Contents lists available at ScienceDirect

NeuroImage

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

Review

Computer-aided extraction of select MRI markers of cerebral small vessel disease: A systematic review

Jiyang Jiang [a,*], Dadong Wang [b], Yang Song [c], Perminder S. Sachdev [a,d], Wei Wen [a,d]

[a] Centre for Healthy Brain Ageing, Discipline of Psychiatry and Mental Health, School of Clinical Medicine, Faculty of Medicine, University of New South Wales, NSW 2052, Australia

[b] Quantitative Imaging Research Team, Data61, CSIRO, Marsfield, NSW 2122, Australia

[c] School of Computer Science and Engineering, University of New South Wales, NSW 2052, Australia

[d] Neuropsychiatric Institute, Prince of Wales Hospital, Randwick, NSW 2031, Australia

ABSTRACT

Cerebral small vessel disease (CSVD) is a major vascular contributor to cognitive impairment in ageing, including dementias. Imaging remains the most promising method for *in vivo* studies of CSVD. To replace the subjective and laborious visual rating approaches, emerging studies have applied state-of-the-art artificial intelligence to extract imaging biomarkers of CSVD from MRI scans. We aimed to summarise published computer-aided methods for the examination of three imaging biomarkers of CSVD, namely cerebral microbleeds (CMB), dilated perivascular spaces (PVS), and lacunes of presumed vascular origin. Seventy classical image processing, classical machine learning, and deep learning studies were identified. Transfer learning and weak supervision techniques have been applied to accommodate the limitations in the training data. While good performance metrics were achieved in local datasets, there have not been generalisable pipelines validated in different research and/or clinical cohorts. Future studies could consider pooling data from multiple sources to increase data size and diversity, and evaluating performance using both image processing metrics and associations with clinical measures.

1. Introduction

Cerebral small vessel disease (CSVD) is highly prevalent in older adults, and is associated with cognitive and neurobehavioural symptoms (Verdelho et al., 2021). Imaging biomarkers of CSVD include recent small subcortical infarcts (RSSI), white matter hyperintensities (WMH), cerebral microbleeds (CMB), dilated perivascular spaces (PVS), and lacunes of presumed vascular origin (hereafter, lacunes) (Wardlaw et al., 2013). Brain atrophy is also frequently observed in CSVD (Wardlaw et al., 2013), but the location and extent are usually comparable with those in ageing and neurodegenerative diseases, making it an unspecific biomarker. WMH are lesions in brain white matter that appear abnormally bright on T2-weighted MRI scans. As a CSVD biomarker with variable sizes and usually widely spread in the brain, WMH has been well documented for its automated segmentation methods (Balakrishnan et al., 2021; Caligiuri et al., 2015), risk factors (Brown et al., 2021), and clinical importance (Debette and Markus, 2010). RSSI are infarcts in the perfusion territory of arterioles with imaging features or clinical symptoms consistent with a recent lesion occurring in the previous few weeks (Wardlaw et al., 2013). They usually evolve into WMH or lacunes, or, more rarely, disappear

and leave normal-appearing brain tissue on MRI. RSSI appear hyperintense on T2-weighted fluid-attenuated inversion recovery (FLAIR) sequences, and hypointense on T1-weighted scans, similar to the contrast of WMH. Most WMH segmentation methods do not explicitly exclude RSSI. Diffusion-weighted imaging can aid the identification of RSSI. While studies have proposed methods to detect, segment and classify stroke lesions (Sarmento et al., 2020), as well as to separate them from WMH (Guerrero et al., 2018), there has not been any study aiming at explicitly segmenting RSSI.

CMB, PVS and lacunes are focal and small imaging biomarkers of CSVD. They share a few characteristics on MRI scans (Table 1): (1) focal lesions with darker or brighter intensities compared to surrounding tissue, and (2) round/ovoid shape (perpendicular plane for PVS). While labour-intensive, error-prone, and subjective visual ratings by experts have been applied to the majority of previous studies, more recent studies have proposed various algorithms to automate the extraction of CMB, PVS and lacunes. However, these methodological studies aimed at proof of concept with limited local data for validation. There have not been any generally accepted, automated pipelines ready for clinical studies of CMB, PVS or lacunes. Such pipelines are essential to advance our understanding on CSVD, especially in the era of big public datasets

* Corresponding author.
  E-mail address: jiyang.jiang@unsw.edu.au (J. Jiang).

Table 1
Definition of cerebral microbleed, perivascular space, and lacune of presumed vascular origin on MRI scans.

Lesion	Usual MRI modalities	Intensity	Shape	Size
Cerebral microbleed	T2*w GRE, SWI	Dark	Round/ovoid	Diameter = 2–5 mm (can be up to 10 mm). Blooming with longer TE.
Perivascular space	T1w, T2w, T2w-FLAIR	Dark on T1w and T2w-FLAIR. Bright on T2w.	Round/ovoid in the plane perpendicular to the course of vessel. Linear in the plane parallel to vessel.	Diameter ≤ 3 mm
Lacune of presumed vascular origin	T1w, T2w, T2w-FLAIR	Dark on T1w and T2w-FLAIR. Bright on T2w. Usually have bright rim on T2w-FLAIR.	Round/ovoid	Diameter = 3–15 mm

T2*w GRE, T2*-weighted gradient recalled echo sequence; SWI, susceptibility-weighted imaging; TE, echo time; T1w, T1-weighted imaging; T2w, T2-weighted imaging; T2w-FLAIR, T2-weighted fluid-attenuated inverse recovery sequence.
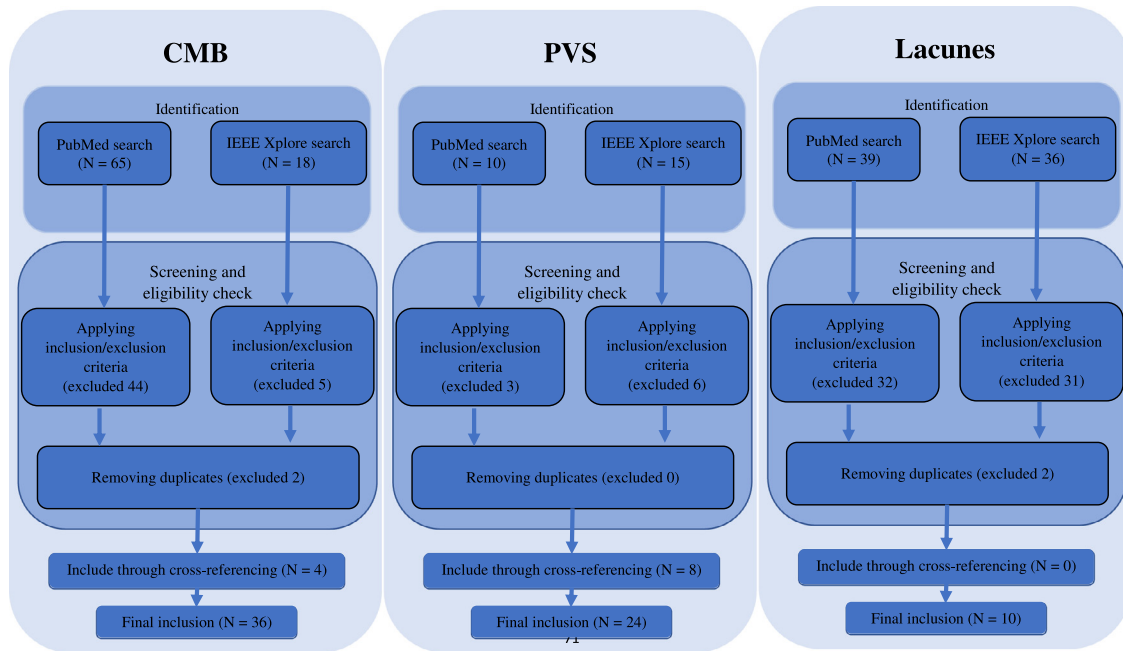

Fig. 1. Literature search process.

