NeuroImage 298 (2024) 120803

Contents lists available at ScienceDirect

NeuroImage

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

Development and validation of a perivascular space segmentation method in multi-center datasets

Peiyu Huang [a,b,1], Lingyun Liu [a,b,1], Yao Zhang [a], Siyan Zhong [c], Peng Liu [d], Hui Hong [a], Shuyue Wang [a], Linyun Xie [a], Miao Lin [a], Yeerfan Jiaerken [a], Xiao Luo [a], Kaicheng Li [a], Qingze Zeng [a], Lei Cui [a], Jixuan Li [a], Yanxing Chen [c], Ruiting Zhang [a,*], for the Alzheimer's Disease Neuroimaging Initiative

[a] Department of Radiology, The Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China
[b] Department of Radiology, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China
[c] Department of Neurology, The Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China
[d] Department of Radiology, Linyi Traditional Chinese Medicine Hospital, Linyi, China

ARTICLE INFO

Keywords:
Perivascular spaces
Magnetic Resonance Imaging
Image Segmentation
Deep Learning

ABSTRACT

Background: Perivascular spaces (PVS) visible on magnetic resonance imaging (MRI) are significant markers associated with various neurological diseases. Although quantitative analysis of PVS may enhance sensitivity and improve consistency across studies, the field lacks a universally validated method for analyzing images from multi-center studies.

Methods: We annotated PVS on multi-center 3D T1-weighted (T1w) images acquired using scanners from three major vendors (Siemens, General Electric, and Philips). A neural network, mcPVS-Net (multi-center PVS segmentation network), was trained using data from 40 subjects and then tested in a separate cohort of 15 subjects. We assessed segmentation accuracy against ground truth masks tailored for each scanner vendor. Additionally, we evaluated the agreement between segmented PVS volumes and visual scores for each scanner. We also explored correlations between PVS volumes and various clinical factors such as age, hypertension, and white matter hyperintensities (WMH) in a larger sample of 1020 subjects. Furthermore, mcPVS-Net was applied to a new dataset comprising both T1w and T2-weighted (T2w) images from a United Imaging scanner to investigate if PVS volumes could discriminate between subjects with differing visual scores. We also compared the mcPVS-Net with a previously published method that segments PVS from T1 images.

Results: In the test dataset, mcPVS-Net achieved a mean DICE coefficient of 0.80, with an average Precision of 0.81 and Recall of 0.79, indicating good specificity and sensitivity. The segmented PVS volumes were significantly associated with visual scores in both the basal ganglia ($r = 0.541$, $p < 0.001$) and white matter regions ($r = 0.706$, $p < 0.001$), and PVS volumes were significantly different among subjects with varying visual scores. Segmentation performance was consistent across different scanner vendors. PVS volumes exhibited significant associations with age, hypertension, and WMH. In the United Imaging scanner dataset, PVS volumes showed good associations with PVS visual scores evaluated on either T1w or T2w images. Compared to a previously published method, mcPVS-Net showed a higher accuracy and improved PVS segmentation in the basal ganglia region.

Conclusion: The mcPVS-Net demonstrated good accuracy for segmenting PVS from 3D T1w images. It may serve as a useful tool for future PVS research.

* Corresponding author at: Department of Radiology, The Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China.
E-mail address: zhangruiting@zju.edu.cn (R. Zhang).
[1] These two authors contributed equally to this study.

https://doi.org/10.1016/j.neuroimage.2024.120803
Received 22 May 2024; Received in revised form 16 August 2024; Accepted 20 August 2024
Available online 23 August 2024

1. Introduction

Perivascular spaces (PVS) are fluid-filled anatomical structures surrounding the brain's penetrating arteries, serving as a vital part of the brain's glymphatic system (Wardlaw et al., 2020). This system is essential for removing metabolic wastes from the brain and participating in inflammatory responses (Hablitz and Nedergaard, 2021). Various pathophysiological factors (Duperron et al., 2023; Francis et al., 2019; Huang et al., 2021) such as genetic mutations, aging, and cerebrovascular diseases can impact the structure and function of PVS, leading to glymphatic dysfunction and accelerated brain degeneration (Nedergaard and Goldman, 2020; Rasmussen et al., 2018).

On human brain magnetic resonance imaging (MRI), PVS manifest as low signals on T1-weighted (T1w) images and high signals on T2-weighted (T2w) images, appearing in linear, elliptical, or circular forms depending on their orientation and the imaging angle (Duering et al., 2023; Wardlaw et al., 2013). A higher prevalence of PVS has been linked to aging (Kim et al., 2022) and cerebral small vessel disease (CSVD) (Francis et al., 2019), establishing them as a core imaging marker for CSVD (Duering et al., 2023; Wardlaw et al., 2013). Moreover, research has identified associations between PVS and various neurological disorders, including Alzheimer's disease (AD) (Zeng et al., 2022), Parkinsons' disease (Li et al., 2020), systemic lupus erythematosus (Miyata et al., 2017), and multiple sclerosis (Miyata et al., 2017), making PVS a significant focus of recent research (Smith, 2022).

Traditionally, clinicians have assessed PVS using visual scores (Doubal et al., 2010; Zhu et al., 2010), which involve counting PVS and converting these counts into a scoring system, typically ranging from 0 to 4. This method requires considerable experience from the observer, is time-consuming, and is subject to floor and ceiling effects. In response, various PVS segmentation methods have been developed to quantify PVS volume and analyze morphological changes more precisely (Barisano et al., 2022; Boutinaud et al., 2021; Choi et al., 2020; Dubost et al., 2019; Jung et al., 2019; Lan et al., 2023; Lian et al., 2018; Ranti et al., 2021; Sepehrband et al., 2019; Sudre et al., 2024). These techniques generally fall into three categories: threshold-based methods (Ramirez et al., 2015), vesselness filter methods (Ballerini et al., 2018; Butler et al., 2023; Donahue et al., 2022, 2021; Lan et al., 2023; Ranti et al., 2021; Sepehrband et al., 2019, 2021; Tidwell et al., 2023), and machine learning approaches (Choi et al., 2020; Lian et al., 2018; Rashid et al., 2023). Studies have shown that quantitative analyses might provide more robust associations with clinical variables (Wang et al., 2021), highlighting the importance of advancing PVS assessment methods in neurological research and diagnosis (Smith, 2022).

Despite technical advances, significant challenges remain in PVS segmentation. Firstly, manual delineation of PVS is time-consuming (especially for 3D high-resolution images), leading many studies to forego voxel-wise evaluation against a ground truth, opting instead to assess correlations with visual scores or numbers (Cai et al., 2015; Schwartz et al., 2019; Sepehrband et al., 2019). Secondly, most machine learning approaches train models on high-quality T1w or T2w images from a single MRI machine (Cai et al., 2015; Jung et al., 2019; Lian et al., 2018; Sudre et al., 2019), which may limit their generalizability across different centers and diseases. Thirdly, the considerable morphological changes in PVS structure in diseased states, along with potential interference from brain lesions such as white matter hyperintensities (WMH), suggest that models developed using healthy adult datasets might not perform adequately with diseased cohorts—a primary focus in clinical research. Figure S1 illustrates how imaging protocols, aging, and brain health status affect the visualization of PVS.

In this study, we aimed to develop a method more suitable for multi-center clinical imaging research. To accomplish this, we chose to train neural networks on 3D T1w images. While 3D T2w images are more sensitive, they are infrequently used in previous studies. Furthermore, many PVS can be identified on these images in young and healthy adults (Figure S1), suggesting that they do not necessarily indicate pathological changes. Moreover, 2D T2w images are prone to segmentation discrepancies due to slice thickness and PVS orientation. On the other hand, a substantial collection of isotropic 3D T1 images from past clinical imaging studies offers convenience for retrospective multi-center analyses. To minimize training bias associated with a single MRI model, we manually delineated PVS structures using images from various MRI machines. We also performed separate validations of the segmentation results for different scanner vendors. Finally, to test the generalizability of our developed model, we used a dataset of elderly community adults collected with a completely different MR machine, which included both T1w and T2w images. This allowed us to assess the sensitivity of our PVS segmentation against both T1w-and T2w-based PVS scores.

2. Methods

2.1. Description of the training and testing datasets

The training and testing data for our model were sourced from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. ADNI is an extensive, multi-center project initiated in 2003, primarily aimed at exploring the relationships among genetic factors, biomarkers, imaging data, cognitive assessments, and clinical evaluations across the full spectrum of Alzheimer's disease, from its preclinical stages through to full dementia. ADNI-3 (ClinicalTrials.gov identifier: NCT02854033) involves 59 sites across North America and includes subjects from three distinct cohorts: normal controls (NC), mild cognitive impairment (MCI), and Alzheimer's disease (AD). Further details about the study protocols and methods are accessible online through the study manual available at www.adni-info.org and http://adni.loni.usc.edu/adni-3/. All subjects and their legal representatives provided written informed consent prior to data collection.

For the training and testing of our model, subjects were pseudo-randomly selected from this extensive database. To enhance the generalizability of our segmentation method, we utilized 3D T1w images obtained from various scanner vendors (as detailed in Table 1) and included both cognitively normal subjects and those with cognitive impairments. Previous research (Griffanti et al., 2016) suggests that training algorithms with images from subjects presenting a higher number of lesions can improve model accuracy, likely due to an increased number of ground truth voxels and a more representative distribution of lesions. This observation aligns with our own experiences. Consequently, we specifically chose subjects with a relatively high number of PVS, despite the significant increase in time required for manual annotation. Ultimately, our training dataset comprised 40 subjects, and our testing dataset included 15 subjects. The data were acquired in 27 centers using 12 different type of scanner models. A detailed list of the MR scanner models can be found in the supplementary file (**Table S1**).

