HHS Public Access

Author manuscript

Neuroimage Rep. Author manuscript; available in PMC 2024 March 07.

Published in final edited form as:

Neuroimage Rep. 2023 March ; 3(1): . doi:10.1016/j.ynirp.2023.100162.

Deep Learning Based Detection of Enlarged Perivascular Spaces on Brain MRI

Tanweer Rashid[a,b,$], **Hangfan Liu**[a,b,$], **Jeffrey B. Ware**[c], **Karl Li**[a], **Jose Rafael Romero**[d], **Elyas Fadaee**[a], **Ilya M. Nasrallah**[b,c], **Saima Hilal**[e], **R. Nick Bryan**[c,f], **Timothy M. Hughes**[g], **Christos Davatzikos**[c], **Lenore Launer**[h], **Sudha Seshadri**[a], **Susan R. Heckbert**[i], **Mohamad Habes**[a,b,*]

[a]Neuroimage Analytics Laboratory and Biggs Institute Neuroimaging Core, Glenn Biggs Institute for Neurodegenerative Disorders, University of Texas Health Science Center at San Antonio, San Antonio, TX, USA

[b]Center for Biomedical Image Computing and Analytics, University of Pennsylvania, Philadelphia, PA, USA

[c]Department of Radiation Oncology, University of Pennsylvania, Philadelphia, PA, USA

[d]Department of Neurology, School of Medicine, Boston University, Boston, MA, USA

[e]Saw Swee Hock School of Public Health, National University of Singapore and National University Health System, Singapore

[f]Department of Diagnostic Medicine, Dell Medical School, University of Texas at Austin, Austin, TX, USA

[g]Department of Internal Medicine and Department of Epidemiology and Prevention, Wake Forest School of Medicine, Winston-Salem, NC, USA

[h]Laboratory of Epidemiology and Population Sciences, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA

[i]Department of Epidemiology and Cardiovascular Health Research Unit, University of Washington, Seattle, WA, USA

Abstract

Deep learning has been demonstrated effective in many neuroimaging applications. However, in many scenarios, the number of imaging sequences capturing information related to small vessel disease lesions is insufficient to support data-driven techniques. Additionally, cohort-based

[*]Corresponding author.: habes@uthscsa.edu.
[$]Equal contributing first authors.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

studies may not always have the optimal or essential imaging sequences for accurate lesion detection. Therefore, it is necessary to determine which imaging sequences are crucial for precise detection. This study introduces a deep learning framework to detect enlarged perivascular spaces (ePVS) and aims to find the optimal combination of MRI sequences for deep learning-based quantification. We implemented an effective lightweight U-Net adapted for ePVS detection and comprehensively investigated different combinations of information from SWI, FLAIR, T1-weighted (T1w), and T2-weighted (T2w) MRI sequences. The experimental results showed that T2w MRI is the most important for accurate ePVS detection, and the incorporation of SWI, FLAIR and T1w MRI in the deep neural network had minor improvements in accuracy and resulted in the highest sensitivity and precision (sensitivity =0.82, precision =0.83). The proposed method achieved comparable accuracy at a minimal time cost compared to manual reading. The proposed automated pipeline enables robust and time-efficient readings of ePVS from MR scans and demonstrates the importance of T2w MRI for ePVS detection and the potential benefits of using multimodal images. Furthermore, the model provides whole-brain maps of ePVS, enabling a better understanding of their clinical correlates compared to the clinical rating methods within only a couple of brain regions.

Keywords

MRI; Deep learning; Enlarged Perivascular Space

1. Introduction

Enlargement of perivascular, or Virchow-Robin, spaces (Doubal, MacLullich, Ferguson, Dennis, & Wardlaw, 2010; Wardlaw et al., 2020) can be a manifestation of cerebral small vessel disease and dysfunction of perivascular drainage routes. Perivascular spaces are fluid-filled spaces that surround arteries, arterioles, veins, and venules (Wardlaw et al., 2013) in the brain. They are generally microscopic in size but with increasing age and/or pathologies may become enlarged and visible, i.e. enlarged perivascular spaces (ePVS) (Hou et al., 2017; Potter, Doubal, et al., 2015; Wardlaw et al., 2020; Wardlaw et al., 2013). Typically, ePVS appear as bright or hyperintense linear or curvilinear structures when running parallel to the imaging plane and ellipsoidal or dot-like when perpendicular to the imaging plane on T2-weighted (T2w) magnetic resonance imaging (MRI) (Wardlaw et al., 2020; Wardlaw et al., 2013). When perivascular spaces are enlarged, they become visible on routine structural MRI, typically with a diameter less than 3mm, but can reach up to 10–20 mm in regions such as the basal ganglia (Wardlaw et al., 2013). While ePVS can be evaluated on T1-weighted (T1w) and T2w sequences, they are easier to visualize and quantify using T2w imaging (Ballerini et al., 2018; Potter, Chappell, Morris, & Wardlaw, 2015).

Many detection/segmentation methods have been proposed (Ballerini et al., 2018; Hou et al., 2017; Lian et al., 2018; Wang et al., 2016; Zhang et al., 2017) which rely on T2w exclusively for detection/segmentation of ePVS. However, it is still unclear if models relying on a single modality such as T2w could account for similar-appearing brain lesions such as white matter hyperintensities (WMH), lacunes and infarcts. WMH are hyperintense on T2w

sequences and can appear as isointense or hypointense on T1w sequences, lacunes are round or ovoid subcortical fluid-filled cavities of between 3 mm and 15 mm in diameter, while infarcts are neuroimaging evidences of recent infarction in the territory of one perforating arteriole (Wardlaw et al., 2013).

In this paper, we aim to evaluate the feasibility and effectiveness of an automated deep learning-based method for segmenting ePVS using multiple MRI sequences from a subset of participants in the Multi-Ethnic Study of Atherosclerosis (MESA) cohort. The brain data collected by the MESA Atrial Fibrillation (AFib) (Austin et al., 2022; Bild et al., 2002; Burke, Lima, Wong, & Narula, 2016; Olson, Bild, Kronmal, & Burke, 2016) ancillary study at Exam 6 offer a unique and rich dataset of high-quality brain MRI at clinical field strength and high spatial resolution (1 mm isotropic images). We aim to evaluate the accuracy and reliability of ePVS segmentation in the presence or absence of T2w MRI, and when T2w is combined with other MRI sequences. We used a variation of our method, previously developed using MESA brain MRI data for fully automated detection of cerebral microbleeds and non-haemorrhage iron deposits in the basal ganglia (Rashid et al., 2021), and investigate the optimal strategy of combining information from susceptibility weighted imaging (SWI), fluid-attenuated inversion recovery (FLAIR), T1w and T2w MRI sequences. A set of ePVS segmentations by a human expert served as the gold standard for model training.

Automation is ideal in large cohort studies for feasibility and to improve reproducibility and to reduce human error (Hurtz et al., 2019). Accurate and reliable methods are also essential for deriving rich datasets from large cohorts to study associations with demographic, cognitive and vascular risk factors (Mohamad Habes et al., 2016; M Habes et al., 2016; Habes et al., 2021), or to refine the development of new methods (Liu, Rashid, & Habes, 2020; Liu et al., 2021).

To the best of our knowledge, this study is among the first to comprehensively evaluate multimodal imaging for ePVS detection with deep learning. The main contributions of this paper include:

1. Development of an effective deep learning scheme with data fusion for accurate ePVS segmentation.
2. Application of the proposed model to the whole brain, instead of selective regions.
3. Investigation of the use of different sequences for optimal performance.