(Madan, 2021). Automatically quantified CSVD biomarkers can also be used to assess the effectiveness of clinical trials and provide objective decision-making tools (deSouza et al., 2019). The current work aimed at systematically reviewing these methodological studies, critically appraising the algorithms from the perspective of generalisability, highlighting promising techniques for future development, and discovering future steps to bridge the gap towards practical and generalisable applications in both research and clinical settings.

The current systematic review included all relevant publications without any constraints on the time when the studies were carried out. However, since there has been a previous systematic review on PVS segmentation (Hernandez Mdel et al., 2013), we focused on studies published after 2013 for PVS. In this published review, Hernandez Mdel et al. (2013) identified 8 publications (6 studies) assessing PVS using 7 computational methods, and 4 studies with automated techniques that can potentially be used for PVS segmentation. Studies have applied classical image processing approaches to consider size, shape, intensity, and location features. Thresholding was widely employed which would affect generalisability. Issues of inconsistent diameters and terminologies, and the significant overlap in shape, intensity, location, and size with lacunes, were identified in the review.

A few terms used in the current work need to be clarified. The current review included all automated and semi-automated studies addressing *regression, detection* and *segmentation* tasks. *Regression* tasks were defined as the prediction of an overall image-level measure, such as the rating or number of lesions. *Detection* tasks find lesions without explicitly out-

lining the boundary between the lesion and surrounding normal tissue. *Segmentation* was defined as a voxel-by-voxel depiction of the lesion. Applied techniques can be categorised into *classical image processing, classical machine learning,* and *deep learning. Classical image processing* techniques referred to traditional computer vision methods without a learning component. In the context of this work, these included various transformation-based methods and intensity-based segmentation techniques. *Classical machine learning* techniques mainly included random forest (RF) and support vector machine (SVM) classifiers in this review. *Deep learning* techniques referred to convolutional neural networks (CNN) and their variants. Some studies of PVS resolved tasks of *image enhancement* which, in the context of this review, means improving the image contrast of PVS, and making PVS more distinguishable from surrounding brain tissue on the image.

2. Methods

2.1. Identification of studies

Because of the technical nature of the review topic, two separate literature searches were conducted in PubMed and IEEE Xplore databases (November 2021). Specifically, the following keywords were used: *microbleed*, perivascular space*, Virchow-Robin space*, lacune*, lacunar infarct*, lacunar stroke, brain infarct, detect*, segment*, extract*, identif*, quantif* (Fig. 1). Reference lists of all relevant articles were also searched for additional studies.

2.2. Inclusion/exclusion criteria

For the systematic review, inclusion criteria were as follows:

- Methodological studies on the extraction of CMB, PVS and/or lacunes, defined in the Standards for Reporting Vascular Changes on Neuroimaging (STRIVE) consensus (Wardlaw et al., 2013).
- Studies using human MRI data for training, validation, and testing.
- Studies were required to apply computer-aided methods in a significant proportion of the algorithm.

Exclusion criteria were as follows:

- Development and/or comparison of imaging sequences (excluded 39 CMB studies).
- Studies using visual detection/rating methods (excluded 5 CMB and 8 PVS studies).
- Studies not published in English (excluded 5 CMB, 1 PVS, and 4 lacune studies).
- Studies on stroke lesion (excluded 59 lacune studies).

3. Results

3.1. Cerebral microbleeds

3.1.1. Overview

Thirty-six studies assessing CMB have been identified (Fig. 1). Table 2 summarised these studies. Many approaches could be divided into two stages: (1) *candidate detection* to screen for CMB candidates with high sensitivity, and (2) *false positive (FP) reduction* to remove false positives from the candidates. In the candidate detection stage, studies have used methods ranging from intensity thresholding (Barnes et al., 2011; Dou et al., 2015; Ghafaryasl et al., 2012), identifying circular/spherical objects (Bian et al., 2013; Chesebro et al., 2021; Fazlollahi et al., 2015, 2014; Morrison et al., 2018), to convolutional neural networks (Al-Masni et al., 2020; Dou et al., 2016). The FP reduction stage usually employs classical machine learning (Barnes et al., 2011; Chen et al., 2015; Dou et al., 2015, 2016; Fazlollahi et al., 2015, 2014; Ghafaryasl et al., 2012; Roy et al., 2015; van den Heuvel et al., 2015, 2016) and deep learning (Al-Masni et al., 2020; Liu et al., 2019) algorithms. Some deep learning (e.g., (Rashid et al., 2021)) and classical image processing (e.g., (Kuijf et al., 2013)) studies applied a single-pass method.

3.1.2. Materials

Study samples – The included studies had a median sample size of 27 participants. Due to the low prevalence of CMB in the general population, studies tended to use diseased patients who were more prone to CMB. These included patients with Cerebral Autosomal Dominant Arteriopathy with Sub-cortical Infarcts and Leukoencephalopathy (CADASIL) (Hong et al., 2019, 2020; Wang et al., 2017, 2019; Zhang et al., 2018a, 2018b, 2016b), stroke/transient ischemic attack (Dou et al., 2015, 2016; Seghier et al., 2011), traumatic brain injury (Liu et al., 2019; Roy et al., 2015; van den Heuvel et al., 2015), mild cognitive impairment and Alzheimer's Disease (Barnes et al., 2011; Fazlollahi et al., 2015, 2014), and gliomas with radiation therapy-induced CMB (Bian et al., 2013; Chen et al., 2019; Morrison et al., 2018).