Table 1
Subject information of the training and testing datasets.

	Age	Sex	Cognitive Impaired	Mean PVS volume, mm^3
Training dataset, N = 40				
Siemens, N = 10	76.8 ± 9.1	2 M/ 8F	4/10	6335
GE, N = 14	71.8 ± 8.4	7 M/ 7F	6/14	3102
Philips, N = 16	75.2 ± 5.8	9 M/ 7F	6/16	5111
Testing dataset, N = 15				
Siemens, N = 6	75.3 ± 8.1	4 M/ 2F	4/6	5436
GE, N = 5	74.2 ± 11.7	3 M/ 2F	2/5	4047
Philips, N = 4	74.3 ± 5.9	3 M/ 1F	4/4	6095

2.2. MRI parameters

MR images in the ADNI3 dataset were acquired across multiple research centers using scanners from various vendors. To ensure consistency, a harmonized protocol was developed, with detailed acquisition parameters available on the ADNI website (https://adni.loni.usc.edu/methods/documents/mri-protocols/). For T1-weighted images, the main parameters were: Echo Time (TE) = min full echo, Repetition Time (TR) = 2300 ms, Inversion Time (TI) = 900 ms, and voxel size = 1 × 1 × 1 mm^3, with a 2X acceleration factor. For FLAIR images, the parameters were: TE = 119 (varies by vendor), TR = 4800 ms, TI = 1650 ms, and voxel size = 1.2 × 1 × 1 mm^3.

2.3. Image preprocessing

As image contrast can influence the visibility of PVS and thus affect PVS annotation, we implemented several preprocessing steps to reduce inter-subject variations. These procedures included bias field correction using ANTS (https://stnava.github.io/ANTs/), brain extraction using SPM 12 (https://www.fil.ion.ucl.ac.uk/spm/software/spm12/), and intensity normalization using a Fuzzy C-means method (https://github.com/jcreinhold/intensity-normalization?tab=readme-ov-file).

2.4. PVS annotation

The quality of ground truth is critical for model development, and it is important to consider several details: (1) On T1 images, differentiating PVS from other normal linear structures (e.g., the edges of brain gray matter on 2D slices or the comb-like structures between the putamen and caudate) and lesions (e.g., white matter hyperintensities, WMH, and lacunes) can be challenging. These situations were meticulously managed by considering the location, size, and morphometry of the suspected voxels, and by registering FLAIR images to T1 images to aid in distinguishing PVS from other lesions. (2) A key distinction between PVS and other disease imaging markers is that PVS is a normal structure in the human brain. PVS in the lowest part of the basal ganglia (BG) can be observed in almost all healthy young adults and does not indicate disease-related pathologies. As such, these were not annotated. (3) Annotating PVS is a complex task. Interestingly, the human visual system possesses an automatic pattern recognition capability that can help identify even vague PVS structures, although it is nearly impossible to outline the voxel at zoomed images; however, this may sometimes lead to false impressions. From a method development perspective, we prioritized robustness and only annotated those voxels that could be clearly identified on both large and small scales.

The detailed procedures for PVS annotation included several steps: (1) Initially, a group meeting was convened to unify opinions regarding the definition of PVS, identification of PVS mimics, and the visual threshold required for delineating a specific PVS. (2) Given the substantial amount of work involved, PVS annotation was divided into three stages. Initially, one postgraduate student (LL, with 2 years of experience) created the preliminary PVS masks. These masks were then refined by another postgraduate student (YZ, with 5 years of experience), followed by final corrections made by a senior neuroimaging researcher (pH, with 15 years of experience). (3) The observers used ITK-SNAP (http://www.itksnap.org) to draw the PVS on a voxel-by-voxel basis. For each image, linear PVS structures that were clearly visible were first outlined on three orthogonal planes. Subsequently, the observers searched for and annotated any missed PVS that appeared as oval or round structures. To aid in differentiating PVS from mimics, co-registered FLAIR images were displayed alongside T1w images. However, given that 3D FLAIR is not readily available in previously acquired datasets, we relied solely on T1 images for model training. The annotation process for a single subject took over 20 h.

2.5. Model training

We employed the nnU-Net framework (Isensee et al., 2021) (https://github.com/MIC-DKFZ/nnUNet) for training our model. The nnU-Net is a deep learning-based segmentation tool that automates the configuration of parameters for preprocessing, network architecture, training, and post-processing for any new task. It has been extensively used and recognized as one of the top segmentation methods across more than 23 public biomedical imaging datasets (Isensee et al., 2021), with few new models surpassing its performance to date.

Given our objective to develop a practical tool for application in multi-center datasets rather than focusing on methodological refinement, we opted for this well-established method. The training was executed on an Nvidia 3090 card (24 G memory). We trained the 3D full-resolution model across 5 folds, with each fold undergoing 1000 epochs. Each epoch took approximately 65 s, culminating in a total training duration of about 91 h. The optimization of other hyperparameters was managed automatically by nnU-Net.

We have named this trained network "mcPVS-Net," which stands for multi-center PVS segmentation network.

2.6. Evaluation against ground truth

We used three metrics (Boutinaud et al., 2021; Lian et al., 2018) to evaluate the accuracy of the mcPVS-Net:

(1) DICE. The DICE coefficient is defined as: DICE=2 × |A∩B| / (|A|+|B|). A & B represent the segmentation result and the ground truth mask, respectively. DICE ranges from 0 to 1, where 0 indicates no overlap between the two masks and 1 indicates perfect overlap.
(2) Precision. The Precision is defined as: Precision=TP / (TP+FP). TP (True Positives) is the number of correct positive voxels made by the model, and FP (False Positives) is the number of incorrect positive voxels made by the model. A higher precision value, closer to 1, indicates that a greater proportion of the positive predictions made by the model are correct, while a lower value indicates that the model is making more incorrect positive predictions.
(3) Recall. The Recall is defined as: Recall=TP / (TP+FN). FN (False Negatives) represents the number of voxels that were incorrectly predicted as negative. A higher recall value, closer to 1, indicates that the model is effectively identifying the actual PVS, while a lower recall value suggests that the model is missing a significant number of PVS.

Additionally, the correlations between the PVS volumes from the ground truth and the segmented masks were calculated for both the training and testing datasets.

2.7. Evaluation in the ADNI3 database

A total of 1090 baseline T1w images were sourced from the ADNI3 database. Images affected by severe artifacts (N = 36) or high noise levels (N = 34) were excluded. The remaining images (N = 1020) underwent the previously described preprocessing procedures, and PVS were segmented using mcPVS-Net. These data were collected from all the participating sites of the ADNI3 project and the subject information could be seen in **Table S2**. Visual evaluation of the segmentation results was conducted for each subject. While various types of errors such as missing PVS or mistaking artifacts for PVS were noted, no manual corrections were made due to the mild nature of most errors and the impracticality of manual correction in such a large dataset.

BG-PVS and white matter (WM) PVS are distinct in their anatomical structures and associated pathologies, hence their links with neurological diseases differ. All T1w images were processed using the Samseg pipeline (https://surfer.nmr.mgh.harvard.edu/fswiki/Samseg), integrated in Freesurfer (version 7.2), which provided comprehensive brain

segmentation including labels for white matter and deep nuclei. Binary masks for the BG and cerebral WM were extracted from these results to separate PVS into specific regions.

To verify whether mcPVS-Net might erroneously classify white matter hyperintensities (WMH) as PVS, WMH were segmented from FLAIR images using an in-house trained deep learning model, which was previously trained on images from 100 subjects with various conditions (healthy, CSVD, AD). Testing on 40 subjects yielded an average DICE coefficient of 0.81. Raw FLAIR images and WMH masks were then registered to corresponding T1 images to calculate overlap rates between WMH and PVS. The volume of WMH (WMHnorm) was also normalized using the intracranial volume (ICV).

To assess the segmentation against visual scores, 300 subjects were selected from the entire dataset using a stratified random method, with 100 images from each scanner type. An experienced radiologist (PL, with 5 years of experience in Radiology) independently evaluated all 300 images in the BG and WM regions based on established standards. The scoring was as follows: in the basal ganglia, PVS were rated 1 if <5 were present, rated 2 for 5–10 PVS, rated 3 for more than 10 countable PVS, and rated 4 when PVS were uncountable. In the white matter, PVS were rated 1 if <10 were present in total, rated 2 for >10 total but no more than 10 in any single section, rated 3 for 10–20 PVS in the section with the most PVS, and rated 4 for >20 PVS in any single section. The correlation between visual scores and PVS volumes was examined using Spearman's correlation, and the differences in PVS volumes among different rating scores were tested using ANOVA with post-hoc analyses. The Dunnett's T3 method was used to control for multiple comparisons, assuming unequal variances. All analyses were conducted both across all 300 subjects and within each scanner type.

The potential associations of between PVS with recognized factors such as age, hypertension, and WMH were explored using multiple linear regression analyses. To adjust for the influence of head size, PVS volumes were normalized against the intracranial volume (ICV): Normalized PVS (PVSnm) = 1000,000 * PVS volume / ICV. PVSnm was set as the dependent variable and the related factors were set as independent variables. Analyses were performed both with and without controlling for scanner vendor.