2. Related Works

Previous ePVS segmentation methods typically adopt conventional machine learning techniques such as vessel enhancement filters (Ballerini et al., 2018) and support vector machines (SVM) (González-Castro et al., 2017). Ballerini et al. trained a model on T2-contrast MR images (Ballerini et al., 2018) and evaluated it by categorical scores (Potter, Chappell, et al., 2015). González-Castro et al. applied SVM classifier with bag of visual words-based descriptors to the T2-weighted MR images with a focus on the

basal ganglia (González-Castro et al., 2017). Wang et al. developed a semi-automatic computational method that extracts ePVS on bilateral ovoid basal ganglia on intensity-normalized T2w MRI (Wang et al., 2016). Meanwhile, some works exploited handcrafted features as predictors, for example, Boespflug et al. used signal intensities and morphologic characterizations including width, volume and linearity (Boespflug et al., 2018), while Ramirez et al. used set localized intensity thresholds for quantification of perivascular spaces (Ramirez et al., 2015), and Zhang et al. proposed vascular feature based structured learning for 3-dimensional ePVS segmentation using T2w data (Zhang et al., 2017). Besides, to facilitate these models, Sepehrband et al. combined T1- and T2w images to enhance PVS contrast to intensify the visibility (Sepehrband et al., 2019).

With the recent success of deep learning techniques (Liu et al., 2021; Mou, Zhang, Fan, Liu, & Wang, 2021; Song & Liu, 2021; Yin et al., 2019), some deep neural network models were proposed for ePVS segmentation. For instance, Boutinaud et al. developed a deep learning algorithm based on an autoencoder and a U-shaped network for the 3-dimensional segmentation of ePVS in deep white matter and basal ganglia using T1-weighted MRI data (Boutinaud et al., 2021), and Lian et al. proposed a fully convolutional neural network using 7T T2-weighted MRI for efficient segmentation of ePVS (Lian et al., 2018), Dubost et al. implemented separate convolutional neural networks for midbrain, hippocampi, basal ganglia and centrum semiovale, trained on T2-contrast MRI to quantify PVS (Dubost et al., 2019), Sudre et al. redesigned the region-based convolutional neural networks model to jointly detect and characterize markers of age-related neurovascular changes (Sudre et al., 2019), and Jung et al. presented a deep 3-dimensional convolutional neural network with densely connected networks with skip connections for ePVS enhancement of 7T MRI (Jung et al., 2019). In general, these works did not investigate how to fully utilize different sources of information for improved ePVS detection on the whole brain based on deep data-driven techniques. Furthermore, these prior studies mostly used 7T MRI which is less available in the clinic compared to the more standard 3T machines. 3T MRI is more conventional and widely available, so a deep learning model tailored for 3T MRI a more cost-effective choice.

3. Materials and Methods

The key point of the proposed scheme is a deep fusion of information from different MRI sequences. An overview of the whole procedure is summarized in Fig. 1. Standard image processing techniques were first applied to raw MRI data of different sequences a subset of the MESA cohort, including inhomogeneity correction, reorientation, smoothing and filtering, brain masking and skull-stripping, followed by gray matter (GM), white matter (WM) and cerebrospinal fluid (CSF) segmentation. Then the participants' MRI were registered to SWI. Preprocessed MRI data were manually segmented to obtain the ground truth used for model training with leave-one-out cross-validation.

3.1. Data

The training data included 21 participants, which are randomly selected from the MESA cohort. For T1w, T2w and FLAIR images, the MESA study collected 3D isotropic MRI scans at 6 different sites with Siemens scanners (Skyra with a 20-channel head coil, Prisma

and Prisma Fit with 32-channel head coil). Our training data included participants from all these sites and all the scanner models, thus ensuring generalizability within the MESA cohort. The MRI scan parameters are shown in Table 1.

The ages of the 21 participants range from 64 to 94 years with an average of 78.7 years, and 12 of them are female. The average total number of individual lesions per participant is more than 683. The ePVS segmentation of these participants was performed by an experienced radiologist (JBW) and served as ground truth for model training (see supplementary material Table S1). The manual segmentation was performed using co-registered T2w, T1w and FLAIR images to ensure reduced likelihood of false positives such as WMH or lacunes (Wardlaw et al., 2013) being present in the ground truth. The average time needed to complete a manual segmentation of ePVS for the whole brain was around 24 hours per participant. For model training and evaluation, we used FLAIR, SWI, T1w and T2w images, which were reoriented, N4 bias corrected (Avants, Tustison, & Song, 2009) and skull-stripped (Doshi et al., 2016). The SWI phase mask was generated from the phase images using a high-pass filter of size 64 × 64 in order to remove artifacts, and the SWI was generated by multiplying the magnitude image with the phase mask (Haacke, Mittal, Wu, Neelavalli, & Cheng, 2009; Haacke, Xu, Cheng, & Reichenbach, 2004. For creation of the reference annotation and machine-based inference, only the SWI image with the shortest echo time (TE=7.5 ms) was used since SWI acquired with longer echo times are noisier. Examples of ePVS on the different sequences are shown in Fig. 2. The MRI scans used in this study have high spatial resolution, making it possible to detect small ePVS, although in clinical settings, the slice thickness is larger to allow for less scan time, so small lesions occurring between slices may not be visible.

3.2. Deep Fusion of Different Sequences

Suppose $f_U: \mathbb{R}^{n \times S} \rightarrow \mathbb{R}^S$ is a nonlinear function with a set of learnable parameters U, where n is the number of MRI Sequences used and S is the size of the images, f maps the n images to voxel-wise labels indicating whether the voxel contains ePVS or not. In this study, f_U is implemented as a multi-channel deep neural network (Rashid et al., 2021), which is a variation of the standard U-Net (Ronneberger, Fischer, & Brox, 2015) and has been demonstrated superior compared to conventional U-Net for small lesions (Rashid et al., 2021). A typical U-Net is made up of a down-sampling or encoding path and a symmetric up-sampling or decoding path. The down-sampling path consists of a series of convolutional blocks, normalization blocks, activation blocks and pooling blocks. The up-sampling path consists of a series of convolutional blocks, normalization blocks, activation blocks and transpose convolutional blocks. The feature map of each corresponding down-sampling path and up-sampling path are concatenated.

The proposed scheme could perform a deep fusion of information from different sequences. The ePVS detection/segmentation model fuses information from T2w, SWI, FLAIR and T1w images through the multi-channel U-Net. It was designed in a scalable manner, i.e., the network using T2 only was basically a single-channel U-Net, and can be easily expanded to include multiple sequences. The manual segmentations by the human expert were used to train the deep learning model using leave-one-out cross-validation. To be specific, in each

iteration of the leave-one-out cross validation, we use data from 20 subjects for network training and data from 1 subject for testing. Also, from the 20 subjects used in training, 4 were used exclusively for within-training validation.

Each 3-dimensional (3D) scan was cut into 2D axial slices, which underwent data augmentation through combinations of geometric transforms such as rotations, translations, up-down and left-right flips. In each experiment, the axial T2w slice (along with the corresponding axial slices in other MRIs) and corresponding axial ground truth slice were augmented. For translations, a set of two random floating-point numbers tx and ty (representing the amount of shift per axis) were generated within the range [−45, 45] and used to translate the image slice(s) and the corresponding slice of the ground truth. This range was chosen empirically so that most of the brain would be visible in the translated image. A total of 10 random floating-point numbers per axis were generated, resulting in 10×10=100 translations for each slice. For rotations, a set of random floating-point numbers d (representing the rotation in degree) were generated within the range [1, 60], and the image slice(s) and the slices with ground truth were rotated using both $+d$ and $-d$. The regions of the crops that were located outside the image matrix were padded with edge values. A total of 16 random floating-point numbers were used, resulting in 16×2=32 rotations. The same set of transforms were applied to the flipped images. For example, a single T2w MRI image having 96 axial slices resulted in 23880 axial slices after data augmentation. These augmented data were fed into the neural network as data samples.