MRI modalities – Hemosiderin deposits in CMB are paramagnetic, leading to heterogenic distribution of magnetic field and quick decay of magnetic resonance signal around CMB (i.e., susceptibility effects) (Shams et al., 2015). T2*-weighted gradient recalled echo (GRE) and susceptibility-weighted imaging (SWI) are the most commonly used imaging modalities for the extraction of CMB. The T2*-weighted GRE sequence is sensitive to susceptibility effects, and has been used to detect CMB for a long time (Greenberg et al., 2009). SWI sequence is introduced later as an alternative to T2*-weighted GRE, because it is heavily weighted by the intrinsic tissue magnetic susceptibility (Ruetten et al., 2019; Shams et al., 2015), resulting in greater reliability and sensitivity

in detecting CMB (Cheng et al., 2013; Shams et al., 2015). However, this is at the cost of increased conspicuity of CMB mimics, including calcium and iron deposits, arteries, and veins (Cheng et al., 2013; Shams et al., 2015). Moreover, due to the greater blooming effects of SWI on paramagnetic substances, CMB tend to be more irregular shaped on SWI (Chesebro et al., 2021; Mittal et al., 2009). Phase images can help separate diamagnetic substances (e.g., calcification) from paramagnetic CMB (Liu et al., 2019). Similarly, quantitative susceptibility mapping (QSM), a further advancement of the phase information to quantify tissue susceptibility (Ruetten et al., 2019), has also been used for the same purpose (Rashid et al., 2021).

MRI preprocessing – MRI preprocessing steps before CMB detection/segmentation were straightforward, mainly including bias field correction (Ghafaryasl et al., 2012; Rashid et al., 2021; van den Heuvel et al., 2016), brain extraction (Al-Masni et al., 2020; Ghafaryasl et al., 2012; Kuijf et al., 2013; Morrison et al., 2018; van den Heuvel et al., 2016), and intensity normalization (Al-Masni et al., 2020; Chen et al., 2015; Fazlollahi et al., 2015; Kuijf et al., 2013). Some studies also implemented data interpolation (Barnes et al., 2011; Chesebro et al., 2021).

Defining true CMB – The definition of gold-standard CMB in most studies is based on the Microbleed Anatomical Rating Scale (MARS) (Gregoire et al., 2009), where definite CMB are defined as small, rounded or circular, well-defined hypointense lesions within brain tissue with a diameter of 2–10 mm. Some computer-aided studies (Bian et al., 2013) also classified possible CMB which were less rounded or circular, less dark, or difficult to distinguish from mimics with confidence (Bian et al., 2013; Gregoire et al., 2009). With dual echo SWI, a missing blooming effect would indicate a non-CMB (Kuijf et al., 2011). Brain Observer MicroBleed Scale (BOMBS) was another scale used in some studies for defining true CMB (Rashid et al., 2021).

3.1.3. Algorithms

Classical image processing. Studies using classical image processing methods to extract CMB usually considered intensity and geometric information.

Intensity. Watershed is an image processing algorithm mainly used for segmentation. The intuitive idea comes from geography where catchment basins are filled up with water (Roerdink and Meijster, 2000). When the water level has reached the highest point where water coming from different basins would meet, dams are built. These dams are called watershed lines or watershed. After Contrast Limited Adaptive Histogram Equalization (CLAHE), Bottom-Hat filtering, and K-means clustering, watershed was applied with active contour segmentation to segment CMB (Tajudin et al., 2017a, 2017b).

Seghier et al. (2011) considered CMB as an extra tissue type in the unified normalisation-segmentation implemented in Statistical Parametric Mapping (SPM). Briefly, priors were adjusted to accommodate the intensity distribution in T2*-weighted images. The contrast between different tissue types was also standardised. These empirical priors were applied in the initial unified normalisation-segmentation. The results were used to modify the priors, including CMB probability thresholding, morphological granulometry thresholding, and removing skull and cerebrospinal fluid regions. The refined priors were then used in the second iteration of unified normalisation-segmentation. Liu et al. (2020b) considered both intensity at the voxel level, and gradients obtained by applying first-order horizontal and vertical differential operators to the images.

Geometry. The 'blooming' effect of CMB in all directions on SWI, influenced by the echo time, makes it a good fit for the radial symmetry transform (RST; Bian et al., 2013; Kuijf et al., 2012) (Fig. 2(A)). In RST, Sobel filters, either 2D (Bian et al., 2013; Morrison et al., 2018) or 3D (Kuijf et al., 2011, 2013, 2012), were applied to SWI data, resulting in positive values for dark-to-bright gradients. Since CMB are hypointense on SWI, only negatively affected voxels should be considered for gradients pointing towards the centre of CMB. Orientation and magnitude projection images were then calculated (Bian et al., 2013;

Table 2
Summary of CMB studies.

Study	method	sample type	Total N of ppts/CMBs	MRI modalities used for segmentation	Task	MRI preproc	Algorithm	Performance
Chen et al. (2015)	DL, FA	transient ischaemic attack	20/117	SWI	D	intensity normalisation	intensity thresholding + CNN + SVM	Patch-wise: Sen = 0.8913; Prec = 0.5616; avg FPs = 6.4
Zhang et al. (2016b)	DL, FA	CADASIL; healthy controls	10	SWI	D		DNN with sparse autoencoder	Patch-wise: Sen = 93.20+/−1.37%; Spec = 93.25+/−1.38%; Acc = 93.22+/−1.37%
Dou et al. (2016)	DL, FA	stroke; normal ageing	320/1149	SWI	D	intensity normalisation	3D FCN + CNN	Patch-wise: Sen = 93.16%; avg FPs = 2.74; Prec = 44.31%
Wang et al. (2017)	DL, FA	CADASIL	10	SWI	D		CNN with rank-based average pooling	Patch-wise: Sen = = 96.94%; Spec = 97.18%; Acc = 97.18%
Lu et al. (2017)	DL, FA		64	SWI	D		CNN	Patch-wise: Sen = 97.29%; Spec = 92.23%; Acc = 96.05%
Zhang et al. (2018a)	DL, FA	CADASIL; healthy controls	20	SWI	S		CNN	Voxel-wise: Sen = 93.05%; Spec = 93.06%; Acc = 93.06%
Zhang et al. (2018b)	DL, FA	CADASIL; healthy controls	20	SWI	S		DNN (one input layer, four sparse autoencoder layers, one softmax layer, and one output layer)	Voxel-wise: Sen = 95.13%; Spec = 93.33%; Acc = 94.23%
Hong et al. (2019)	DL, FA	CADASIL	10	SWI	D		CNN (ResNet50) with transfer learning	Patch-wise: Sen = 95.71+/−1.044%; Spec = 99.21+/−0.076%; Acc = 97.46+/−0.524%
Liu et al. (2019)	DL, FA	haemodialysis; tbi; stroke; healthy controls	220/1641	SWI	D	bias field correction; warp to standard space	3D FRST + ResNet (SWI and high-pass filtered phase image)	Patch-wise: Sen = 95.8%; Prec = 70.9%; avg FPs = 1.6
Wang et al. (2019)	DL, FA	CADASIL; healthy controls	20		D		DenseNet; transfer learning (freeze earlier layers and train later layers only)	Patch-wise: Sen = 97.78%; Spec = 97.64%; Acc = 97.71%; Prec = 97.65%
Chen et al. (2019)	DL, FA	individuals who had giomas, underwent radiation therapy, with confirmed radiation-induced CMBs	73	SWI	D		3D ResNet	Patch-wise: Prec = 71.9%; 94.7% of true CMBs were correctly identified
Hong et al. (2020)	DL, FA	CADASIL	10	SWI	D		CNN	Patch-wise: Sen = 99.74+/−0.13%; Spec = 96.89+/−0.26%; Acc = 98.32+/−0.15%
Al-Masni et al. (2020)	DL, FA		179/760	SWI	D	brain extraction; data augmentation	YOLO detection + 3D CNN	Patch-wise: avg FPs per subject = 1.42 (high resolution data); avg FPs per subject = 1.89 (low resolution data)
Rashid et al. (2021)	DL, FA	community	24 / 4 had no CMB; 13 had 1 or 2; 6 had 3–8; 1 had >100	T2w, SWI, QSM	S	bias field correction; coregistration; brain extraction; intensity normalisation; data augmentation.	U-Net with a higher number of resolution layers and padded convolutions.	Voxel-wise: Sen = 0.84–0.88; Prec = 0.40–0.59
Lu et al. (2021b)	DL, FA			SWI	D		CNN + extreme learning machine	Patch-wise: Sen = 94.53%; Spec = 96.10%; Acc = 95.25%
Myung et al. (2021)	DL, FA		186/1133	GRE	D	skull stripping; data augmentation	YOLO detection +/- CSF filtering	Patch-wise: Sen = 59.69–80.96%; Prec = 60.98–79.75%; avg FPs per subject = 2.15–6.57
Li et al. (2021)	DL, FA		58	SWI	D	bias field correction	CNN with transfer learning (SSD-512/VGG pre-trained on ImageNet dataset) + feature enhancement	Patch-wise: Sen = 90%; Prec = 79.7%