2.8. Evaluation of segmentation consistency in a longitudinal dataset

To test the consistency of mcPVS-Net for segmentation in longitudinal data, we analyzed 90 subjects (30 for each scanner) with 3 timepoints (baseline, 1 year and 2 years follow up) from the ADNI3 database. First, we created a common template for each subject and registered the T1w images in all timepoints to the subject-specific common template using the antsMultivariateTemplateConstruction2 command embedded in the ANTs tool. Default registration parameters were used, and nonlinear registrations were performed to create the template. Then, all images went through the preprocessing and segmentation pipelines as we described above. We extracted PVS volumes from every subject at each timepoint. Intra-class correlation coefficients were used to assess segmentation consistency across timepoints.

2.9. Evaluation in a new dataset with high-resolution T2w images

Finally, we evaluated whether mcPVS-Net could be directly applied to another independent dataset (Dataset2, $N = 173$). In brief, subjects in this dataset were community-dwelling middle-to-old age adults enrolled through advertisements in communities and online. Subject information could be seen in **Table S3**. All T1w and T2 images were acquired using a United Imaging 790 3T scanner. The T1w image had a voxel size of $1 \times 1 \times 1mm^3$, and the T2w image had a voxel size of $0.8 \times 0.8 \times 0.8mm^3$.

PVS was segmented from T1w images using the pipeline as described above. We annotated PVS in 5 subjects, and calculated the DICE, Recall, and Precision values. T1-based PVS scores were evaluated by the same observer (PL) using identical criteria. Additionally, as T2w may have better sensitivity at detecting PVS, we also evaluated PVS visual scores on T2w images. An experienced Radiologist (RZ, 8 years of experience) evaluated the PVS score according to another widely used T2w-based criteria (Doubal et al., 2010). In brief, 0=no PVS, 1 ≤ 10 PVS, 2 = 11 to 20 PVS, 3 = 21 to 40 PVS, and 4=more than 40 PVS. The numbers refer to PVS on one side of the brain with the highest number of PVS.

Correlation between visual scores and PVS volume were examined using Spearman's correlation, and the PVS volume difference among groups with varying PVS scores was tested using ANOVA with post-hoc analyses. The Dunnett's T3 method was used to control for multiple comparisons, assuming unequal variances. Groups with less than 5 subjects were excluded from statistical analyses. We also tested the association between PVSnm and age, hypertension, and WMH volume using multiple regression analysis.

2.10. Comparison with a previous published method

Currently there is one published and openly available deep-learning method that segment PVS from T1 images (Boutinaud et al., 2021). We compared the performance of this neural network model with mcPVS-Net on the 15 testing images. Following descriptions in the paper, we performed brain segmentation, intensity normalization, and cropping. The preprocessed images were then segmented using the provided neural network model, and the results were transformed back into the original space for comparison with results produced by mcPVS-Net. This method produces probabilistic maps, and the authors recommend a threshold of 0.5. We tested the threshold from 0.1 to 0.9 with an interval of 0.1, confirming that the threshold of 0.5 could produce results with the highest DICE. Finally, we compared the results using DICE, Precision, and Recall. Visual assessments were also performed to identify the pattern of inconsistent segmentations.

3. Results

3.1. Evaluation against ground truth

Evaluation against ground truth showed a mean DICE of 0.84 in the training dataset and a mean DICE of 0.80 in the testing dataset (Table 2). The average Precision in the training dataset was 0.91 and 0.81 in the testing dataset, indicating that about 80~90 percent of the voxels detected by the model were true PVS voxels. The Recall metric suggests that the model correctly detected about 80 % of the ground truth voxels. When separating all data according to scanner vendors, the segmentation metrics remained similar across different vendors. The correlation coefficients between the ground truth and predicted PVS volumes were 0.997 in the training dataset and 0.970 in the testing dataset (both $p < 0.001$, Fig. 1).

3.2. Evaluation in the ADNI3 database

In 300 subjects, the PVS volume was closely associated with visual scores in both the BG ($r = 0.541$, $p < 0.001$) and WM ($r = 0.706$, $p <$

Table 2
Model evaluation against ground truth.

	DICE	Precision	Recall
Training dataset N = 40			
ALL	0.84 ± 0.06	0.91 ± 0.04	0.79 ± 0.08
Siemens	0.84 ± 0.03	0.90 ± 0.04	0.79 ± 0.03
GE	0.82 ± 0.08	0.92 ± 0.03	0.75 ± 0.11
Philips	0.86 ± 0.05	0.90 ± 0.04	0.82 ± 0.06
Testing dataset N = 15			
ALL	0.80 ± 0.09	0.81 ± 0.09	0.79 ± 0.12
Siemens	0.76 ± 0.10	0.74 ± 0.08	0.78 ± 0.14
GE	0.85 ± 0.09	0.90 ± 0.06	0.81 ± 0.12
Philips	0.80 ± 0.09	0.81 ± 0.07	0.79 ± 0.11

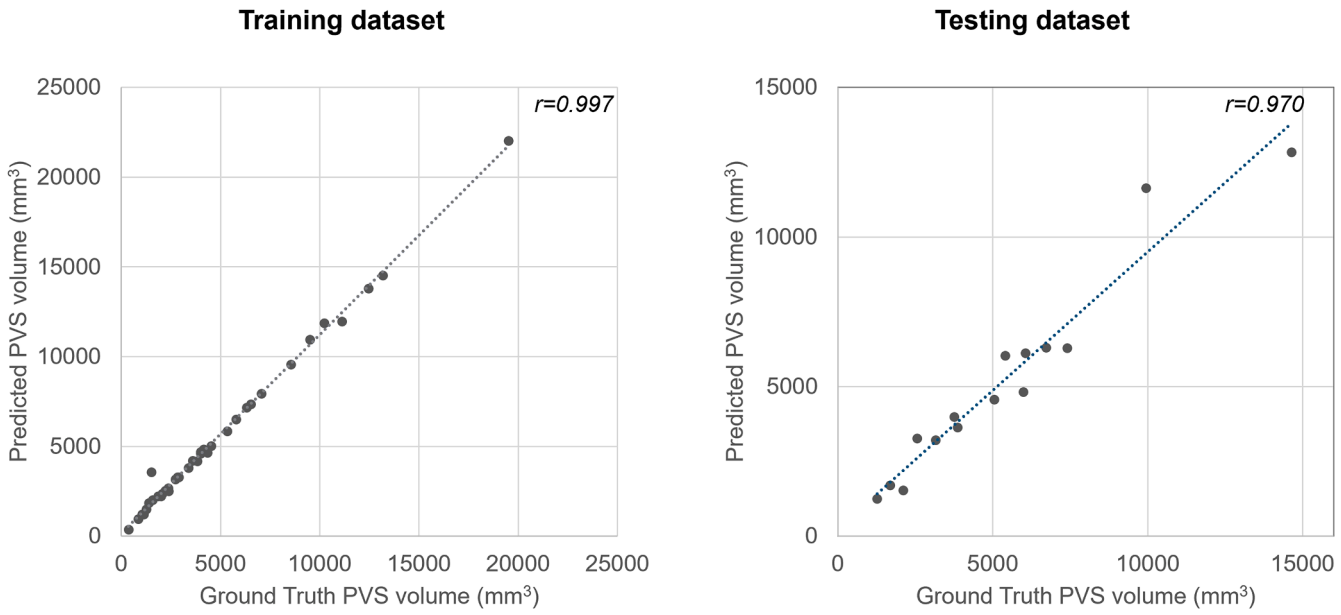

Fig. 1. Correlation between the ground truth and predicted PVS volumes.

0.001) region. In Siemens images, the correlation coefficients were 0.512 and 0.657; in GE images, the correlation coefficients were 0.576 and 0.754; in Philips images, the correlation coefficients were 0.630 and 0.726; all correlations had a $p < 0.001$.

As shown in Fig. 2, PVS volume robustly increased from subjects with low to high PVS scores, and the PVS volumes were significantly different between almost all pairs of scores. An exception is the comparison between score 1 and 2 in the WM region. This might be due to the small number of subjects with score 1 in the whole group ($N = 14$) and different scanner groups. A list of PVS volumes corresponding to different visual scores can be seen in Supplementary **Table S4**.

In all 1020 subjects from ADNI3 **(Supplementary Table S2)**, the median of the overlapping rate between PVS and WMH in the total dataset was 1.33 % (interquartile range: 0.28 %~3.6 %), suggesting a very low error rate. Regression analyses showed that age, hypertension, and WMH were all associated with BG-PVSnm volume (Table 3), and age and WMH were associated with WM-PVSnm. The model R^2, standardized β, and p values were very close between models with and without controlling for the scanner vendor.

Some segmentation results were demonstrated in Fig. 3. There are apparent differences in the tissue contrast among images from three vendors. Nonetheless, mcPVS-Net generated satisfactory results in all three subjects.

3.3. Segmentation consistency in a longitudinal dataset

The ICC across three timepoints in all 90 subjects was 0.966. In Siemens, GE, and Philips subsets, the ICCs were 0.954, 0.927, and 0.978, respectively. The p values for all the ICC analyses were <0.001.