We aimed to train multi-class models, where predicted classes were background and ePVS, using the following combinations of imaging sequences: (1) T2w-only, (2) T2w and FLAIR, (3) T2w, T1w and FLAIR, (4) T2w, T1w, FLAIR and SWI, (5) T2w and T1w, (6) FLAIR only, (7) T1w only, (8) T1w and FLAIR. The training time for the deep learning models was 3 to 8 days. Once model training was completed, the trained models can predict whole-brain ePVS in less than 30 seconds.

3.3. Analysis of Detection Results

The accuracy of these models was based on three parameters: sensitivity S, precision P, and magnitude accuracy A, which are defined as

$$S = TP/(TP + FN),$$

$$P = TP/(TP + FP),$$

$$A = \sqrt{S^2 + P^2},$$

where TP is the number of true positives, FN is the number of false negatives, and FP stands for false positives.

We also selected metrics effective for small lesions like ePVS where shape information and volume is important. The ePVS could be as small as one voxel. The analysis included Bland-Altman plots and scatterplots of ePVS count and volumes (prediction vs expert labelled data), as well as sensitivity and precision based on center of mass of the lesions. We also assessed performance using intra-class correlation coefficients (ICC) (Shrout & Fleiss, 1979), volumetric similarity (Ramaswamy Reddy, Prasad, & Reddy, 2013), area under the curve (AUC) from receiver operator curves, Hausdorff distance (Rockafellar & Wets, 2009) and Mahalanobis distance (Xiang, Nie, & Zhang, 2008). For ICC, we used the method of (Shrout & Fleiss, 1979) with a two-way random model, absolute agreement, single measure. Hausdorff distance calculates the distance between two point sets that correspond to ground truth labels and segmentations respectively, while Mahalanobis distance is a multivariate distance metric that measures the distance between a point and a distribution and is particularly effective for classification on highly imbalanced datasets. The mean metrics are obtained by averaging over subjects, e.g., suppose S_i is the sensitivity obtained by testing subject i ($i = 1,2, \ldots, 21$), then average sensitivity $\bar{S} = \frac{1}{21}\sum_{i=1}^{21} S_i$.

4. Results

The mean evaluation metrics with corresponding standard errors of all subjects including sensitivity, precision, magnitude accuracy, ICC, volumetric similarity, AUC, Hausdorff distance and Mahanabolis distance are shown in Table 2. The results indicate that T2w MRI is the most informative, with the best performance of any single sequence and near optimal for several measures. For most measures, the combination of T2w, FLAIR, T1w and SWI achieved the best performance. Adding SWI to the combination of the other 3 sequences offered minimal overall gain but improved ICC.

Fig. 3 displays the correlations between the number of predicted lesions and that of ground-truth lesions. The highest correlations are achieved by using T2w. Fig. 4 plots the points located by pairs (S, P) from all the participants, and indicates that by including T2w, FLAIR and T1w the model could attain highest magnitude accuracy, which is reflected by the distance between ($\bar{S}$,$\bar{P}$) and (0, 0) in the figure, where $\bar{S}$ and $\bar{P}$ are the median sensitivity and median precision respectively. Fig. 5 shows the Bland-Altman plots of number of lesions, demonstrating that the mean difference between the prediction and the gold standard as well as the random fluctuations around the mean reached the minimal when using T2w only, and remained low when incorporating other sequences. Fig. 6 displays the correlations between the volume of predicted lesions and volume of the ground truth, reaffirming the importance of using T2w for ePVS segmentation. Fig. 7 shows the Bland-Altman plots of lesion volumes, indicating that combination of T2w, FLAIR, T1w and SWI could attain better results than using FLAIR, T1w only, since the mean difference and the fluctuations were minimal when combining T2w, T1w, FLAIR and SWI, and were significantly smaller when T2w is included.

Based on such observations, we can see that although T1 and FLAIR are more standard research sequences, for ePVS ratings using these two sequences only are not nearly as accurate as including T2w, and incorporating other sequences did not improve results

significantly. However, utilization of information from different modalities enables the model to effectively distinguish ePVS from mimics like white matter lesions and lacunes, as demonstrated in Fig. 8.

5. Discussion and Conclusions

Enlarged perivascular spaces (ePVS) are increasingly recognized as a subclinical biomarker for brain health and disease, including cerebrovascular disease, and therefore quantification is of interest to the research community. Manual quantification of individual ePVS is extremely time consuming (Ramirez et al., 2015; Wang et al., 2016), operator-dependent and may not reflect accurately the true burden of ePVS. Data-driven automated systems, including deep learning models, provide a promising way to generate robust, reproducible, and rapid quantification of ePVS from brain MRI scans, and when training dataset is limited, light-weight networks can be sufficient to achieve accurate prediction (Peng, Gong, Beckmann, Vedaldi, & Smith, 2021).

Automated ePVS quantification is challenging due to the existence of mimics like lacunes and white matter lesions, which may lead to false positive measurements. Furthermore, in many scenarios, the number of neuroimaging data samples could be insufficient to support data-driven systems. Such problems still remain in recently published deep learning methods. In general, there are several limitations: 1) It is still under question as to whether one single modality could be sufficiently informative for ePVS quantification; 2) The advantages of combining different sequences for the application is not investigated; 3) Existing methods generally use 7T MRI whereas 3T MRI is more available and accessible in practice; 4) Existing methods were only applied to selective regions rather than the whole brain.

To address these issues, this study aims to fully exploit the informative 3T MRI data available by jointly utilizing different sequences, and investigate the optimal strategy of fusing information from different sequences in the deep learning framework for ePVS segmentation, which could be applied to the whole brain. Specifically, since the number of data samples is often limited, it is of great importance to make full use of the data available, and the fusion of information from different sequences could be an effective solution. The deep learning model adopts a light-weight multi-channel variation of the U-Net tailored for the application. The experimental results demonstrate that the combination of T2w, FLAIR, T1w and SWI leads to best segmentation performance, and that performance with T1w alone is worse than T2w alone for detecting ePVS. Our results suggest that if quantification of ePVS is of interest, prospective research studies should include T2w imaging in a brain MRI protocol. T1w images, which are by far more prevalent in research studies due to utility in brain tissue segmentation, should be expected to provide less accurate quantification of ePVS.

For regional evaluation, we derived several regions based on the existing MUlti-atlas region Segmentation utilizing Ensembles (MUSE) (Doshi et al., 2016), as shown in Fig. S3 of the supplementary material. Based on these regions, we did the same sensitivity and precision calculations for each individual region in all the experiments. The metrics including mean

sensitivity and precision etc. are in Table S2 ~ S8 of the supplementary material. We can see that in the basal ganglia, the sensitivity and precision is high for all experiments, even when using only T1 or FLAIR. This suggests we can get reliable and accurate ePVS readings in the basal ganglia using only T1 and/or FLAIR. On the other hand, we see that the sensitivity and precision is poor in the hippocampus and temporal regions. This is because of false positives due to the presence of blood vessels prevalent in those regions. Currently, the most clinically relevant regions for ePVS readings are the basal ganglia, centrum semiovale and maybe the midbrain (Wardlaw et al., 2020; Wardlaw et al., 2013). So our experiments are showing that our models can make accurate predictions in the basal ganglia and the centrum semiovale, even when T2w is absent.

