMH after incorporating spatial features into CNN (Ghafoorian et al., 2017; Rachmadi et al., 2018). The inclusion of spatial features is based on the idea that some regions of the brain are more likely to suffer from a higher WMH burden than others. However, this assumption might be biased because the prior probabilities are mainly determined by large continuous WMHs. Due to the fact that large WMHs are typically located in the periventricular and deep white matter regions, it may be difficult for location-sensitive methods to segment small individual lesions that are located in subcortical areas (Kuijf et al., 2019). The use of coordinates and voxel-based prior probabilities may also suffer from missed coregistration problems and high computational burdens. In our study, we introduced anatomical-based spatial features, modified prior probabilities of WMH occurrence were assigned to 18 brain regions by setting a threshold for normalized WMH volumes. This approach makes our network less sensitive to specific locations while still benefitting from additional anatomical knowledge. This approach is also less sensitive to missed coregistration. Anatomical-based spatial features have the same dimension as the intensities that are extracted from T1 and T2-FLAIR image, which, as shown in our results, did not dramatically increase the computational burden in either training or testing procedures. However, when stage-wise network configuration is used for incorporating spatial features, a large increase in time cost would occur during model training. Furthermore, as the spatial features are handcrafted, a nonlinear coregistration procedure is needed for preprocessing, which leads to a major increase in the computational burden when segmenting a new instance. New methods for nonlinear coregistration have been developed in recent years, such as diffeomorphic Demons algorithm and deep learning-based methods (Dalca et al., 2018). The time cost for running these algorithms on a GPU can be reduced to less than 30 swithout strongly compromising the accuracy. These methods should be used to improve the performance of our method in the future.

4.3. Stage-wise network configuration

Recently, several deep learning-based frameworks were specifically designed for fusing the features from multimodality data and were successfully applied to dementia diagnosis and MR image synthesis (Zhou et al., 2020, 2019). In these frameworks, stage-wise configurations were designed to solve the problems of heterogeneous data and sample limitations. In each stage, a deep neural network was designed to learn the latent data representations for different combinations of modalities by using the maximum number of available samples. Through stage-wise configuration, the network should be able to exploit the underlying correlations among multiple modalities, while also capturing modality specific information (Zhou et al., 2020). Our proposed ^{A}U-Net-two-step configuration presents similar idea for dealing with heterogeneous features. For WMH segmentation in this paper, two heterogeneous features were involved: intensities extracted from MRI images and handcrafted anatomical-based spatial features. Intensities are voxel-based features whereas spatial features are ROI-based, they have different ranges (the intensity of a voxel labeled as WMH can be larger than 4 after normalization, whereas spatial features have a range of [0, 1]) and distributions. Moreover, these two features have different levels of discriminative ability for lesion segmentation: intensity features play a major role in identifying WMHs, whereas spatial features