3.4. Evaluation in Dataset2

The mean DICE, Precision, and Recall values were 0.67, 0.79, and 0.59, respectively. A demonstration of PVS segmentation in Dataset2

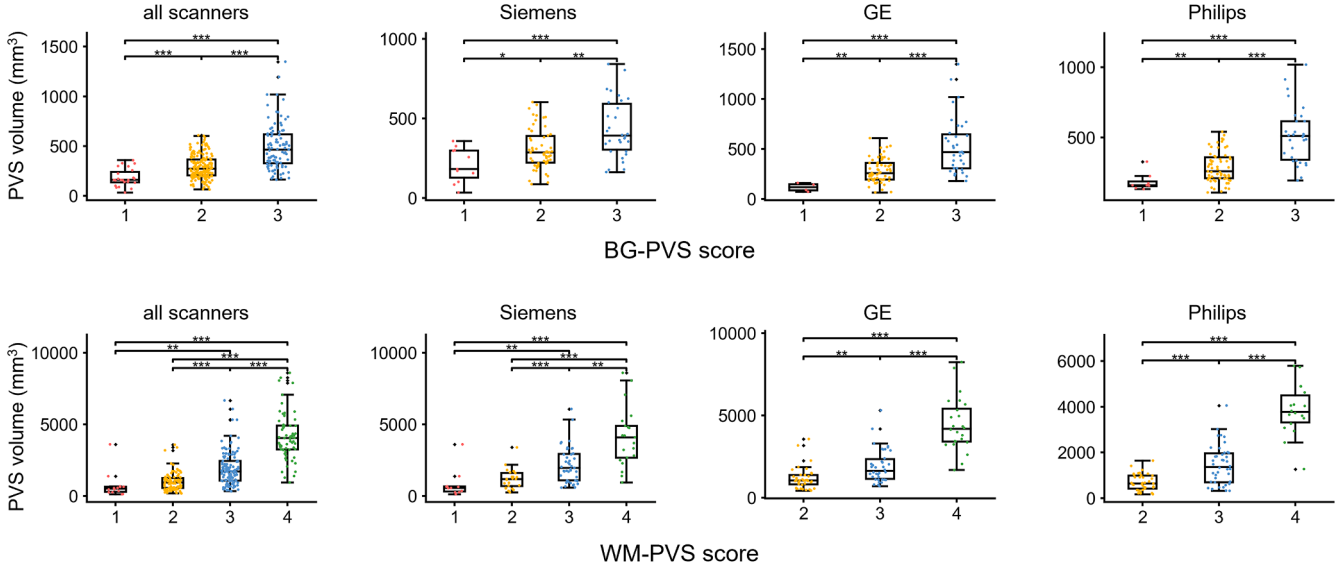

Fig. 2. Differences of PVS volume among subjects with different visual scores. *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$. Multiple comparison correction was performed using the Dunnett T3 Method. Groups with less than 5 subjects were excluded from statistical analyses.

Table 3
Associations between PVSnm and related risk factors.

Variables	BG-PVSnm				WM-PVSnm			
	without scanner variable		with scanner variable		without scanner variable		with scanner variable	
	β	p	β	p	β	p	β	p
Age	0.176	**<0.001**	0.169	**<0.001**	0.102	**0.003**	−0.101	**0.003**
Hypertension	0.062	**0.037**	0.064	**0.031**	−0.016	0.611	−0.019	0.559
WMHf	0.256	**<0.001**	0.263	**<0.001**	−0.081	**0.020**	−0.080	**0.020**
Model R^2	0.142(0.001)		0.145($p < 0.001$)		0.011(0.058)		0.010(0.015)	

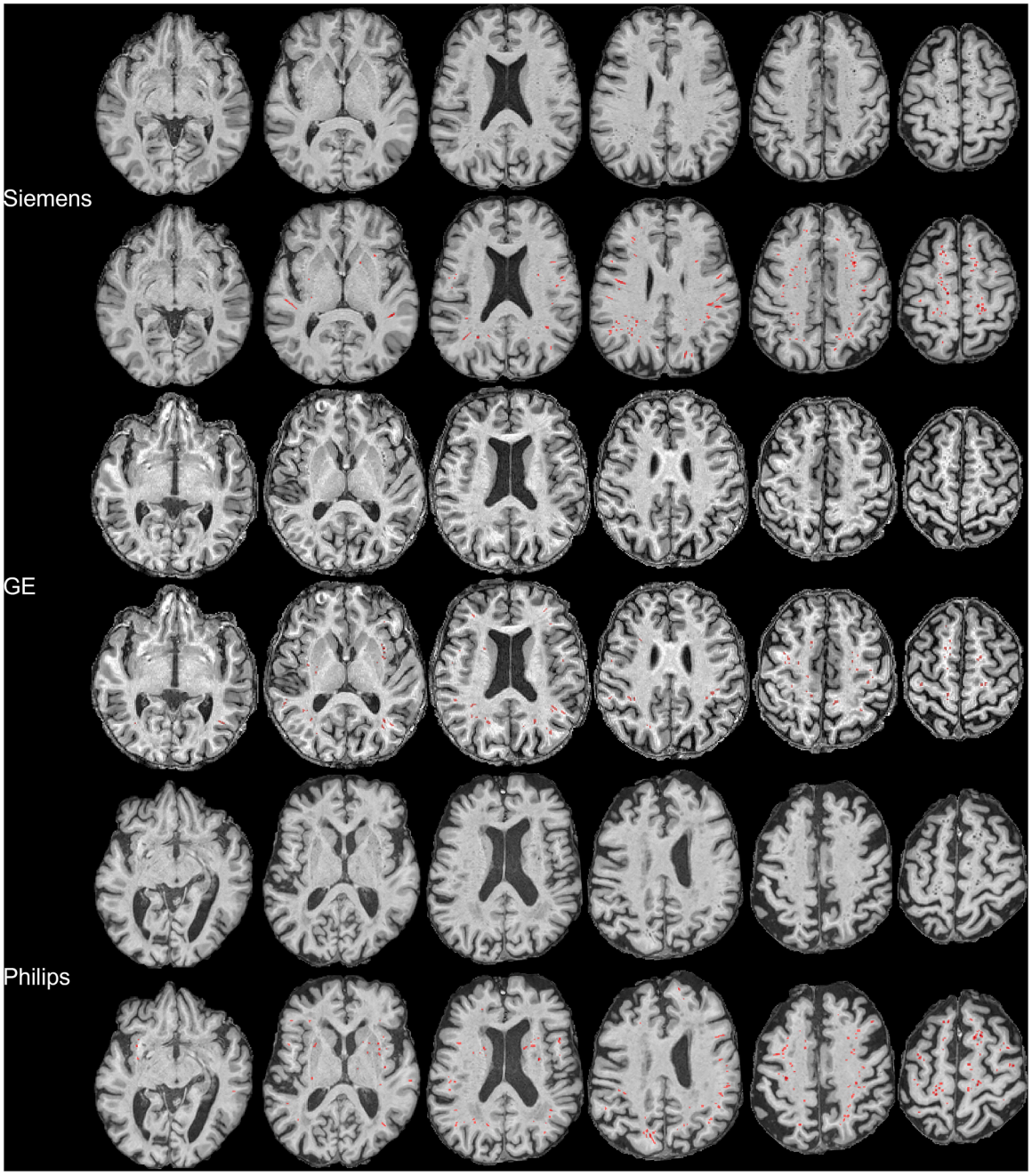

Fig. 3. Demonstration of PVS segmentation in images from different scanners.

can be seen in Fig. 4. The PVS volumes were significantly correlated with visual scores evaluated on T1w image (BG: $r = 0.401$; WM: $r = 0.809$; $p < 0.001$ for both). ANOVA analysis among subjects with different T1-PVS scores showed significant differences in PVS volumes (Fig. 4, $p <$ 0.001 for both BG and WM region). Post-hoc analysis showed significant differences in almost all pairs of scores except for 3–4 in the WM region. Please note that there was apparent difference between the two groups, the non-significant result might be due to the small number of subjects

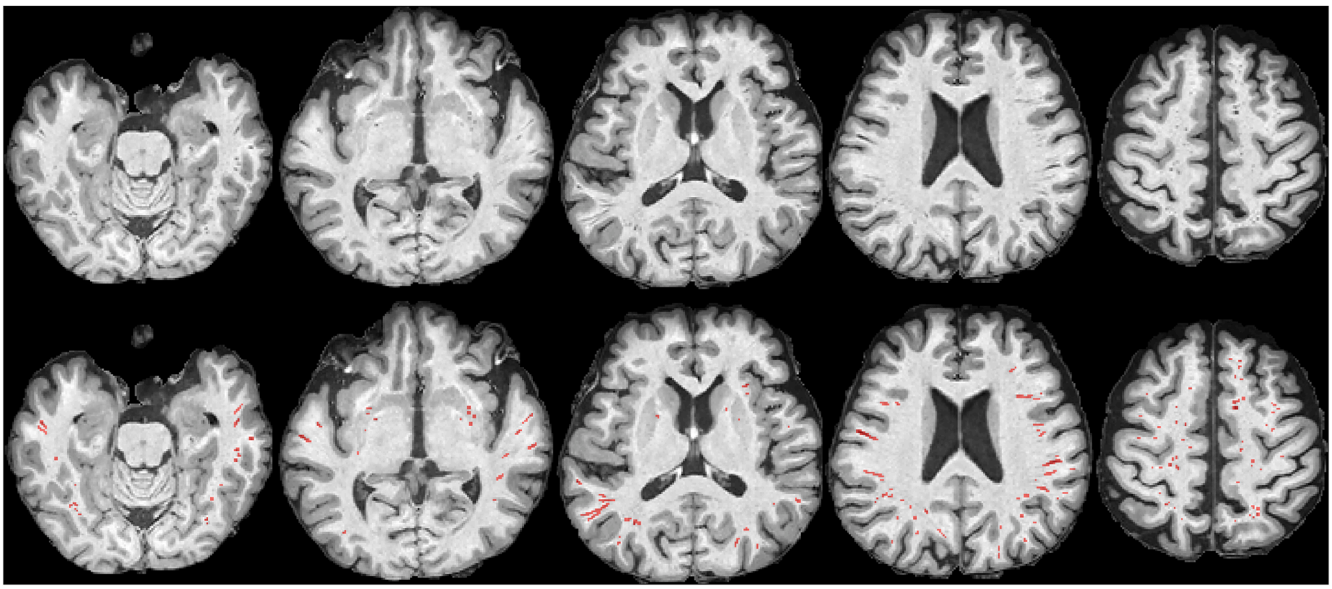

Fig. 4. Demonstration of PVS segmentation in one subject from Dataset2.