In this work we chose 2D slices over 3D because the ground truth ePVS labels were manually labeled for each slice separately. Although theoretically using 3D samples could utilize context from adjacent slices, in practice such connection between adjacent slices might be weak or even misleading when the ground-truth labels are produced slice-wise manually. From this perspective, the main benefit of using 2D dataset is that it could better fit the characteristics of real-world expert-segmented labels. This might be the reason why 2D datasets were demonstrated to be better than the 3D counterparts in some previous works (Srikrishna et al., 2022). Another practical reason for using 2D slices in lieu of 3D images is the exponentially larger GPU memory required to hold all the parameters of a 3D-capable deep learning model. Smaller models may fit commercial GPU memory but may not be capable of extracting the features necessary for effective learning and segmentation.

Although the dataset we used is a subset of the MESA brain MRI study at Exam 6, with available ground truth data, we strived to alleviate the issue of overfitting to the extent possible by performing leave-one-out cross-validation when evaluating our models. Our cross-validation strategy has shown great generalization to the available dataset, but additional experiments with external datasets might improve the generalization of our model. However, that comes with challenges, namely, other studies usually have recruited participants with pathology and may not include similar MRI sequences like T2w, FLAIR and SWI, and may not have the expert-segmented ePVS labels.

In conclusion, the proposed automated pipeline enables robust and time efficient readings of ePVS from MR scans, and demonstrated the importance of T2w MRI for ePVS detection and the insignificant benefit of using multimodal images. It may also provide a potential way to alleviate the issues brought by limitation of data samples. The automated pipeline will help in generating a rich variable set in MESA that will enable examination of ePVS in relation to other risk factors. A limitation of the study is that manual ePVS segmentation from only one expert is available.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This study was supported in part by the National Institute of Health (NIH) grant P30AG066546 (South Texas Alzheimer's Disease Research Center) and grant numbers 1R01AG080821, 5R01HL127659, 1U24AG074855 and the San Antonio Medical Foundation grant SAMF – 1000003860. This MESA research was supported by contracts 75N92020D00001, HHSN268201500 003I, N01-HC-95159, 75N92020D00005, N01-HC-95160, 75N92020D00002, N01-HC-95161, 75N92020D00003, N01-HC-95162, 75N92020D00006, N01-HC-95163, 75N92020D00004, N01-HC-95164, 75N92020D00007, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168 and N01-HC-95169 and grant HL127659 from the National Heart, Lung, and Blood Institute, and by grants UL1-TR-000040, UL1-TR-001079, and UL1-TR-001420 from the National Center for Advancing Translational Sciences (NCATS). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The authors thank the other investigators, the staff, and the participants of the MESA study for their valuable contributions. A full list of participating MESA investigators and institutions can be found at http://www.mesa-nhlbi.org.

REFERENCES

Austin TR, Nasrallah IM, Erus G, Desiderio LM, Chen LY, Greenland P, … Longstreth W Jr (2022). Association of Brain Volumes and White Matter Injury With Race, Ethnicity, and Cardiovascular Risk Factors: The Multi-Ethnic Study of Atherosclerosis. Journal of the American Heart Association, e023159. [PubMed: 35352569]

Avants BB, Tustison N, & Song G (2009). Advanced normalization tools (ANTS). Insight j, 2(365), 1–35.

Ballerini L, Lovreglio R, Hernández M. d. C. V., Ramirez J, MacIntosh BJ, Black SE, & Wardlaw JM (2018). Perivascular spaces segmentation in brain MRI using optimal 3D filtering. Scientific Reports, 8(1), 1–11. [PubMed: 29311619]

Bild DE, Bluemke DA, Burke GL, Detrano R, Diez Roux AV, Folsom AR, … Liu K (2002). Multi-ethnic study of atherosclerosis: objectives and design. American journal of epidemiology, 156(9), 871–881. [PubMed: 12397006]

Boespflug EL, Schwartz DL, Lahna D, Pollock J, Iliff JJ, Kaye JA, … Silbert LC (2018). MR imaging–based multimodal autoidentification of perivascular spaces (mMAPS): automated morphologic segmentation of enlarged perivascular spaces at clinical field strength. Radiology, 286(2), 632–642. [PubMed: 28853674]

Boutinaud P, Tsuchida A, Laurent A, Adonias F, Hanifehlou Z, Nozais V, … 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. Frontiers in neuroinformatics, 15, 29.

Burke G, Lima J, Wong ND, & Narula J (2016). The multiethnic study of atherosclerosis. Global *heart,* 11(3).

Doshi J, Erus G, Ou Y, Resnick SM, Gur RC, Gur RE, … Initiative A s. N. (2016). MUSE: MUlti-atlas region Segmentation utilizing Ensembles of registration algorithms and parameters, and locally optimal atlas selection. Neuroimage, 127, 186–195. [PubMed: 26679328]

Doubal FN, MacLullich AM, Ferguson KJ, Dennis MS, & Wardlaw JM (2010). Enlarged perivascular spaces on MRI are a feature of cerebral small vessel disease. Stroke, 41(3), 450–454. [PubMed: 20056930]

Dubost F, Yilmaz P, Adams H, Bortsova G, Ikram MA, Niessen W, … de Bruijne M (2019). Enlarged perivascular spaces in brain MRI: Automated quantification in four regions. Neuroimage, 185, 534–544. [PubMed: 30326293]

González-Castro V, Valdés Hernández M. d. C., Chappell FM, Armitage PA, Makin S, & Wardlaw JM (2017). Reliability of an automatic classifier for brain enlarged perivascular spaces burden and comparison with human performance. Clinical Science, 131(13), 1465–1481. [PubMed: 28468952]

Habes M, Erus G, Toledo JB, Zhang T, Bryan N, Launer LJ, … Van der Auwera S (2016). White matter hyperintensities and imaging patterns of brain ageing in the general population. Brain, 139(4), 1164–1179. [PubMed: 26912649]

Habes M, Janowitz D, Erus G, Toledo J, Resnick S, Doshi J, … Hosten N (2016). Advanced brain aging: relationship with epidemiologic and genetic risk factors, and overlap with Alzheimer disease atrophy patterns. Translational Psychiatry, 6(4), e775–e775. [PubMed: 27045845]

Habes M, Pomponio R, Shou H, Doshi J, Mamourian E, Erus G, … Bilgel M (2021). The Brain Chart of Aging: Machine-learning analytics reveals links between brain aging, white matter disease, amyloid burden, and cognition in the iSTAGING consortium of 10,216 harmonized MR scans. Alzheimer's & Dementia, 17(1), 89–102.

Hou Y, Park SH, Wang Q, Zhang J, Zong X, Lin W, & Shen D (2017). Enhancement of perivascular spaces in 7 T MR image using Haar transform of non-local cubes and block-matching filtering. Scientific Reports, 7(1), 1–12. [PubMed: 28127051]

Hurtz S, Chow N, Watson AE, Somme JH, Goukasian N, Hwang KS, … Petersen RC (2019). Automated and manual hippocampal segmentation techniques: Comparison of results, reproducibility and clinical applicability. NeuroImage: Clinical, 21, 101574. [PubMed: 30553759]

Jung E, Chikontwe P, Zong X, Lin W, Shen D, & Park SH (2019). Enhancement of perivascular spaces using densely connected deep convolutional neural network. IEEE Access, 7, 18382–18391. [PubMed: 30956927]

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. Medical Image Analysis, 46, 106–117. [PubMed: 29518675]

Liu H, Rashid T, & Habes M (2020). Cerebral Microbleed Detection Via Fourier Descriptor with Dual Domain Distribution Modeling Paper presented at the 2020 IEEE 17th International Symposium on Biomedical Imaging Workshops (ISBI Workshops).