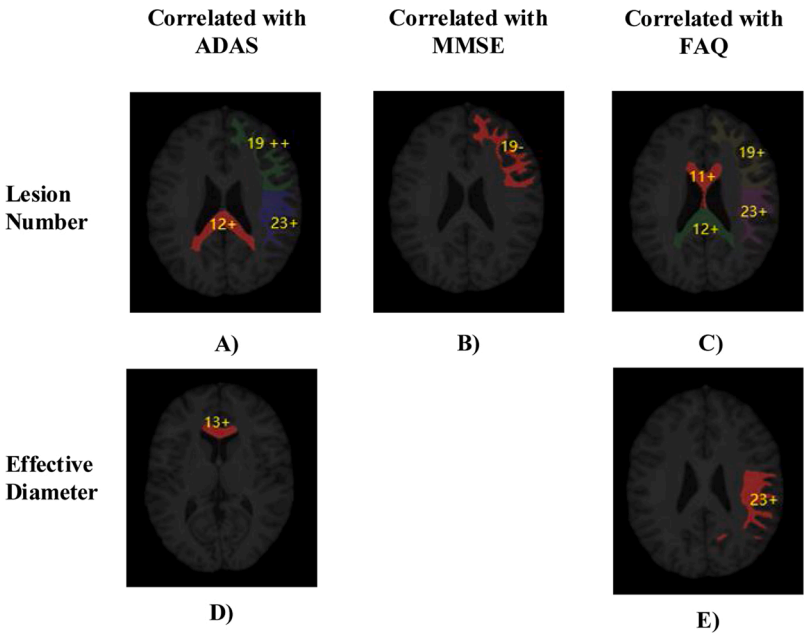

Fig. 7. Associations between the characteristics of individual WMH lesions and cognitive impairments. Regional lesion numbers were significantly correlated with A) ADAS in regions of: right frontal (label 19), right parietal subcortical white matter (label 23) and splenium of corpus callosum (label 12), B) MMSE in regions of right frontal subcortical white matter, C) FAQ in regions of: right frontal, right parietal subcortical white matter, body of corpus callosum (label 11) and splenium of corpus callosum. Mean effective diameters of WMH lesions were significantly correlated with D) ADAS in genu of corpus callosum (label 13), E) FAQ in right parietal subcortical white matter. +: positive correlation with $p < 0.05$, ++: positive correlation with $p < 0.01$, -: negative correlation with $p < 0.05$. Abbreviations: ADAS: Alzheimer's Disease Assessment Scale, MMSE: Mini Mental State Examination, FAQ: Functional Assessment Questionnaire.

should provide complementary information for difficult decisions such as detecting subcortical small lesions and avoiding false positives in typical areas (gray matter, periventricular areas and corpus callosum). To resolve the heterogeneity problem, we designed a two-step U-Net framework. In the first stage, our network was trained by using intensities that were extracted from T1 and T2-FLAIR patches, after which the high-level features were combined with anatomical-based spatial features for the training of a second U-Net in stage 2. As shown by our results, the ^{A}U-Net-two-step configuration achieved much better results, when compared with direct concatenation of intensities and spatial features in the input layer (such as ^{A}U-Net-Spatial, and ^{A}U-Net-Spatial + refine), which indicates that the stage-wise deep neural network configuration may be a powerful framework for resolving problems of multimodality data fusion.

4.4. Cognition associated WMH burden

In this study, we found different WMH accumulation patterns (measured by WMH volumes) that were associated with normal aging and cognitive impairments. Age-associated WMH accumulation was mainly located in regions of deep white matter, whereas cognition-associated WMH burdens infected more functional regions (ranging from deep white matter to subcortical white matter). These results are consistent with results from previous studies (Frey et al., 2019; Habes et al., 2016, 2018). Recent studies have also shown that WMH burden in strategic white matter tracts, such as the forceps minor, forceps major, superior longitudinal fasciculus or anterior thalamic radiation (ATR), is more relevant in explaining the variance in cognitive functioning, when compared to global WMH volume (Biesbroek et al., 2017). As these fiber tracts are typically distributed in regions of anterior deep white matter, and corpus callosum, as well as frontal, occipital and parietal subcortical white matter, the significant correlation between lesion distributions and cognition that was detected by our methods may be explained by the injury of the specific fiber connectivity. An epidemiological study confirmed that, within elderly and neurodegenerative cohorts, the majority of WMHs are individual small lesions with effective diameters <=3 mm (Ghafoorian et al., 2015). Moreover, in longitudinal research, individual small WMH lesions have been observed to grow over time (Frey et al., 2019; Ghafoorian et al., 2015), thus, individual WMH located in strategic white matter tracts should be an important risk factor for future cognitive impairments. Using our methods, the growth of individual WMH lesions within important brain regions can be automatically traced by measuring the number and effective diameters of lesions, which should be important for the further exploration of lesion-symptom mapping relationships in neurodegeneration.