($N = 6$) in group 4.

The PVS volumes were also correlated with T2-based PVS scores (Fig. 5, BG: $r = 0.567$; WM: $r = 0.673$; both $p < 0.001$). ANOVA showed significant differences among groups ($p < 0.001$ for both BG and WM region), and post-hoc analysis showed significant differences in all pairs of scores.

Age (β=0.188, $p = 0.021$) and WMHnorm (β=0.319, $p < 0.001$) were associated with higher BG-PVS volume, and hypertension (β=0.188, $p = 0.020$) was associated with higher WM-PVS volume. A list of PVS volumes corresponding to different visual scores can be seen in Supplementary **Table S5**.

3.5. Segmentation in subjects with imaging artifacts and lacunes

Artifacts are common in clinical imaging involving older adults and patients. We found that mcPVS-Net was relatively robust against mild-to-moderate artifacts. As shown in Fig. 6A, many imaging artifacts can be seen in the frontal and occipital lobe of the subject. The mcPVS-Net only segmented a small number of voxels as PVS.

Furthermore, lacunes are also a common imaging occurrence, which can be difficult to be differentiated from PVS even for experienced researchers. Fig. 6B and 6C showed several cases where mcPVS-Net correctly avoided lacunes.

3.6. Comparison with a previous deep learning method

Using Boutinaud' method, mean DICE in the testing data ranged from 0.40 to 0.51, with the highest value derived from threshold 0.5. Under this threshold, the mean DICE value was 0.51, the mean Precision was 0.54, and the mean Recall was 0.52. We examined the results and found that the two methods were generally consistent in WM PVS segmentation (Fig. 7A), but there were two major differences. First,

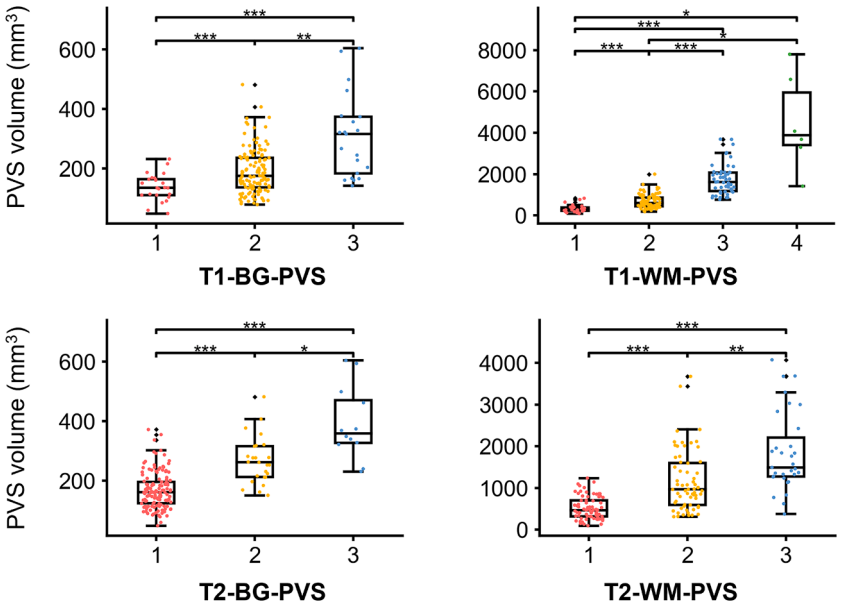

Fig. 5. Differences of PVS volume among subjects with different visual scores evaluated on T1 or T2 weighted images. *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$. Multiple comparison correction was performed using the Dunnett T3 Method. Groups with less than 5 subjects were excluded from statistical analyses.

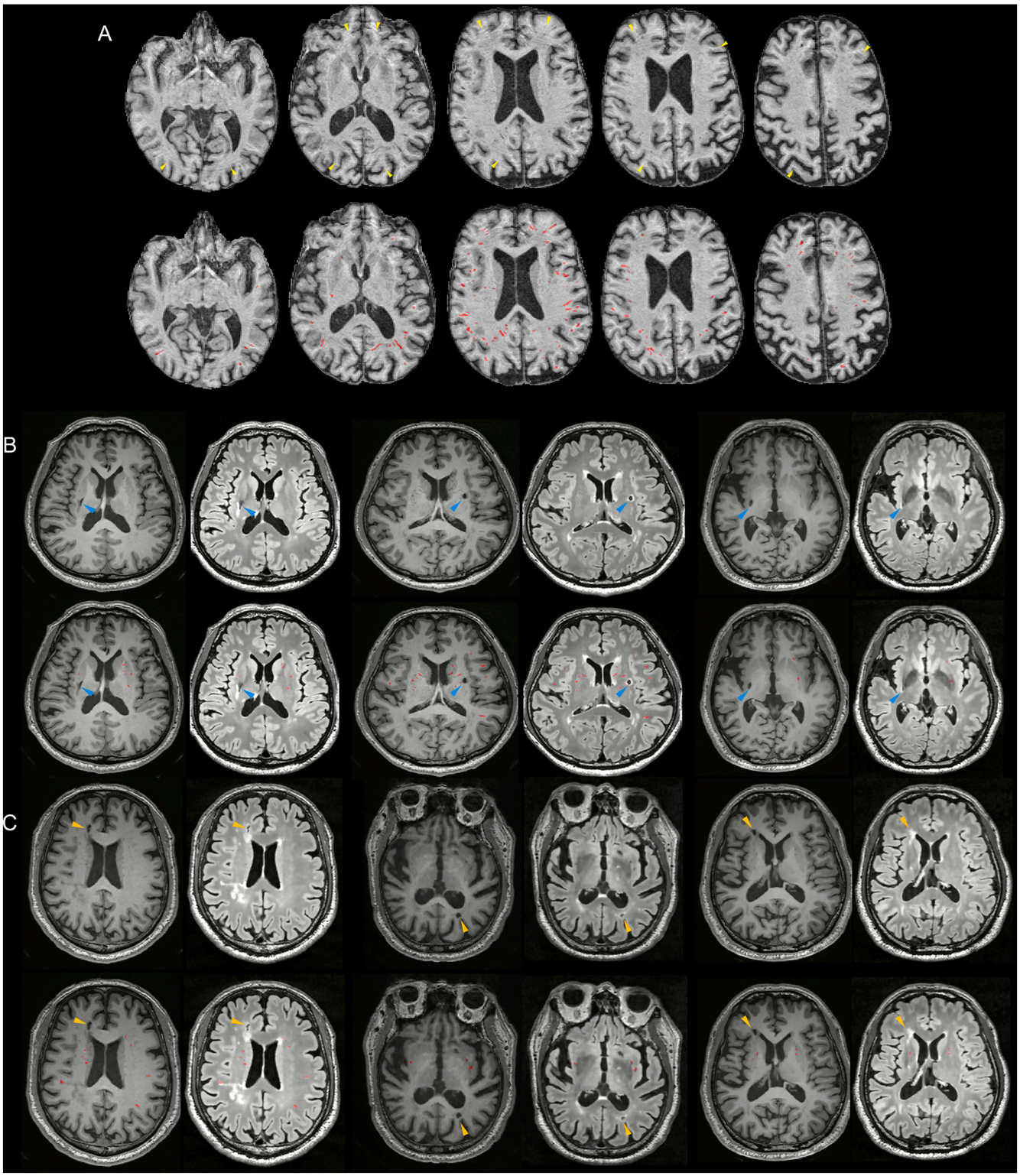

Fig. 6. (A) PVS segmentation in a subject with widespread imaging artifacts (yellow arrowhead) from the ADNI database. The mcPVS-Net had a low error rate, ignoring most of the dark artifact lines. (B) mcPVS-Net avoided lacunes in the BG region (blue arrowhead) in subjects from the United imaging dataset. (C) mcPVS-Net avoided lacunes in various WM regions (orange arrowhead) in subjects from the United imaging dataset.

Boutinaud' method segmented large PVS in the inferior part of BG (Fig. 7B) while mcPVS-Net didn't. Second, the previous method neglected some PVS in the BG region (Fig. 7C). Although lowering the threshold to 0.1 might remedy some errors, many PVS were still missing, and a low threshold significantly increased false positives in other brain regions.