Liu H, Rashid T, Ware J, Jensen P, Austin T, Nasrallah I, … Habes M (2021). Adaptive Squeeze-and-Shrink Image Denoising for Improving Deep Detection of Cerebral Microbleeds Paper presented at the International Conference on Medical Image Computing and Computer-Assisted Intervention.

Mou C, Zhang J, Fan X, Liu H, & Wang R (2021). COLA-Net: Collaborative Attention Network for Image Restoration. IEEE Transactions on Multimedia.

Olson JL, Bild DE, Kronmal RA, & Burke GL (2016). Legacy of MESA. Global heart, 11(3), 269–274. [PubMed: 27741974]

Peng H, Gong W, Beckmann CF, Vedaldi A, & Smith SM (2021). Accurate brain age prediction with lightweight deep neural networks. Medical image analysis, 68, 101871. [PubMed: 33197716]

Potter GM, Chappell FM, Morris Z, & Wardlaw JM (2015). Cerebral perivascular spaces visible on magnetic resonance imaging: development of a qualitative rating scale and its observer reliability. Cerebrovascular Diseases, 39(3–4), 224–231. [PubMed: 25823458]

Potter GM, Doubal FN, Jackson CA, Chappell FM, Sudlow CL, Dennis MS, & Wardlaw JM (2015). Enlarged perivascular spaces and cerebral small vessel disease. International Journal of Stroke, 10(3), 376–381. [PubMed: 23692610]

Ramaswamy Reddy A, Prasad E, & Reddy L (2013). Comparative analysis of brain tumor detection using different segmentation techniques. International Journal of Computer Applications, 82(14), 0975–8887.

Ramirez J, Berezuk C, McNeely AA, Scott CJ, Gao F, & Black SE (2015). Visible Virchow-Robin spaces on magnetic resonance imaging of Alzheimer's disease patients and normal elderly from the Sunnybrook Dementia Study. Journal of Alzheimer's Disease, 43(2), 415–424.

Rashid T, Abdulkadir A, Nasrallah IM, Ware JB, Liu H, Spincemaille P, … Habes M (2021). DEEPMIR: A DEEP neural network for differential detection of cerebral Microbleeds and IRon deposits in MRI. Scientific Reports

Rockafellar RT, & Wets RJ-B (2009). Variational analysis (Vol. 317): Springer Science & Business Media.

Ronneberger O, Fischer P, & Brox T (2015). U-net: Convolutional networks for biomedical image segmentation Paper presented at the International Conference on Medical image computing and computer-assisted intervention.

Sepehrband F, Barisano G, Sheikh-Bahaei N, Cabeen RP, Choupan J, Law M, & Toga AW (2019). Image processing approaches to enhance perivascular space visibility and quantification using MRI. Scientific Reports, 9(1), 1–12. [PubMed: 30626917]

Shrout PE, & Fleiss JL (1979). Intraclass correlations: uses in assessing rater reliability. Psychological bulletin, 86(2), 420. [PubMed: 18839484]

Song Q, & Liu H (2021). Deep Gradient Prior Regularized Robust Video Super-Resolution. Electronics, 10(14), 1641.

Srikrishna M, Heckemann RA, Pereira JB, Volpe G, Zettergren A, Kern S, … Schöll M (2022). Comparison of two-dimensional-and three-dimensional-based U-Net architectures for brain tissue classification in one-dimensional brain CT. Frontiers in Computational Neuroscience, 15, 123.

Sudre CH, Anson BG, Ingala S, Lane CD, Jimenez D, Haider L, … Jäger RH (2019). 3D multirater RCNN for multimodal multiclass detection and characterisation of extremely small objects Paper presented at the International Conference on Medical Imaging with Deep Learning.

Wang X, Hernández M. d. C. V., Doubal F, Chappell FM, Piper RJ, Deary IJ, & Wardlaw JM (2016). Development and initial evaluation of a semi-automatic approach to assess perivascular spaces on conventional magnetic resonance images. Journal of neuroscience methods, 257, 34–44. [PubMed: 26416614]

Wardlaw JM, Benveniste H, Nedergaard M, Zlokovic BV, Mestre H, Lee H, … MacIntosh BJ (2020). Perivascular spaces in the brain: anatomy, physiology and pathology. Nature Reviews Neurology, 16(3), 137–153. [PubMed: 32094487]

Wardlaw JM, Smith EE, Biessels GJ, Cordonnier C, Fazekas F, Frayne R, … Benavente OR (2013). Neuroimaging standards for research into small vessel disease and its contribution to ageing and neurodegeneration. The Lancet Neurology, 12(8), 822–838. [PubMed: 23867200]

Xiang S, Nie F, & Zhang C (2008). Learning a Mahalanobis distance metric for data clustering and classification. Pattern recognition, 41(12), 3600–3612.

Yin S, Peng Q, Li H, Zhang Z, You X, Liu H, … Fan Y (2019). Multi-instance deep learning with graph convolutional neural networks for diagnosis of kidney diseases using ultrasound imaging. Paper presented at the Uncertainty for Safe Utilization of Machine Learning in Medical Imaging and Clinical Image-Based Procedures

Zhang J, Gao Y, Park SH, Zong X, Lin W, & Shen D (2017). Structured learning for 3-D perivascular space segmentation using vascular features. IEEE Transactions on Biomedical Engineering, 64(12), 2803–2812. [PubMed: 28362579]

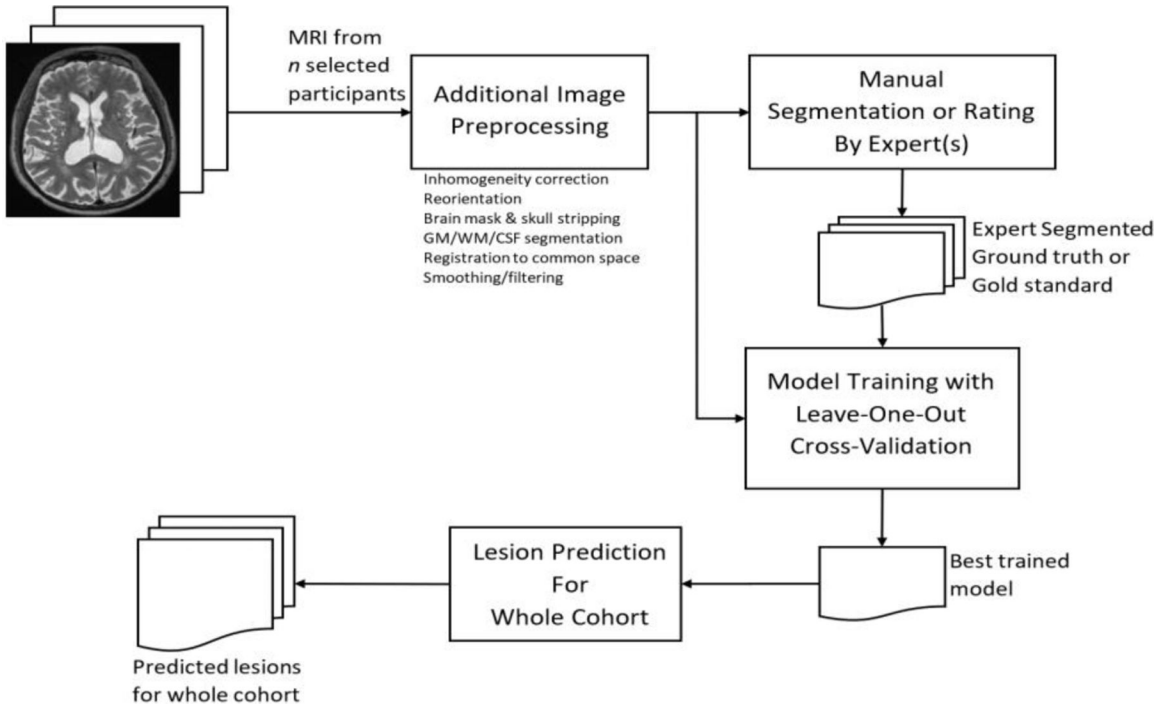

Fig. 1.
Overview of the proposed ePVS detection/segmentation procedure.