Table 2 (*continued*)

Study	method	sample type	Total N of ppts/CMBs	MRI modalities used for segmentation	Task	MRI preproc	Algorithm	Performance
Afzal et al. (2022)	DL, FA		20	SWI	D	skull stripping	CNN with transfer learning (AlexNet and ResNet50 pre-trained on ImageNet dataset)	Patch-wise: AlexNet: accuracy=97.26%, 1.8% FP rate; ResNet50: accuracy=97.89%, 1.1% FP rate
Barnes et al. (2011)	ML, SA	MCI; early AD	126 CMBs	SWI	D	brain extraction; interpolation	intensity thresholding + SVM (covariance matrix shape, intensity, and size features) + manual review	Patch-wise: Sen = 85.8%, Spec = 95.9% (before review); Sen = 81.7%, Spec = 100% (after review)
Ghafaryasl et al. (2012)	ML, FA		237/631	GRE, PD	D	bias field correction, coregistration, brain extraction	intensity and size thresholding + LDC/QDC/Parzen (size, shape, intensity features) + LDC/QDC/Parzen/linear SVM (Hessian-based shape, intensity on PD)	Patch-wise: Sen = 91%; avg FPs per subject = 4.1 FPs
Fazlollahi et al. (2014)	ML, FA	AD; MCI; healthy	41/104	SWI	D	tissue segmentation; brain extraction; bias field correction; intensity normalisation; inverting image contrast so that CMBs and vessels appear hyperintense; applying gradient based anisotropic diffusion and adaptive histogram equalisation	multi-scale Laplacian of Gaussian spherical/semi-spherical object detection + multi-layer RF (Radon-based shape features)	Patch-wise: Sen = 92.04%; avg 6.7 FPs per true CMB, and 16.84 false CMBs per subject
Fazlollahi et al. (2015)	ML, FA	AD; MCI; healthy	66/231	SWI	D	same as Fazlollahi et al. (2014)	multi-scale Laplacian of Gaussian spherical/semi-spherical object detection + multi-layer RF (Radon- and Hessian-based shape features)	Patch-wise: Sen = 87%; avg false detection rate of 27.1 CMBs per ppt (possible and definite CMBs); Sen = 93%; avg false detection rate of 10.0 CMBs per ppt (definite CMBs).
Roy et al. (2015)	ML, FA	mild to severe traumatic brain injury	27	SWI	S	SWI phase enhancement; brain extraction; bias field correction; registering T1w to SWI to obtain WM mask	RST + RF (intensity, RST shape features) + size thresholding	Cluster-wise: Sen = 85.6%; Spec = 99.5%; 1.73 FPs per true CMB.
Dou et al. (2015)	ML, SA		44/615	SWI	D	intensity normalisation	intensity and size thresholding + RF (intensity features) + 2-layer ISA for feature extraction + SVM	Patch-wise: Sen = 89.44%; avg FPs of 7.7 FPs per subject, and 0.9 FPs per CMB
van den Heuvel et al. (2015, 2016)	ML, SA	moderate to severe traumatic brain injury; healthy controls	33–51	T1w, SWI	D	tissue segmentation; coregistration; bias field correction; intensity normalisation	intensity thresholding + voxel-level RF (intensity,shape) + region growing + region-level RF (intensity, shape, size) + manual correction	Patch-wise: Sen = 89.1 ± 0.8% / 93.2 ± 1.0% with on average 25.9 ± 0.8 / 12.9 ± 0.8 FPs per TBI patient
Seghier et al. (2011)	IP, FA/SA	unselected, consecutive patients to Stroke Service	30	GRE	S		consider CMB as an extra class in unified normalisation-segmentation in SPM; a flat prior for CMB in the first pass, then modify subject-specific prior; 2nd pass with refined priors	Voxel-wise: Kappa = 0.43 (0.65 after manual correction); ICC = 0.71 (0.87 after manual correction)

5

Table 2 (*continued*)

Study	method	sample type	Total N of ppts/CMBs	MRI modalities used for segmentation	Task	MRI preproc	Algorithm	Performance
Kuijf et al. (2011)	IP, SA	individuals with CMB, no other brain pathology	2	dual-echo SWI	S	Minimal intensity projection post-processing for both echoes simultaneously; Processed images with a slab thickness of 2 mm (no overlap), and images with a slab thickness of 4 mm and a −2 mm gap between the slices (2 mm overlap) were used for both scans.	3D RST	Feasibility study; Human rating time was reduced from 30 to 1.5 min per scan using the proposed method.
Kuijf et al. (2012)	IP, SA	individuals with CMB	18/54	dual-echo SWI	S	brain masking containing GM and WM; intensity normalisation	3D RST	Cluster-wise: Sen = 71.2%. Required visual rating time reduced from 30 to 2 min.
Kuijf et al. (2013)	IP, SA	consecutive patients to memory clinic	72	GRE	S		3D RST	Cluster-wise: Sen = 65–84% depending on settings
Bian et al. (2013)	IP, SA	individuals who had giomas, underwent radiation therapy, with confirmed radiation-induced CMBs	15	SWI	S	brain extraction, intensity normalisation	2D FRST + intensity thresholding on FRST map + vessel mask screening (using FRST orientation projection map) + 3D region growing + 2D geometric feature examination (area, circularity)	Cluster-wise: Sen for definite CMBs = 95.4%; Sen for possible CMBs = 77.5%; Sen for total = 86.5%
Tajudin et al. (2017a, 2017b)	IP, FA				S		CLAHE to improve contrast + Bottom-Hat filtering + K-means + watershed + active contour segmentation	
Morrison et al. (2018)	IP, SA	individuals who had giomas, underwent radiation therapy, with confirmed radiation-induced CMBs	15/248	GRE	D/S	brain extraction, intensity normalisation	Bian et al. (2013) + volume segmentation using local intensity thresholding + 2D geometric feature examination (circularity, centroid-to-seed point distance) + size thresholding + manual judgement	Cluster-wise (detection): Sen = 86.7% (215/248 CMBs detected); Voxel-wise (segmentation): mean Jaccard Index = 0.98±0.01.
Liu et al. (2020b)	IP, FA		20		S		dual domain (gradient and voxel) distribution + distribution of Fourier coefficients	Voxel-wise: Sen = 85.2%; Prec = 3.2%
Chesebro et al. (2021)	IP, SA	community	78/64	SWI, GRE	S	brain extraction, coregistration, applying lobar masks, upsampling by factor of 3	Sobel filter + Canny edge detection + circular Hough transform + entropy of neighbourhood + Frangi filter-based analysis + manual correction	Cluster-wise: overall sen = 95% (combining SWI and GRE).