4. Discussion

In this study, we developed mcPVS-Net, a neural network model that can segment PVS from multi-center 3D T1 images. Tests against ground truth PVS masks showed good segmentation accuracy. There were good correlations between PVS volumes and visual scores in both the BG and

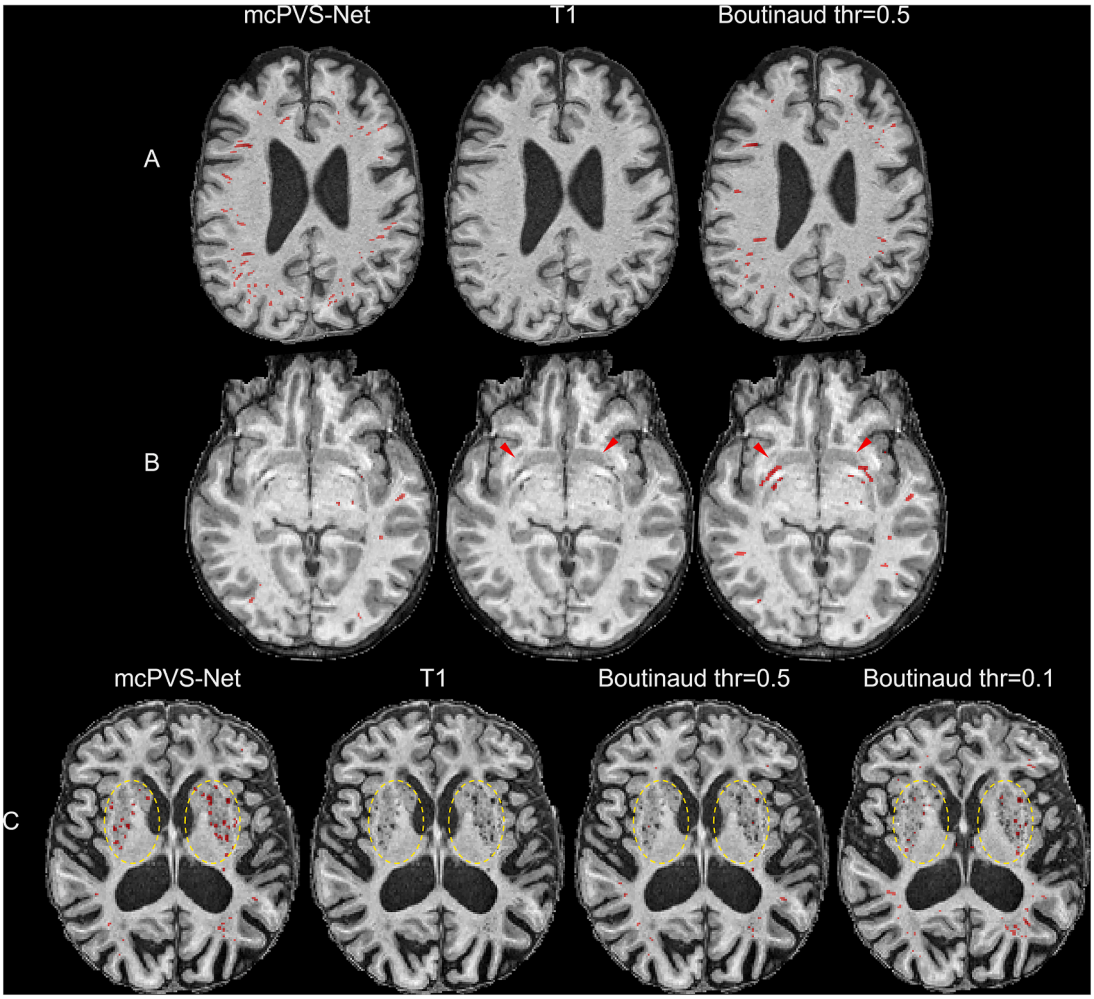

Fig. 7. (A) Our method and Boutinaud's method performed similar in the WM region. (B) mcPVS-Net did not segment PVS in the lower BG region. (C) mcPVS-Net segmented more PVS in the BG region.

WM region and the performance across different scanner vendors was similar. Associations between PVS volume and several well-recognized factors of PVS dilation were validated in the whole ADNI3 dataset. Furthermore, we tested the generalizability of mcPVS-Net in another completely new dataset, which also demonstrated good accuracy and the PVS volume can differentiate different visual scores evaluated using either T1 or T2-weighted images. In addition, the model can produce robust results when imaging artifacts and other brain lesions existed. The mcPVS-Net may serve as a useful tool for future multi-center clinical imaging studies.

Recently, several reviews systematically analyzed the related literatures and pointed out a few limitations (Pham et al., 2022; Waymont et al., 2024), including a lack of validation against ground truth, difficult to be generalized, lack of differentiation between PVS and other lesions, etc. Many previous studies adopting thresholding or filtering methods (Ballerini et al., 2018; Ramirez et al., 2015) evaluated model performance against visual assessments, but not ground truth masks. Among deep learning based segmentation methods, Lian et al. (2018) and Jung et al. (2019) trained models on images acquired on 7T MR scanners, which are not readily available in most clinical sites. Sudre et al. (2019) trained a model based on 1 mm isotropic T1w, T2w, and FLAIR from 14 subjects and evaluated the model on 2 subjects. Rashid et al. (2023) trained multi-class models based on high-resolution T1w, T2w, FLAIR, and SWI images on 21 subjects, but the main purpose was to compare the contribution of different sequences and only leave-one-out cross--validation was performed. Dubost et al. (2019) trained a neural network based on high-resolution T2 weighted images, but the evaluation was against visual scores. Considering that many previous projects, such as the ADNI and UK biobank, had not acquired high resolution T2 images, these models may have limited applicability. Williamson et al. (2022) trained a neural network based on thick-slice T2w images from multiple centers and the goal was to detect none-to-mild versus moderate-to-severe EPVS, so no quantitative information could be derived.

There is one openly available deep learning method that segments PVS solely based on 3D T1w images (Boutinaud et al., 2021). The authors used a class of DL methods based on autoencoders and U-shaped networks. They trained the model using T1w images (1 mm isotropic) of 40 subjects acquired using an identical Siemens Prisma 3T scanner, and the model was tested in 10 images with the same acquisition conditions. The DICE coefficient for PVS segmentation was 0.51 for WM and 0.66 for BG. The authors applied the method to another MRI database (1 mm isotropic, Philips 3T scanner) and obtained a similar distribution of PVS load. However, no further comparison against ground truth or visual scores were performed in the new dataset. Unlike on T2w images, PVS on T1w images are relatively sparse, therefore, the authors combined the autoencoders and U-shaped networks to overcome this issue. Here we tried to mitigate this issue through training on subjects with large amount of PVS. Relative advantages of the current mcPVS-Net include training and testing on multi-center images from three scanner vendors, validation of the associations between PVS volume and well-recognized risk factors, additional validation against T1w and T2w scores in a

completely new dataset, and evaluation of interferences from artifacts and other lesions. Because there are significant differences between the vesselness approach and the deep learning approach, we did not compare our method with studies using the vesselness approach.

With the rise of clinical interests on PVS, many researchers try to apply quantitative PVS measures in clinical studies. However, few studies performed evaluation against ground truth in multi-center datasets. We evaluated the quality of PVS segmentation in images from three different vendors and found that PVS volumes were well correlated with visual scores across the different vendors. We also showed that introducing the scanner variable almost did not improve prediction power in the large ADNI3 dataset. Furthermore, direct application in another completely new dataset showed good segmentation accuracy, and the segmented PVS volumes were significantly different among groups with different visual scores. Possibly due to a smaller sample size and a younger mean age, some relationships between risk factors and PVS volumes were not statistically significant. Nonetheless, these results suggest good model generalizability, supporting the use of mcPVS-Net in large scale studies that may contain 3D T1w images from various scanners.

A potential concern regarding PVS segmentation based on T1w images is the reduced sensitivity compared to T2w-based approaches. Interestingly, we found that PVS volume derived from T1w images could differentiate subjects with different T2w-PVS scores, suggesting that the current approach has a sensitivity like traditional T2w based visual assessment. It should be noted that although there are good correlations between T1w- and T2w-based visual scores (Adams et al., 2015; Evans et al., 2023), T2w images are still superior for detecting small PVS. PVS segmentation based on T2w images should be able to detect more subtle changes. Nonetheless, comparison with T2w-based PVS segmentation will introduce other methodological issues beyond the scope of this study. Furthermore, a recent study using Boutinaud's method found that BG-PVS volume segmented from T1w images could predict CSVD progression, confirming the clinical value of T1-based PVS segmentation (Li et al., 2024).

Imaging artifacts and contamination from other lesions are common in clinical imaging of older adults and patients. In this study, we only excluded those patients with severe head motion that caused apparent blurring in images. In subjects with mild-to-moderate head motion, we found that the mcPVS-Net could avoid most of the artifacts. Errors only occurred when the location and orientation of artifacts were like real PVS tubes. Furthermore, the segmented PVS masks had a very low overlapping rate with WMH and could avoid typical lacune lesions. We noticed that a few small lacunes that looked like PVS were wrongly segmented as PVS. Nonetheless, differentiation of PVS and small lacunes are very difficult based on solely T1 images. Although adding FLAIR to training modalities may increase the accuracy regarding lacune differentiation, we had not done this due to (1) many past imaging projects had not acquired 3D FLAIR images; (2) lacunes are relatively scarce for most datasets and removing them from the PVS masks is not difficult. Training on 3D-T1 images alone, on the other hand, improves the applicability of our model. In general, the mcPVS-Net was robust to these interferences. We demonstrated the association between PVS and risk factors without applying any postprocessing corrections, confirming the robustness of mcPVS-Net.