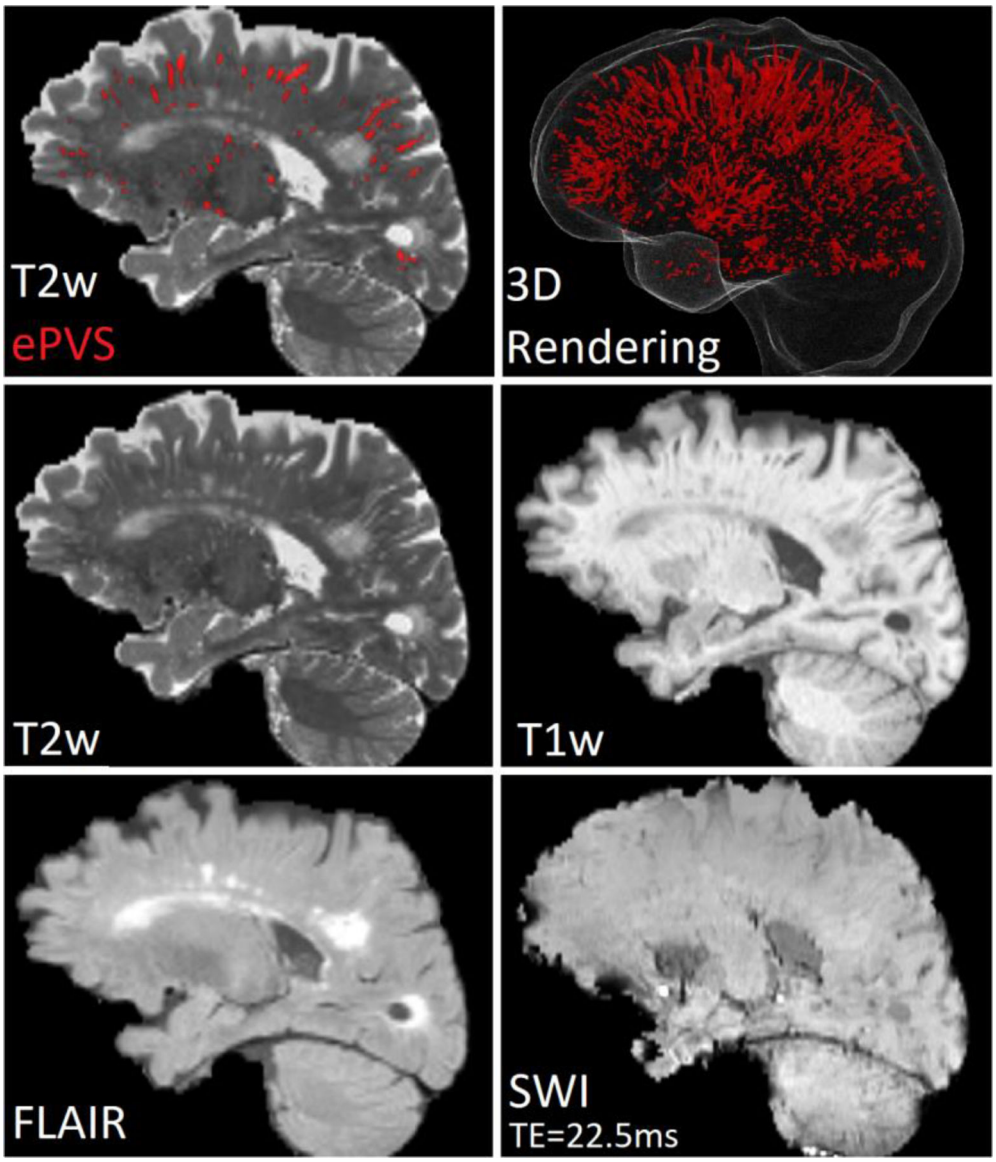

Fig. 2.

Examples of ePVS in different MRI sequences. From top to bottom and left to right: a) A T2w image with ePVS regions marked red. b) 3D rendering of ePVSs, which are marked red. c) The original T2w image without label. d) Corresponding T1w image. e) Corresponding FLAIR image. f) Corresponding SWI image when time of echo is 22.5ms.

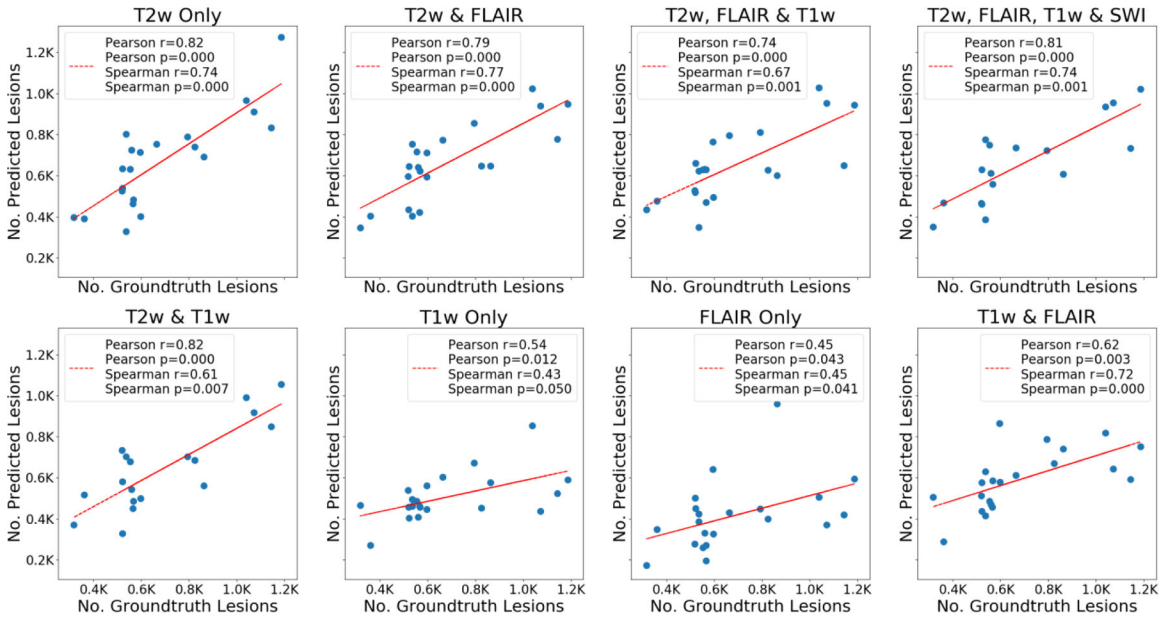

Fig. 3.

Scatterplots of number of groundtruth ePVS vs. number of predicted ePVS per subject, based on which the Person correlation and Spearman correlation (*r*) and the corresponding *p*-values (*p*) are calculated.