DL, deep learning; FA, fully automated; ML, machine learning; SA, semi-automated; IP, image processing; CADASIL, cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy; MCI, mild cognitive impairment; AD, Alzheimer's Disease; SWI, susceptibility-weighted imaging; QSM, quantitative susceptibility mapping; GRE, gradient echo sequence; PD, proton density-weighted imaging; S, segmentation; D, detection; CNN, convolutional neural network; SVM, support vector machine; FCN, fully convolutional network; RST, radial symmetry transform; FRST, fast RST; YOLO, You Only Look Once; CSF, cerebrospinal fluid; SSD-512/VGG, VGG-based Single Shot Multi-box Detector network; LDC, linear discriminant classifier; QDC, quadratic discriminant classifier; RF, random forest; ISA, independent subspace analysis; Sen, sensitivity; Spec, specificity; Acc, accuracy; Prec, precision.
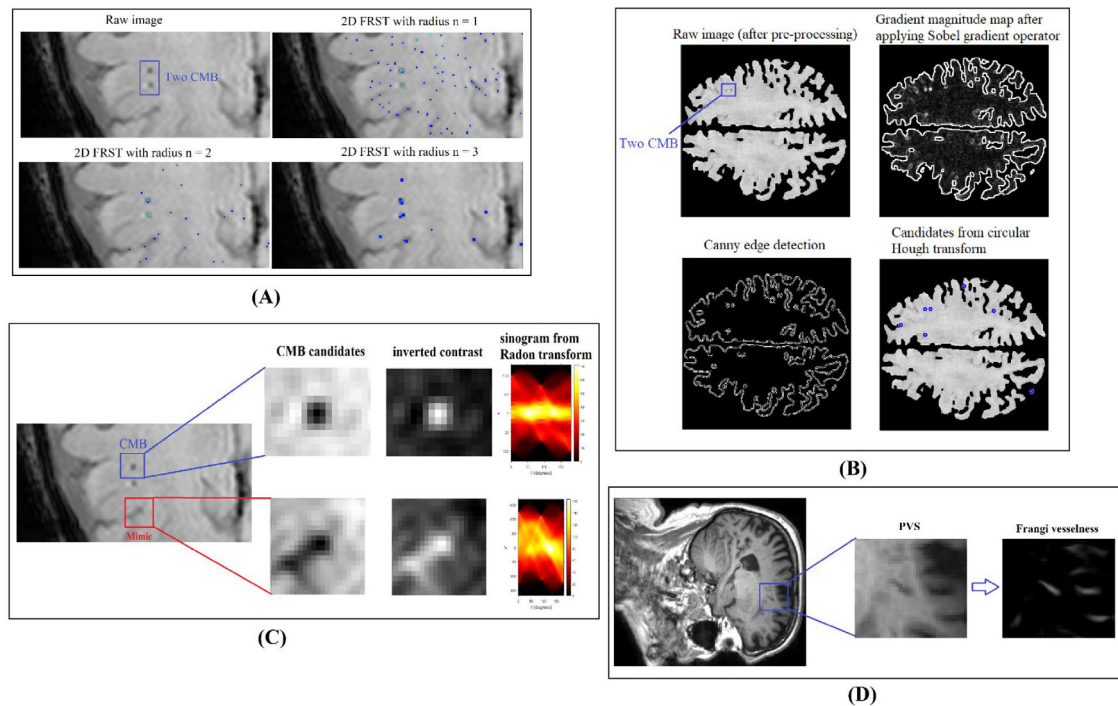
Fig. 2. Illustration of some classical image processing techniques. (A) 2D fast radial symmetry transform (RST); RST highlights points of interest in an image using local radial symmetry. The highlighted points depend on user-defined radiuses. The figure shows the candidate points detected using radiuses of 1–3. (B) Sobel gradient map, Canny edge detection, and circular Hough transform, following (Chesebro et al., 2021); Sobel operator aims to generate the gradient of an image. Canny edge detection operator is then applied to outline the edges from the gradient map. Circular Hough transform is an image transform to extract circular objects from an image, even if the circular is incomplete. (C) Radon transform; Radon transform can convert a volume on a image through summing voxel intensities along different orientations that are defined by a set of angles. (D) Frangi vesselness filter. Frangi vesselness filter is based on the Hessian matrix to compute the likeliness of a region containing vessels or tubular objects.

Kuijf et al., 2012), followed by the application of RST for each radius n ∈ N. The final transform was computed as the sum of all radial symmetry contributions. The output from RST was a voxel-by-voxel map with intensities indicating local sphericalness (Kuijf et al., 2013). The centre of CMB would have high intensities (i.e., high radial symmetry values) on the map.

Chesebro et al. (2021) applied Canny edge detection on the 2D gradient image (derived by applying Sobel filter to SWI or T2*-weighted GRE), followed by a circular Hough transform to detect circles (Fig. 2(B)). After excluding circles located on the edge of the brain, within cerebrospinal fluid, or with unrealistic sizes, the remaining circles were output as CMB candidates. Four metrics for each candidate CMB region were then used to clean FPs: 3D image entropy, 2D image entropy of the maximum intensity projection, and the volume and compactness of the central blob resulting from Frangi filter. Moreover, Liu et al. (Liu et al., 2020b) used Fourier descriptors to capture shape information.

Classical machine learning. SVM (Barnes et al., 2011; Dou et al., 2015) and RF (Dou et al., 2015; Fazlollahi et al., 2015, 2014; Roy et al., 2015; van den Heuvel et al., 2016) were the most commonly used classical machine learning algorithms in assessing CMB. Intensity, size and shape features were usually considered for the classification of CMB vs. non-CMB.