This study is subject to several limitations. First, due to the vast amount of work for PVS annotation, we only annotated 60 T1w volumes and the model was only tested in 20 subjects. While the sample size is already larger than all previous PVS segmentation studies, it is still relatively small compared to many other imaging segmentation tasks. This is mainly due to the difficulty of PVS annotations on high-resolution images. While the number of images volumes was small, there were over 30,000 (over 200 slices, 3 orthogonal planes for each subject) image slices and countless small PVS objects. Second, we did not evaluate inter-rater or intra-rater reliability for the same reasons mentioned above. It should be noted that the identification and annotation of PVS are not easy tasks and may be subject to observer bias. Similarly, segmentation accuracy in the longitudinal dataset was evaluated using PVS volumes rather than ground truth masks. Therefore, variations in longitudinal segmentations may exist. Third, we chose to use only T1w images to train the neural network and this strategy has inherent shortcomings, such as a relatively low sensitivity to PVS, difficult to differentiate other lesions, etc. Despite these issues, we showed that the sensitivity of mcPVS-Net is similar to traditional T2-weighted scoring method and is robust to common interferences. On the other hand, this strategy allows the mcPVS-Net to be applied to many existing datasets that only acquired high-resolution T1 images. Fourth, due to our annotation strategies, the model may overlook some PVS structures with very weak contrast to adjacent white matter (Figure S3). According to our tests, this strategy does not significantly influence the model's clinical applicability but can substantially improve annotation consistency. Fifth, we assessed segmentation accuracy in scanners from different vendors, but there would still be differences related to scanner models. Due to the large number of scanner models, assessing segmentation accuracy for each model would require a much larger number of annotated cases. Image quality also depends on other complex parameters that influence spatial resolution, tissue contrast, etc. Considering that acquisition protocols have been harmonized in the ADNI project, PVS segmentation accuracy in other datasets may decrease. Since we demonstrated appropriate segmentation in our local dataset, we expect decent performance for images acquired using similar protocols. However, applications in significantly different images, such as those acquired using 1.5T or 7T scanners, would need further validation. Careful checking for incorrect segmentation due to imaging artifacts and other unexpected lesions is also necessary. Lastly, although we tested the association between PVS volume and visual score in images from different vendors, we had not evaluated the segmentation consistency using travel subject. The site effect should be carefully examined in future application studies.

5. Conclusion

We successfully developed mcPVS-Net, a neural network model that segments PVS from 3D-T1w images. Validations against ground truth masks and visual scores demonstrated good accuracy and generalizability. The mcPVS-Net may serve as a useful tool for future studies.

Availability of data and materials

Data used in preparation of this article were obtained from the Alzheimer's disease Neuroimaging Initiative (ADNI) database (www.adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

The datasets analyzed during the current study are not publicly available due to subject privacy policies. For potential data sharing, we will need a formal data sharing agreement and approval from the requesting researcher's local ethics committee.

The trained model is available at https://github.com/huangnick/mcPVS-Net/tree/main.

Ethics approval and consent to participate

The study was approved by the Medical Ethics Committee of the Second Affiliated Hospital, Zhejiang University School of Medicine. All participants in this study were included after written informed consent.

Funding

This work was supported by the National Natural Science Foundation

of China (Grant No: 82371907; 82101987; 82271936; 82202090; 82302138) and the Natural Science Foundation of Zhejiang Province (Grant No: Z24H180002 & LQ20H180015), China Postdoctoral Science Foundation (Grant No: 2019M662083). The fundings agencies had no further role in study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication.

CRediT authorship contribution statement

Peiyu Huang: Writing – original draft, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization. **Lingyun Liu:** Writing – original draft, Validation, Investigation, Formal analysis. **Yao Zhang:** Investigation, Formal analysis, Data curation. **Siyan Zhong:** Investigation, Formal analysis, Data curation. **Peng Liu:** Formal analysis, Data curation. **Hui Hong:** Investigation, Data curation. **Shuyue Wang:** Investigation, Data curation. **Linyun Xie:** Investigation, Data curation. **Miao Lin:** Visualization, Investigation, Data curation. **Yeerfan Jiaerken:** Investigation, Data curation. **Xiao Luo:** Investigation, Data curation. **Kaicheng Li:** Investigation. **Qingze Zeng:** Investigation. **Lei Cui:** Investigation. **Jixuan Li:** Investigation. **Yanxing Chen:** Writing – review & editing, Funding acquisition, Conceptualization. **Ruiting Zhang:** Writing – review & editing, Writing – original draft, Supervision, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

We have no conflict of interest to declare.

Data availability

The authors do not have permission to share data.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2024.120803.

References

Adams, H.H., Hilal, S., Schwingenschuh, P., Wittfeld, K., van der Lee, S.J., DeCarli, C., Vernooij, M.W., Katschnig-Winter, P., Habes, M., Chen, C., 2015. A priori collaboration in population imaging: the uniform neuro-imaging of virchow-robin Spaces Enlargement consortium. Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring 1, 513–520.

Ballerini, L., Lovreglio, R., Valdés Hernández, M.d.C., Ramirez, J., MacIntosh, B.J., Black, S.E., Wardlaw, J.M., 2018. Perivascular spaces segmentation in brain MRI using optimal 3D filtering. Sci. Rep. 8, 2132.

Barisano, G., Lynch, K.M., Sibilia, F., Lan, H., Shih, N.-C., Sepehrband, F., Choupan, J., 2022. Imaging perivascular space structure and function using brain MRI. Neuroimage 257, 119329.

Boutinaud, P., Tsuchida, A., Laurent, A., Adonias, F., Hanifehlou, Z., Nozais, V., Verrecchia, V., Lampe, L., Zhang, J., Zhu, Y.-C., 2021. 3D segmentation of perivascular spaces on T1-weighted 3 Tesla MR images with a convolutional autoencoder and a U-shaped neural network. Front Neuroinform 15, 641600.

Butler, T., Zhou, L., Ozsahin, I., Wang, X.H., Garetti, J., Zetterberg, H., Blennow, K., Jamison, K., de Leon, M.J., Li, Y., 2023. Glymphatic clearance estimated using diffusion tensor imaging along perivascular spaces is reduced after traumatic brain injury and correlates with plasma neurofilament light, a biomarker of injury severity. Brain Commun. 5, fcad134.

Cai, K., Tain, R., Das, S., Damen, F.C., Sui, Y., Valyi-Nagy, T., Elliott, M.A., Zhou, X.J., 2015. The feasibility of quantitative MRI of perivascular spaces at 7 T. J. Neurosci. Methods 256, 151–156.

Choi, Y., Nam, Y., Choi, Y., Kim, J., Jang, J., Ahn, K.J., Kim, B.s., Shin, N.Y., 2020. MRI-visible dilated perivascular spaces in healthy young adults: a twin heritability study. Hum. Brain Mapp. 41, 5313–5324.

Donahue, E.K., Bui, V., Foreman, R.P., Duran, J.J., Venkadesh, S., Choupan, J., Van Horn, J.D., Alger, J.R., Jakowec, M.W., Petzinger, G.M., 2022. Magnetic resonance spectroscopy shows associations between neurometabolite levels and perivascular space volume in Parkinson's disease: a pilot and feasibility study. Neuroreport 33, 291–296.

Donahue, E.K., Murdos, A., Jakowec, M.W., Sheikh-Bahaei, N., Toga, A.W., Petzinger, G. M., Sepehrband, F., 2021. Global and regional changes in perivascular space in idiopathic and familial Parkinson's disease. Mov. Disord. 36, 1126–1136.

Doubal, F.N., MacLullich, A.M., Ferguson, K.J., Dennis, M.S., Wardlaw, J.M., 2010. Enlarged perivascular spaces on MRI are a feature of cerebral small vessel disease. Stroke 41, 450–454.

Dubost, F., Yilmaz, P., Adams, H., Bortsova, G., Ikram, M.A., Niessen, W., Vernooij, M., de Bruijne, M., 2019. Enlarged perivascular spaces in brain MRI: automated quantification in four regions. Neuroimage 185, 534–544.

Duering, M., Biessels, G.J., Brodtmann, A., Chen, C., Cordonnier, C., de Leeuw, F.E., Debette, S., Frayne, R., Jouvent, E., Rost, N.S., Ter Telgte, A., Al-Shahi Salman, R., Backes, W.H., Bae, H.J., Brown, R., Chabriat, H., De Luca, A., deCarli, C., Dewenter, A., Doubal, F.N., Ewers, M., Field, T.S., Ganesh, A., Greenberg, S., Helmer, K.G., Hilal, S., Jochems, A.C.C., Jokinen, H., Kuijf, H., Lam, B.Y.K., Lebenberg, J., MacIntosh, B.J., Maillard, P., Mok, V.C.T., Pantoni, L., Rudilosso, S., Satizabal, C.L., Schirmer, M.D., Schmidt, R., Smith, C., Staals, J., Thrippleton, M.J., van Veluw, S.J., Vemuri, P., Wang, Y., Werring, D., Zedde, M., Akinyemi, R.O., Del Brutto, O.H., Markus, H.S., Zhu, Y.C., Smith, E.E., Dichgans, M., Wardlaw, J.M., 2023. Neuroimaging standards for research into small vessel disease-advances since 2013. Lancet Neurol. 22, 602–618.

Duperron, M.-G., Knol, M.J., Le Grand, Q., Evans, T.E., Mishra, A., Tsuchida, A., Roshchupkin, G., Konuma, T., Trégouët, D.-A., Romero, J.R., 2023. Genomics of perivascular space burden unravels early mechanisms of cerebral small vessel disease. Nat. Med. 29, 950–962.