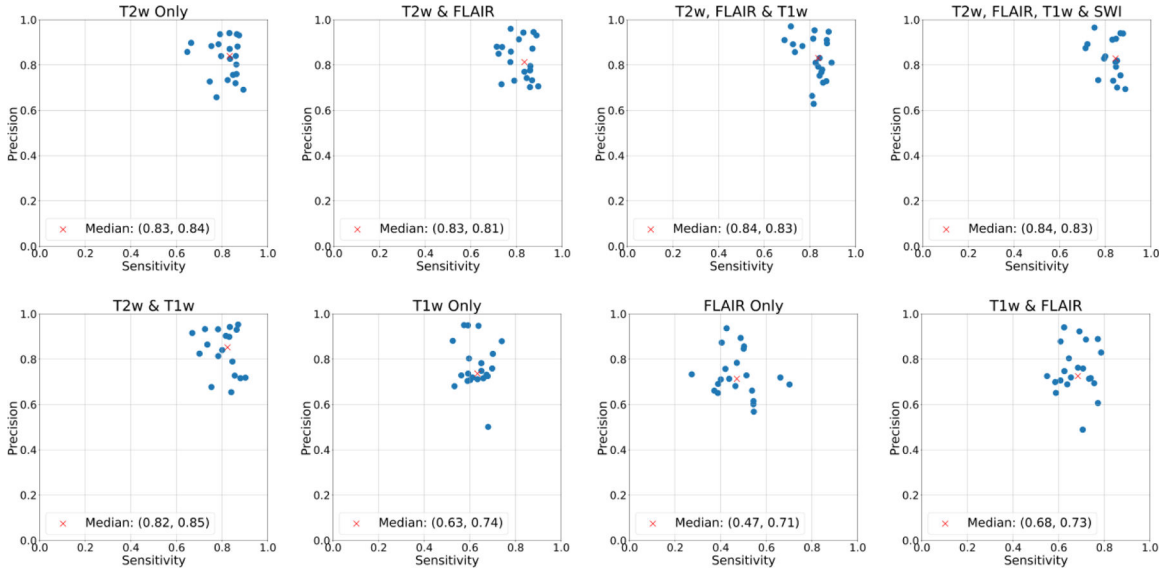

Fig. 4.

Scatterplots of sensitivity vs precision per subject along with the corresponding (median sensitivity, median precision) for different combination of the sequences.

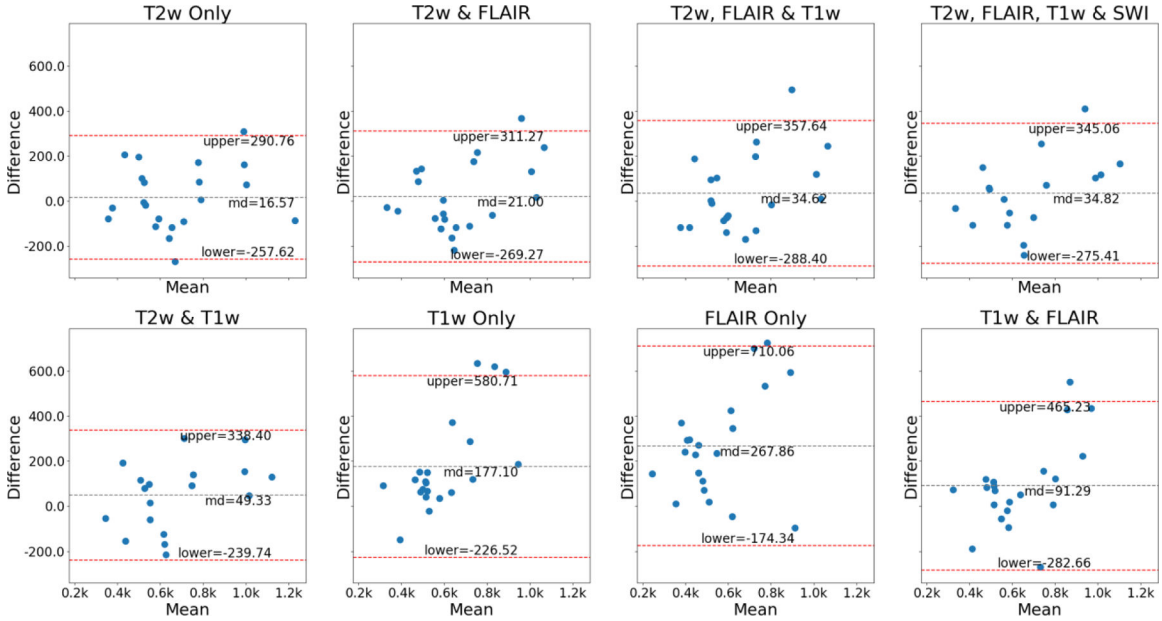

Fig. 5.

Bland-Altman plots of numbers of ePVS for different combinations of the sequences. The plots show the differences between the numbers of predicted ePVSs and those of the groundtruth ePVSs.

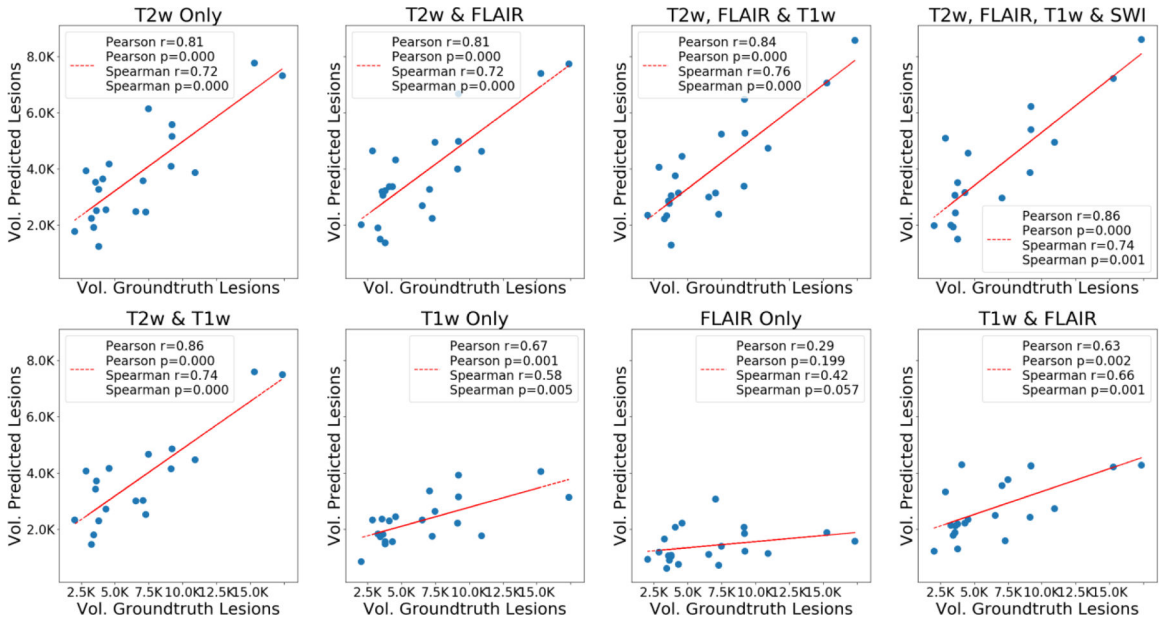

Fig. 6.

Scatterplots of volumes of groundtruth ePVS vs. volumes of predicted ePVS per subject, based on which the Person correlation and Spearman correlation (*r*) and the corresponding *p*-values (*p*) are calculated.