Intensity and size features. Intensity features considered in classical machine learning studies of CMB usually included average, minimum, maximum, median and/or standard deviation of intensities in the candidate (Barnes et al., 2011; van den Heuvel et al., 2016). Some studies also considered intensities of the voxels surrounding the candidate to take into account the contrast of CMB with surrounding voxels (Ghafaryasl et al., 2012). CMB usually have a diameter of 2–10 mm (Gregoire et al., 2009; Wardlaw et al., 2013). Therefore, size features,

including size of the candidate and the bounding box, were also considered (Ghafaryasl et al., 2012).

Geometric features. Shape is an important descriptor for detecting CMB and separating true CMB from mimics. Unlike the straightforward intensity and size features, geometric features usually require sophisticated transforms.

Fazlollahi et al. (2015, 2014) applied Radon transform to describe the shape of candidate objects and distinguish CMBs from vessel mimics (Fig. 2(C)). Specifically, the mean and standard deviation profiles across different projection angles were used as shape features. The length of mean and standard deviation profiles was standardized to account for different sizes of candidates.

The Hessian matrix and resultant measures have been used in both candidate detection and FP reduction stages. In the candidate detection stage, after applying multi-scale Laplacian of Gaussian filters, Fazlollahi et al. screened the resultant data for spherical objects (Fazlollahi et al., 2015, 2014). This was done through (1) transforming the question of detecting a 3D object into a question of detecting lines in orthogonal dimensions, (2) identifying the centre of the object by averaging the normalized line responses from the previous step and extracting local maxima, and (3) assessing the anisotropy of eigenvalues from the Hessian matrix of the identified centre for the detection of spherical objects. A lower anisotropy threshold was applied to account for the partial volume effects and presence of noise. The results were used as candidates for further reduction of FPs.

In the FP reduction stage, the three eigenvalues of the Hessian matrix and different synthetic measures from these eigenvalues have been proposed as shape descriptors:

- Ratio of eigenvalues (Ghafaryasl et al., 2012).
- Sphericalness defined as the volume to largest cross-sectional area (Fazlollahi et al., 2015, 2014).

- Largest cross section defined as the ratio of the largest 2 eigenvalues (Fazlollahi et al., 2015, 2014).
- Fractional anisotropy (Fazlollahi et al., 2015, 2014).
- Determinant of eigenvalues (Ghafaryasl et al., 2012; van den Heuvel et al., 2016).
- Trace defined as the sum of the diagonal in the Hessian matrix (Ghafaryasl et al., 2012; van den Heuvel et al., 2016).
- Frangi vesselness (Ghafaryasl et al., 2012; van den Heuvel et al., 2016).
- Local sphericalness defined as the equalness of 2 principal curvatures (van den Heuvel et al., 2015, 2016).
- To distinguish the distorted (e.g., elongated) CMB in SWI from vessels, a feature to describe the candidates' orientation, defined by the angle between the eigenvector with the largest eigenvalue and the xy-plane, was also employed (Fazlollahi et al., 2015, 2014).

A few other geometric descriptors have been proposed to describe CMB. van den Heuvel et al. (2015, 2016) used the convolution outcome of the scan with a spherical kernel as a shape feature. The kernel was designed to be negative in the middle, and positive on the edge, resulting in positive response when convolving with a dark spherical object. Since the kernel had been normalized, there is no response when convolving with an image with constant intensities.

In addition, as introduced in Section 3.1.3.1 'Classical image processing', RST outputs voxel-wise maps with intensities indicating local radial symmetry, and can therefore be used as shape features in classical machine learning approaches (Roy et al., 2015). The three eigenvalues of the covariance matrix and relative anisotropy calculated from these eigenvalues were considered as shape features in a previous study (Barnes et al., 2011). Moreover, the compactness, defined as the ratio between the volume of candidates and the volume of their bounding boxes, was used as a geometric feature (Ghafaryasl et al., 2012).

Deep learning. CNNs have been widely used in recent studies of CMB extraction (Table 2). Rashid et al. (2021) applied U-Net with a few modifications to adapt to characteristics in segmenting CMB. These modifications included 1) padded, instead of unpadded, convolutions, and 2) a higher number of resolution layers (6 as opposed to 5). 3D CNN was used in a few studies to take into account the contextual information in adjacent slices (Chen et al., 2019; Dou et al., 2016). Some recent studies also applied transfer learning techniques to address the issue of limited training data (Afzal et al., 2022; Hong et al., 2019; Li et al., 2021). Zhang et al. (2018b, 2016b) introduced sparse autoencoder layers to the CNN model for CMB detection.

CMB mimics. Vessel mimics are one of the most common FPs in detecting CMB. In addition to the shape features to distinguish vessels from CMB (e.g., Radon transform, vesselness using Hessian matrix), features from other imaging modalities have also been used. For example, Ghafaryasl et al. (2012) extracted average, minimum, maximum, median, and standard deviation of intensities within candidates on proton density (PD)-weighted images. Blood vessels are shown as hypointensity on PD-weighted scans, while CMB are not visible on these images. The phase information (either high-pass filtered phase images (Liu et al., 2019) or the calculated QSM images (Rashid et al., 2021)) was used to exclude diamagnetic calcification mimics.

Limited training data and class imbalance. Under-sampling (e.g., Zhang et al., 2016b) and augmentation (e.g., Afzal et al., 2022) were the most common approach to address the issues of class imbalance and small sample sizes. Dou et al. (2015) applied a stacked 2-layer convolutional independent subspace analysis (ISA) network to find discriminative features of CMB from limited training data in an unsupervised fashion, which were then used in a SVM classifier. Transfer learning with pre-trained models was also used in some recent deep learning studies to address the issue of limited training data. ResNet50 (Afzal et al., 2022;

Hong et al., 2019), AlexNet (Afzal et al., 2022), and VGG-based Single Shot Multi-box Detector (SSD)-512 network (Li et al., 2021) pre-trained on the ImageNet dataset, have been used to initialise the optimisation.

After the initial candidate detection, Fazlollahi et al. (2015) used consecutive independent RF classifiers with bagging and random undersampling. Low to high posterior probability thresholds were applied to the output from the classifiers to progressively exclude obvious non-CMB candidates, to handle the imbalanced dataset (i.e., more non-CMBs than non-CMBs) and decrease the number of non-informative candidates without distinctive features. These techniques have proven to be beneficial for effectively reducing FPs.