Regional WMH accumulation may be a relatively severe consequence of the continuous and chronic pathological progress of white matter tract injury (Prins and Scheltens, 2015). A longitudinal study in 2018 concluded that white matter damage can be detected by diffusion tensor imaging (DTI) before the appearance of WMH lesions (van Leijsen et al., 2018). By providing finer spatial resolution utilizing multimodality MR imaging technique, researchers have also explored the spatial relationship between diffusion characteristics along white matter tracts and the distance from WMHs (Reginold et al., 2018). Several clinical neuroimaging biomarkers were found to be associated with neurodegeneration, including: small subcortical infarctions, lacunes of presumed vascular origins, WMHs of presumed vascular origins, perivascular spaces, cerebral microbleeds, and brain atrophy (Wardlaw et al., 2013b). These biomarkers may be interrelated in their development because they share similar risk factors and etiologies; additionally, they may have both combined and independent impacts on cognitive functioning (Prins and Scheltens, 2015). For example, WMH accumulation in certain strategic regions can cause the aggravation of cortical neurodegeneration over time, as was proposed by a study in 2018 (Habes et al., 2018). These results suggest that an AD-related neurodegeneration spectrum may exist that combines the interrelation and coupling of pathologies that are observed in multimodality neuroimaging. Subtle impairments of white matter tracts caused by chronic ischemia or other risk factors may be present in the early phase. Regional WMH lesions may reflect the accumulation effect of impairments in fiber connectivity, and are further involved in neurodegeneration by triggering or aggravating other related cerebral pathologies.

4.5. Limitations and future directions

Several limitations of this study should be noted. First, the manually annotated training set only consisted of 60 subjects, which is rather small for the deep learning-based method. Transfer learning can be used to resolve this problem in the future. Second, we only used the U-Net based architecture for the design of segmentation framework. Although we confirmed that anatomical knowledge can improve WMH

segmentation, the optimal network configurations should be further developed. Third, for the investigation of cognition-associated WMH burden, novel and multiple neuroimaging techniques should be used to characterize the location of WMHs within white matter network. Finally, as the majority of subjects present a small to medium WMH burden, especially in the stage of prodromal neurodegeneration (Biesbroek et al., 2017; Prins and Scheltens, 2015), in order to assess the associations between lesion locations and cognitive impairments of specific domains, large multicenter studies that include thousands of individuals are needed to provide coverage in rarely affected brain regions and to provide powerful generalizations. Longitudinal studies are also important for relating new incident lesions in strategic regions with cognitive decline, as well as for exploring their interactions with other pathologies in neurodegeneration, such as brain atrophy.

5. Conclusion

In this study, we proposed ^{A}U-Net framework to simultaneously segment and quantify WMH burden. This method significantly improved WMH segmentation performance, especially in cohorts with small and medium WMH burdens, and achieved comparable good performance when compared with the state-of-the-art method. Different network architectures for incorporating anatomical-based spatial features were evaluated, a stage-wise configuration achieved the best performance. After segmentation, anatomical-based descriptions of WMH burdens can be immediately quantified. By applying ^{A}U-Net to ADNI dataset, we detected different patterns of WMH accumulation that were associated with normal aging and cognitive impairments. Our findings also suggest that the characteristics of individual WMH lesions were still significantly correlated with cognition after controlling for total lesion volumes. ^{A}U-Net can be applied to other deep learning-based methods with the expectation of improving WMH segmentation; more importantly, it may facilitate the identification of lesions in cognition evaluation studies, based on anatomical knowledge.

Funding and acknowledgements

This study was supported by grants from National Key Research and Development Program of China (2018YFC1312000), the National Natural Science Foundation of China (61801145), and the Basic Research Project of Shenzhen Science and Technology Program (JCYJ20180306171800589).

CRediT authorship contribution statement

Li Liang: Conceptualization, Formal analysis, Methodology, Writing - original draft. **Pengzheng Zhou:** Methodology, Validation, Formal analysis. **Wanxin Lu:** Methodology, Visualization. **Xutao Guo:** Software. **Chenfei Ye:** Visualization, Validation. **Haiyan Lv:** Software. **Wang Tong:** Funding acquisition, Supervision. **Ting Ma:** Conceptualization, Funding acquisition, Supervision, Data curation, Project administration, Writing - review & editing.

Declaration of Competing Interest

The authors report no declarations of interest.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.compmedimag.2021.101873.