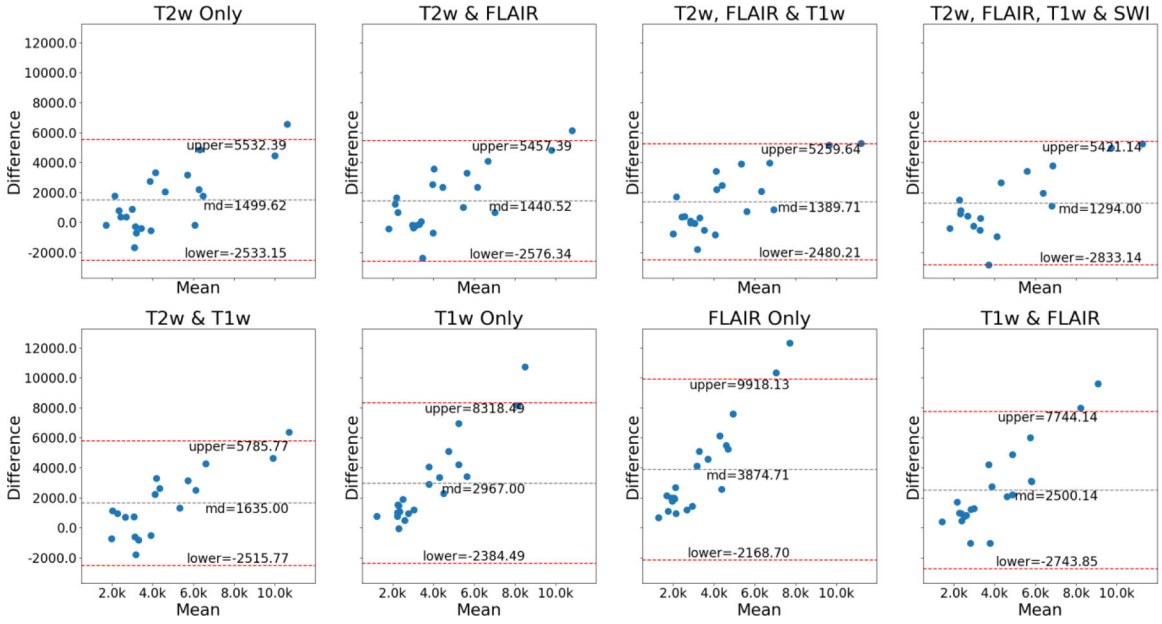

Fig. 7.

Bland-Altman plots of volume of of ePVS for different combinations of the sequences. The plots show the differences between the volumes of predicted ePVSs and those of the groundtruth ePVSs.

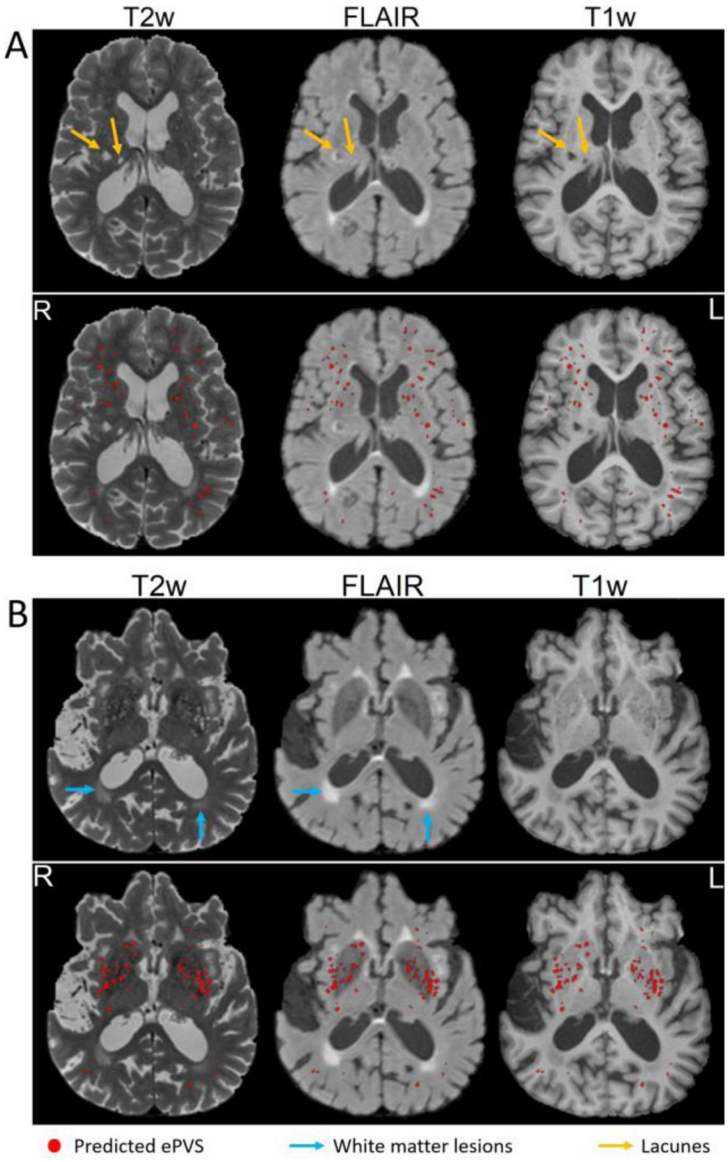

Fig. 8.

Examples of the predicted ePVS with the presence of white matter lesions and lacunes on T2w, FLAIR and T1w images. The proposed model is able to distinguish ePVS from such mimics.

Table 1 –

MRI Scanner parameters.

MRI Modalities	TR (ms)	TE (ms)	FOV (mm)	Flip Angle	Slice Thickness (mm)	No of Slices	Matrix	Scan Duration
T1w	1900	2.93	250	9	1	176	256×256	4:26
T2w	3200	408	250	120	1	176	256×256	4:08
FLAIR	6000 (TI=2200)	289	250	120	1	160	258×221	4:14
SWI	35	7.5, 15, 22.5, and 30	256	15	1.5	96	256×192	6:00

Table 2 -

Subject-wise evaluation for different combinations of the sequences. The best scores are marked as red and the second best as blue. The corresponding standard errors are shown when applicable.

Expts	Avg Sensitivity	Avg Precision	Avg Mag Accuracy	Avg Similarity Volumetric	Avg AUC	Average Hausdorff Distance	Average Mahanabolis Distance	ICC (#Lesions)	ICC (Volume)
T2w	0.81±0.01	0.83±0.02	1.16±0.02	0.81±0.03	0.72±0.01	1.41±0.08	0.17±0.02	0.83	0.59
T2w+FLAIR	0.82±0.01	0.82±0.02	1.16±0.01	0.81±0.03	0.73±0.01	1.38±0.08	0.16±0.02	0.77	0.60
T2w+FLAIR+T1w	0.82±0.01	0.83±0.02	1.17±0.01	0.82±0.03	0.74±0.01	1.27±0.07	0.17±0.02	0.70	0.63
T2w+FLAIR+T1w+SWI	0.82±0.02	0.83±0.02	1.17±0.02	0.82±0.03	0.74±0.01	1.28±0.07	0.16±0.01	0.77	0.67
T2w+T1w	0.80±0.02	0.84±0.02	1.16±0.02	0.78±0.02	0.71±0.01	1.40±0.09	0.18±0.01	0.79	0.58
T1w	0.63±0.01	0.77±0.02	1.00±0.02	0.66±0.04	0.59±0.01	2.49±0.11	0.24±0.02	0.30	0.18
FLAIR	0.47±0.02	0.73±0.02	0.88±0.02	0.48±0.04	0.53±0.02	3.59±0.09	0.35±0.04	0.24	0.05
T1w+FLAIR	0.68±0.02	0.75±0.02	1.02±0.02	0.71±0.03	0.61±0.01	2.30±0.11	0.23±0.02	0.50	0.25