3.2. Perivascular spaces

3.2.1. Overview

Twenty-four studies assessing PVS have been identified (Fig. 1), and summarised in Table 3. Out of the 24 studies, 8 studies employed classical image processing methods (Ballerini et al., 2020, 2016, 2018; Boespflug et al., 2018; Liu et al., 2020a; Schwartz et al., 2019; Sepehrband et al., 2019; Wang et al., 2016), 5 applied classical machine learning techniques (SVM (González-Castro et al., 2016, 2017) and RF (Park et al., 2016; Zhang et al., 2016a, 2017)), 1 study combined classical image processing and machine learning (RF, (Hou et al., 2017)), and 10 used deep learning models (Boutinaud et al., 2021; Dubost et al., 2019a, 2020, 2017; Dubost et al., 2019b; Jung et al., 2019; Lian et al., 2018a, 2018b; Sudre et al., 2019; Yang et al., 2021). Due to the high prevalence of PVS in the general population and the large number of PVS within a scan, studies usually focus on certain brain regions, especially basal ganglia (Boutinaud et al., 2021; Dubost et al., 2019a, 2020, 2017; Dubost et al., 2019b; González-Castro et al., 2016, 2017; Wang et al., 2016; Yang et al., 2021) and centrum semiovale (Ballerini et al., 2020, 2016, 2018; Dubost et al., 2020, 2019b; Schwartz et al., 2019), because arterioles in these regions are relatively straight appearing as straight lines or perforating through the imaging plane (Wang et al., 2016). Midbrain, hippocampi, and deep corona radiata were also included in some studies (Ballerini et al., 2020; Dubost et al., 2019a, 2020).

There were also studies including wider regions, such as white matter (Boutinaud et al., 2021; Liu et al., 2020a) and the whole brain (Boespflug et al., 2018; Lian et al., 2018a, 2018b; Park et al., 2016; Sepehrband et al., 2019; Sudre et al., 2019). PVS studies have addressed regression (rating scores (Yang et al., 2021), number of PVS (Dubost et al., 2019a, 2019b)), detection (Dubost et al., 2020, 2017; Sudre et al., 2019), or segmentation (Ballerini et al., 2020, 2016, 2018; Boespflug et al., 2018; Boutinaud et al., 2021; González-Castro et al., 2016, 2017; Lian et al., 2018a, 2018b; Liu et al., 2020a; Park et al., 2016; Schwartz et al., 2019; Wang et al., 2016; Zhang et al., 2016a, 2017) tasks. A few studies addressed image enhancement tasks for PVS (Hou et al., 2017; Jung et al., 2019; Sepehrband et al., 2019).

3.2.2. Materials

Study samples – Studies have used population-based cohorts (Ballerini et al., 2020, 2016; Boespflug et al., 2018; Dubost et al., 2019a, 2017, 2019b; Schwartz et al., 2019), young healthy individuals (Boutinaud et al., 2021; Hou et al., 2017; Jung et al., 2019; Lian et al., 2018a, 2018b; Park et al., 2016; Sepehrband et al., 2019), and patients with subjective memory decline/mild cognitive decline/dementia (Ballerini et al., 2018; Schwartz et al., 2019; Yang et al., 2021), stroke (González-Castro et al., 2016, 2017; Wang et al., 2016), and idiopathic generalised epilepsy (Liu et al., 2020a). The median sample size is 65 (range, 1–2202).

MRI modalities – All studies used T1-weighted, T2-weighted, and/or T2-weighted FLAIR sequences to assess PVS, except for 3 studies that employed PD-weighted MRI with similar contrast to T2-weighted images (Boespflug et al., 2018; Dubost et al., 2019a, 2017). A few studies tried to enhance PVS contrast on images (Hou et al., 2017; Jung et al., 2019) and segment (Lian et al., 2018a, 2018b; Park et al., 2016;

Table 3
Summary of PVS studies.

Study	Method	Sample type	Total N of ppts	MRI modalities	Brain region/s	Task	MRI preproc	Algorithm	Performance
Wang et al. (2016)	IP, SA	mild stroke patients	100	T2w	BG	S	coregistration different modalities, brain extraction	adaptive intensity thresholding	BG PVS count associated with WMH and global atrophy
Ballerini et al. (2016)	IP, FA	community	24	T2w	CS	S	brain extraction, tissue segmentation, locating CS, reslicing	3D Frangi filter. Ordered Logit Model (OLM) to optimise Frangi filter parameters.	correlated with neuroradiological assessments
Hou et al. (2017)	IP + ML, FA	healthy volunteers (25–37 yo)	17	T2w	whole brain	E/S		Haar transform-based non-local line singularity representation	Voxel-wise: Dice = 0.75; Sen = 0.77; PPV = 0.73
Boespflug et al. (2018)	IP, FA	community-based non-demented	14 * 2 timepoints	T1w, FLAIR, T2w, PD	whole brain	S	tissue segmentation, bias field correction, registering other modalities to T1w	based on 1) relative normalised intensity in WM, ventricle and cortex on T1w, FLAIR, T2w, and PD data. 2) Morphologic (width, volume, linearity) characterization of each cluster.	correlations with manual counts =0.54–0.69
Ballerini et al. (2018)	IP, FA	dementia, stroke	68	T1w, T2w	CS	S	tissue segmentation, CSF+GM mask to reduce FPs, hole filling in WM, identifying CS ROI, reslicing.	same as Ballerini et al. (2016): 3D Frangi filter with OLM optimisaion.	PVS counts/volumes from automated method are correlated with visual rating.
Sepehrband et al. (2019)	IP, FA	HCP, healthy young	100	T1w, T2w	whole brain	E/S	correction for gradient nonlinearity, readout, bias field; AC-PC alignment; warp to MNI; tissue segmentation.	adaptive non-local mean filtering applied to T1w and T2w, enhanced contrast calculated as T1w/T2w, multi-scale Frangi vesselness filters	better PVS visibility, more PVS identified by rater, PVS count on enhanced and original image were correlated, high inter-rater and test-retest reliability using enhanced image, high correlation with visual counts = 0.72; test-retest (r = 0.87 (single slice), 0.77 (whole brain)); PPV = 77.5–87.5%
Schwartz et al. (2019)	IP, FA	community, ADNI	44	T1w, FLAIR	CS	S	tissue segmentation, bias field correction	local intensity inhomogeneity search on white matter-masked T1-weighted data, constrained by FLAIR hyperintensities, as well as width, volume, and linearity measurements	
Ballerini et al. (2020)	IP, FA	community (mean age, 72.6)	700	T2w	CS, deep corona radiata	S	generating individual CS mask through warping a CS mask in standard space, reslicing	same as Ballerini et al. (2016): 3D Frangi filter with OLM optimisaion.	PVS count/volume/length/width/size from automated method correlated with visual rating (r = 0.47–0.61); associated with hypertension, stroke, WMH.
Liu et al. (2020a)	IP, FA	children with idiopathic generalised epilepsy/controls	62	T2w	WM above ventricles	S	skull stripping, tissue segmentation	2D Frangi filter	
González-Castro et al. (2016)	ML, FA	stroke patients	264	T2w	BG	D	brain extraction, tissue segmentation, subcortical structure extraction, detecting BG ROI, CLAHE, num of bright blobs to determine best slice for PVS.	Bag of visual words (BoW) with Scale Invariant Feature Transform (SIFT).	Patch-wise: Acc = 82.34%; TNR = 80.77%; TPR = 83.95%.
Park et al. (2016)	ML, FA	healthy volunteers (25–37 yo)	17	T2w	whole brain	S	define candidates by anatomy and vesselness	randomised Haar features, sequential learning.	Voxel-wise: Dice = 0.63±0.05; sen = 0.59±0.08; PPV = 0.68±0.06
González-Castro et al. (2017)	ML, FA	stroke patients	264	T2w	BG	D	same as González-Castro et al. (2016)	compared 3 descriptors: 1) Wavelet transform coefficients, 2) local binary patterns, and 3) BoW with SIFT.	Patch-wise: Acc (SVM+BoW+SIFT) = 81.16% - best amongst 3; high agreement with visual ratings.
Zhang et al. (2016a, 2017)	ML, FA		19	T2w	whole brain	S		entropy-based sampling to extract informative samples, responses to 3 filters (steerable filters, optimally oriented flux, Frangi filter) as features.	Voxel-wise: Dice = 0.66; Sen = 0.65; PPV = 0.68