References

Avants, B.B., Epstein, C.L., Grossman, M., Gee, J.C., 2008. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. Med. Image Anal. 12, 26–41.
Biesbroek, J.M., Weaver, N.A., Biessels, G.J., 2017. Lesion location and cognitive impact of cerebral small vessel disease. Clin. Sci. 131, 715–728.
Brickman, A.M., Tosto, G., Gutierrez, J., Andrews, H., Gu, Y., Narkhede, A., et al., 2018. An MRI measure of degenerative and cerebrovascular pathology in Alzheimer disease. Neurology 91, e1402–e1412.
Dadar, M., Maranzano, J., Ducharme, S., Collins, D.L., Alzheimer's Disease Neuroimaging, I, 2019. White matter in different regions evolves differently during progression to dementia. Neurobiol. Aging 76, 71–79.
Dalca, A.V., Balakrishnan, G., Guttag, J., Sabuncu, M.R., 2018. Unsupervised learning for fast probabilistic diffeomorphic registration. Med. Image. Comput. Assist. Intervention 2018 (Pt I 11070), 729–738.
Damulina, A., Pirpamer, L., Seiler, S., Benke, T., Dal-Bianco, P., Ransmayr, G., et al., 2019. White matter hyperintensities in alzheimer's disease: a lesion probability mapping study. J. Alzheimers Dis. 68, 789–796.
Frey, B.M., Petersen, M., Mayer, C., Schulz, M., Cheng, B., Thomalla, G., 2019. Characterization of white matter hyperintensities in large-scale MRI-Studies. Front. Neurol. 10.
Ghafoorian, M., Karssemijer, N., van Uden, I., de Leeuw, F.E., Heskes, T., Marchiori, E., Platel, B., 2015. small White matterlesiondetection in cerebralsmallvesseldisease. Medical Imaging 2015: Computer-Aided Diagnosis, 9414.
Ghafoorian, M., Karssemeijer, N., Heskes, T., van Uden, I.W.M., Sanchez, C.I., Litjens, G., et al., 2017. Location sensitive deep convolutional neural networks for segmentation of white matter hyperintensities. Sci. Rep. 7.
Glasser, M.F., Coalson, T.S., Robinson, E.C., Hacker, C.D., Harwell, J., Yacoub, E., et al., 2016. A multi-modal parcellation of human cerebral cortex. Nature 536, 171-+.
Griffanti, L., Zamboni, G., Khan, A., Li, L., Bonifacio, G., Sundaresan, V., et al., 2016. BIANCA (Brain Intensity AbNormality Classification Algorithm): a new tool for automated segmentation of white matter hyperintensities. Neuroimage 141, 191–205.
Habes, M., Erus, G., Toledo, J.B., Zhang, T., Bryan, N., Launer, L.J., et al., 2016. White matter hyperintensities and imaging patterns of brain ageing in the general population. Brain 139, 1164–1179.
Habes, M., Sotiras, A., Erus, G., Toledo, J.B., Janowitz, D., Wolk, D.A., et al., 2018. White matter lesions: spatial heterogeneity, links to risk factors, cognition, genetics, and atrophy. Neurology 91, e964–e975.
Hua, K., Zhang, J., Wakana, S., Jiang, H., Li, X., Reich, D.S., et al., 2008. Tract probability maps in stereotaxic spaces: analyses of white matter anatomy and tract-specific quantification. Neuroimage 39, 336–347.
Jenkinson, M., Smith, S., 2001. A global optimisation method for robust affine registration of brain images. Med. Image Anal. 5, 143–156.
Jenkinson, M., Bannister, P., Brady, M., Smith, S., 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage 17, 825–841.
Kuijf, H.J., Casamitjana, A., Collins, D.L., Dadar, M., Georgiou, A., Ghafoorian, M., et al., 2019. Standardized assessment of automatic segmentation of White Matter Hyperintensities and results of the WMH segmentation challenge. IEEE Trans. Med. Imaging 38, 2556–2568.
Lee, S., Viqar, F., Zimmerman, M.E., Narkhede, A., Tosto, G., Benzinger, T.L.S., et al., 2016. White matter hyperintensities are a core feature of alzheimer's disease: evidence from the dominantly inherited alzheimer network. Ann. Neurol. 79, 929–939.
Li, J.Q., Tan, L., Wang, H.F., Tan, M.S., Tan, L., Xu, W., et al., 2016. Risk factors for predicting progression from mild cognitive impairment to Alzheimer's disease: a systematic review and meta-analysis of cohort studies. J. Neurol. Neurosurg. Psychiatr. 87, 476–484.
Li, H., Jiang, G., Zhang, J., Wang, R., Wang, Z., Zheng, W.S., Menze, B., 2018. Fully convolutional network ensembles for white matter hyperintensities segmentation in MR images. Neuroimage 183, 650–665.
Prins, N.D., Scheltens, P., 2015. White matter hyperintensities, cognitive impairment and dementia: an update. Nat. Rev. Neurol. 11, 157–165.
Puelles, L., Harrison, M., Paxinos, G., Watson, C., 2013. A developmental ontology for the mammalian brain based on the prosomeric model. Trends Neurosci. 36, 570–578.
Rachmadi, M.F., Valdes-Hernandez, M.D.C., Agan, M.L.F., Di Perri, C., Komura, T., Alzheimer's Disease Neuroimaging, I, 2018. Segmentation of white matter hyperintensities using convolutional neural networks with global spatial information in routine clinical brain MRI with none or mild vascular pathology. Comput. Med. Imaging Graph. 66, 28–43.
Reginold, W., Sam, K., Poublanc, J., Fisher, J., Crawley, A., Mikulis, D.J., 2018. Impact of white matter hyperintensities on surrounding white matter tracts. Neuroradiology 60, 933–944.
Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: convolutional networks for biomedical image segmentation. Med. Image Comput. Comput. Assist. Interv. (Pt Iii 9351), 234–241.
Smith, S.M., 2002. Fast robust automated brain extraction. Hum. Brain Mapp. 17, 143–155.
Tang, X., 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, e65591.