Evans, T.E., Knol, M.J., Schwingenschuh, P., Wittfeld, K., Hilal, S., Ikram, M.A., Dubost, F., Van Wijnen, K.M., Katschnig, P., Yilmaz, P., 2023. Determinants of perivascular spaces in the general population: a pooled cohort analysis of individual participant data. Neurology 100, e107–e122.

Francis, F., Ballerini, L., Wardlaw, J.M., 2019. Perivascular spaces and their associations with risk factors, clinical disorders and neuroimaging features: a systematic review and meta-analysis. Int. J. Stroke 14, 359–371.

Griffanti, L., Zamboni, G., Khan, A., Li, L., Bonifacio, G., Sundaresan, V., Schulz, U.G., Kuker, W., Battaglini, M., Rothwell, P.M., 2016. BIANCA (Brain Intensity AbNormality Classification Algorithm): a new tool for automated segmentation of white matter hyperintensities. Neuroimage 141, 191–205.

Hablitz, L.M., Nedergaard, M., 2021. The glymphatic system. Curr. Biol. 31, R1371–R1375.

Huang, P., Zhu, Z., Zhang, R., Wu, X., Jiaerken, Y., Wang, S., Yu, W., Hong, H., Lian, C., Li, K., 2021. Factors associated with the dilation of perivascular space in healthy elderly subjects. Front. Aging Neurosci. 13, 624732.

Isensee, F., Jaeger, P.F., Kohl, S.A., Petersen, J., Maier-Hein, K.H., 2021. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nat. Methods 18, 203–211.

Jung, E., Chikontwe, P., Zong, X., Lin, W., Shen, D., Park, S.H., 2019. Enhancement of perivascular spaces using densely connected deep convolutional neural network. IEEE Access 7, 18382–18391.

Kim, H.G., Shin, N.-Y., Nam, Y., Yun, E., Yoon, U., Lee, H.S., Ahn, K.J., 2022. MRI-visible dilated perivascular space in the brain by age: the human connectome project. Radiology 306, e213254.

Lan, H., Lynch, K.M., Custer, R., Shih, N.C., Sherlock, P., Toga, A.W., Sepehrband, F., Choupan, J., 2023. Weakly supervised perivascular spaces segmentation with salient guidance of Frangi filter. Magn. Reson. Med. 89, 2419–2431.

Li, H., Jacob, M.A., Cai, M., Kessels, R.P., Norris, D.G., Duering, M., De Leeuw, F.-E., Tuladhar, A.M., 2024. Perivascular spaces, diffusivity along perivascular spaces, and free water in cerebral small vessel disease. Neurology 102, e209306.

Li, Y., Zhu, Z., Chen, J., Zhang, M., Yang, Y., Huang, P., 2020. Dilated perivascular space in the midbrain may reflect dopamine neuronal degeneration in Parkinson's disease. Front. Aging Neurosci. 12, 161.

Lian, C., Zhang, J., Liu, M., Zong, X., Hung, S.C., Lin, W., Shen, D., 2018. Multi-channel multi-scale fully convolutional network for 3D perivascular spaces segmentation in 7T MR images. Med. Image Anal. 46, 106–117.

Miyata, M., Kakeda, S., Iwata, S., Nakayamada, S., Ide, S., Watanabe, K., Moriya, J., Tanaka, Y., Korogi, Y., 2017. Enlarged perivascular spaces are associated with the disease activity in systemic lupus erythematosus. Sci. Rep. 7, 12566.

Nedergaard, M., Goldman, S.A., 2020. Glymphatic failure as a final common pathway to dementia. Science 370, 50–56.

Pham, W., Lynch, M., Spitz, G., O'Brien, T., Vivash, L., Sinclair, B., Law, M., 2022. A critical guide to the automated quantification of perivascular spaces in magnetic resonance imaging. Front. Neurosci. 16, 1021311.

Ramirez, J., Berezuk, C., McNeely, A.A., Scott, C.J., Gao, F., Black, S.E., 2015. Visible Virchow-Robin spaces on magnetic resonance imaging of Alzheimer's disease patients and normal elderly from the Sunnybrook Dementia Study. J. Alzheimers Dis. 43, 415–424.

Ranti, D., Murrough, J., Balchandani, P., Morris, L., 2021. Trauma exposure and perivascular spaces in depression: a 7-Tesla MRI study. Biol. Psychiatry 89, S188–S189.

Rashid, T., Liu, H., Ware, J.B., Li, K., Romero, J.R., Fadaee, E., Nasrallah, I.M., Hilal, S., Bryan, R.N., Hughes, T.M., 2023. Deep learning based detection of enlarged perivascular spaces on brain MRI. Neuroimage: Reports 3, 100162.

Rasmussen, M.K., Mestre, H., Nedergaard, M., 2018. The glymphatic pathway in neurological disorders. Lancet Neurol. 17, 1016–1024.

Schwartz, D.L., Boespflug, E.L., Lahna, D.L., Pollock, J., Roese, N.E., Silbert, L.C., 2019. Autoidentification of perivascular spaces in white matter using clinical field strength T1 and FLAIR MR imaging. Neuroimage 202, 116126.

Sepehrband, F., Barisano, G., Sheikh-Bahaei, N., Cabeen, R.P., Choupan, J., Law, M., Toga, A.W., 2019. Image processing approaches to enhance perivascular space visibility and quantification using MRI. Sci. Rep. 9, 12351.

Sepehrband, F., Barisano, G., Sheikh-Bahaei, N., Choupan, J., Cabeen, R.P., Lynch, K.M., Crawford, M.S., Lan, H., Mack, W.J., Chui, H.C., 2021. Volumetric distribution of perivascular space in relation to mild cognitive impairment. Neurobiol. Aging 99, 28–43.

Smith, E., 2022. Perivascular Spaces: clinically Relevant but Underappreciated. AAN Enterprises 1019–1020.

Sudre, C.H., Anson, B.G., Ingala, S., Lane, C.D., Jimenez, D., Haider, L., Varsavsky, T., Smith, L., Ourselin, S., Jäger, R.H., 2019. 3D multirater RCNN for multimodal multiclass detection and characterisation of extremely small objects. In: International Conference on Medical Imaging with Deep Learning. PMLR, pp. 447–456.

Sudre, C.H., Van Wijnen, K., Dubost, F., Adams, H., Atkinson, D., Barkhof, F., Birhanu, M. A., Bron, E.E., Camarasa, R., Chaturvedi, N., 2024. Where is VALDO? VAscular lesions detection and segmentation challenge at MICCAI 2021. Med. Image Anal. 91, 103029.

Tidwell, J., Taylor, J., Collins, H., Chamberlin, J., Barisano, G., Sepehrband, F., Turner, M., Gauthier, G., Mulder, E., Gerlach, D., 2023. Longitudinal changes in cerebral perfusion, perivascular space volume, and ventricular volume in a healthy cohort undergoing a spaceflight analog. Am. J. Neuroradiol. 44, 1026–1031.

Wang, S., Huang, P., Zhang, R., Hong, H., Jiaerken, Y., Lian, C., Yu, X., Luo, X., Li, K., Zeng, Q., 2021. Quantity and morphology of perivascular spaces: associations with vascular risk factors and cerebral small vessel disease. J. Magn. Reson. Imaging 54, 1326–1336.

Wardlaw, J.M., Benveniste, H., Nedergaard, M., Zlokovic, B.V., Mestre, H., Lee, H., Doubal, F.N., Brown, R., Ramirez, J., MacIntosh, B.J., 2020. Perivascular spaces in the brain: anatomy, physiology and pathology. Nat. Rev. Neurol. 16, 137–153.

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., Oostenbrugge, R., Pantoni, L., Speck, O., Stephan, B.C., Teipel, S., Viswanathan, A., Werring, D., Chen, C., Smith, C., van Buchem, M., Norrving, B., Gorelick, P.B., Dichgans, M., nEuroimaging, S.T.f.R.V.c.o., 2013. Neuroimaging standards for research into small vessel disease and its contribution to ageing and neurodegeneration. Lancet Neurol. 12, 822–838.

Waymont, J.M., Valdes Hernandez, M.D.C., Bernal, J., Duarte Coello, R., Brown, R., Ballerini, L., Chappell, F.M., Wardlaw, J.M., 2024. A systematic review and meta-analysis of automated methods for quantifying enlarged perivascular spaces in the Brain. medRxiv, 2024.2003. 2004.24303705.

Williamson, B.J., Khandwala, V., Wang, D., Maloney, T., Sucharew, H., Horn, P., Haverbusch, M., Alwell, K., Gangatirkar, S., Mahammedi, A., 2022. Automated grading of enlarged perivascular spaces in clinical imaging data of an acute stroke cohort using an interpretable, 3D deep learning framework. Sci. Rep. 12, 788.

Zeng, Q., Li, K., Luo, X., Wang, S., Xu, X., Jiaerken, Y., Liu, X., Hong, L., Hong, H., Li, Z., 2022. The association of enlarged perivascular space with microglia-related inflammation and Alzheimer's pathology in cognitively normal elderly. Neurobiol. Dis. 170, 105755.

Zhu, Y.-C., Tzourio, C., Soumaré, A., Mazoyer, B., Dufouil, C., Chabriat, H., 2010. Severity of dilated Virchow-Robin spaces is associated with age, blood pressure, and MRI markers of small vessel disease: a population-based study. Stroke 41, 2483–2490.