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 Leijsen, E.M.C., Bergkamp, M.I., van Uden, I.W.M., Ghafoorian, M., van der Holst, H. M., Norris, D.G., et al., 2018. Progression of white matter hyperintensities preceded by heterogeneous decline of microstructural integrity. Stroke 49, 1386–1393.
Wallin, A., Roman, G.C., Esiri, M., Kettunen, P., Svensson, J., Paraskevas, G.P., Kapaki, E., 2018. Update on vascular cognitive impairment associated with subcortical small-vessel disease. J. Alzheimers Dis. 62, 1417–1441.
Wardlaw, J.M., Smith, C., Dichgans, M., 2013a. Mechanisms of sporadic cerebral small vessel disease: insights from neuroimaging. Lancet Neurol. 12, 483–497.
Wardlaw, J.M., Smith, E.E., Biessels, G.J., Cordonnier, C., Fazekas, F., Frayne, R., et al., 2013b. Neuroimaging standards for research into small vessel disease and its contribution to ageing and neurodegeneration. Lancet Neurol. 12, 822–838.
Wu, D., Ma, T., Ceritoglu, C., Li, Y., Chotiyanonta, J., Hou, Z., et al., 2016. Resource atlases for multi-atlas brain segmentations with multiple ontology levels based on T1-weighted MRI. Neuroimage 125, 120–130.
Wu, D., Albert, M., Soldan, A., Pettigrew, C., Oishi, K., Tomogane, Y., et al., 2019. Multi-atlas based detection and localization (MADL) for location-dependent quantification of white matter hyperintensities. Neuroimage Clin. 22, 101772.
Zhang, Y.J., Zhang, J.Y., Oishi, K., Faria, A.V., Jiang, H.Y., Li, X., et al., 2010. Atlas-guided tract reconstruction for automated and comprehensive examination of the white matter anatomy. Neuroimage 52, 1289–1301.
Zhou, T., Thung, K.H., Zhu, X.F., Shen, D.G., 2019. Effective feature learning and fusion of multimodality data using stage-wise deep neural network for dementia diagnosis. Hum. Brain Mapp. 40, 1001–1016.
Zhou, T., Fu, H.Z., Chen, G., Shen, J.B., Shao, L., 2020. Hi-net: hybrid-fusion network for multi-modal MR image synthesis. IEEE Trans. Med. Imaging 39, 2772–2781.