(continued on next page)

Table 3 (*continued*)

Study	Method	Sample type	Total N of ppts	MRI modalities	Brain region/s	Task	MRI preproc	Algorithm	Performance
Dubost et al. (2017)	DL, FA	community	1642	PD	BG	D	Warp to MNI, ROI masks from FreeSurfer, intensity normalisation	Compare to Dubost et al. (2020): deeper architecture, no residual connections.	Cluster-wise: Sen = 62%; avg 1.5 FPs per image
Lian et al. (2018a, 2018b)	DL, FA	healthy young to middle age	17–20	T2w	whole brain	S		multi-channel FCN; 2 inputs: original T2w, and T2w after denoising and enhencing for tubular structures (BM4D, Haar transform-based line sigularity representation); auto-context knowledge (recursive use of PVS probability maps from initial run).	Voxel-wise: Dice = 0.78±0.07 / 0.77±0.06; Sen = 0.74±0.12 / 0.74±0.13; PPV = 0.83±0.05 / 0.83±0.06
Sudre et al. (2019)	DL, FA		16	T1w, T2w, FLAIR	whole brain	D/S	bias field correction, skull stripping, z-tranform to WM region statistics	2D RCNN for multi-class multi-instance simultaneous detection and segmentation of PVS and lacunes.	Cluster-wise: Sen = 72.7%; median overlap of 59% between detected bounding box and ground truth when all raters agreed, and an overlap of 30% for more uncertain cases (>= 1 rater considering non-lesion).
Dubost et al. (2019b)	DL, FA	community	2115	T2w	midbrain, hippocampi, BG, CS	R/D	Bias field correction, ROI mask from FreeSurfer, individual mask adjustment, intensity normalisation.	Compare to Dubost et al. (2019a): a) simpler and lighter, b) including skip connections, and c) global pooling instead of 2 fully connected layers.	ICC against visual rating = 0.75–0.88; ICC of test-retest = 0.82–0.93; similar associations with 20 PVS determinants to visual scores.
Dubost et al. (2019a)	DL, FA	community	2017	PD	BG	R/D	Warp to MNI, ROI masks from FreeSurfer, intensity normalisation	regression CNN to predict num of PVS, with more convolutional than pooling layers and no final activation.	ICC against visual rating = 0.74; ICC of test-retest = 0.93; correlated with age similar to visual.
Jung et al. (2019)	DL, FA	healthy volunteers (25–37 yo)	17	T2w	whole brain	E		3D CNN with dense skip connections to alleviate the gradient vanishing problem.	peak SNR = 47.98±2.66 (PVS), 44.59±2.23 (WM); structural similarity with gold standard = 0.975±0.01 (PVS), 0.98±0.01 (WM).
Dubost et al. (2020)	DL, FA		2202	T2w	midbrain, hippocampi, BG, CS	D	FreeSurfer segmentation of ROI, intensity normalisation	Train a regression CNN with total num of PVS. During inference, attention maps of PVS were calculated as linear combination of feature maps of the layer before global pooling layer using weights from the following fully connected layer.	Cluster-wise: Sen = 62.1 ± 8.7%; avg 2.33±1.71 FPs per scan; avg 2.44±2.01 FNs per scan
Yang et al. (2021)	DL, FA	young healthy/ dementia/MCI/ subjective memory impairment	96	T2w	BG	R	intensity normalisation, identify BG slices, Haar to enhance visibility.	CNN	Accuracy = 87.7% (image-level), 80.3% (subject-level)
Boutinaud et al. (2021)	DL, FA	university students	1832	T1w	BG, deep WM	S	tissue segmentation, skull stripping, intensity normalisation, generating BG mask	3D convolutional autoencoder trained on visual rating data, transferring learnt weights to U-Net	Dice (voxel-level) = 0.51 (DWM), 0.66 (BG); Dice (cluster-level) = 0.64 (DWM), 0.71 (BG); Dice > 0.9 for large PVS; R^2 > 0.38 for DWM, and 0.02 for BG against visual rating.

FA, fully automated methods; SA, semi-automated methods; DL, deep learning; ML, machine learning; IP, image processing; HCP, human connectome project; ADNI, Alzheimer's Disease Neuroimaging Initiative; BG, basal ganglia; CS, centrum semiovale; S, segmentation; E, image enhancement; D, detection; R, regression; CSF, cerebrospinal fluid; GM, grey matter; WM, white matter; FP, false positive; ROI, region of interest; CNN, convolutional neural network; ICC, intraclass correlation coefficient; TNR, true negative rate; TPR, true positive rate.

10

Zhang et al., 2016a, 2017) PVS from scans acquired from 7T MRI scanners.

MRI preprocessing – Since a few computer-aided methods for extracting PVS focused on particular brain regions (e.g., basal ganglia, centrum semiovale), these masks were generated in addition to the preprocessing steps commonly used in studies of CMB and lacunes (within-subject coregistration of image modalities, bias field correction, brain extraction, intensity normalisation). Subcortical grey matter structures segmented in FreeSurfer had been used to derive masks of the basal ganglia which included thalamus, thalamus-proper, caudate, putamen, pallidum, and accumbens (Boutinaud et al., 2021). Other methods had detected landmarks in the insular cortex to construct a polygon around the basal ganglia region (González-Castro et al., 2016, 2017). Cerebrospinal fluid regions were then removed from the polygon to generate a mask of the basal ganglia. Centrum semiovale was usually defined as white matter regions above the lateral ventricles (Ballerini et al., 2016). Some studies registered a centrum semiovale mask in standard space to individual space (Ballerini et al., 2020). Others used FreeSurfer segmentations (Dubost et al., 2019b), or first detected lateral ventricles and then defined white matter above them as centrum semiovale (Ballerini et al., 2018).

Defining true PVS – While studies applied various protocols to the manual annotations and visual ratings of PVS (Ballerini et al., 2020; Boespflug et al., 2018; Boutinaud et al., 2021; Dubost et al., 2020), PVS are usually defined as round, ovoid, or linear lesions (diameter < 3 mm) with intensity characteristics similar to cerebrospinal fluid. The pattern of PVS in different brain regions may show slight differences with hippocampal PVS being rounder, and PVS in centrum semiovale being more elongated (Dubost et al., 2019b). Since PVS follow the trajectory of a vessel, the PVS shape on a scan slice depends on the imaging plane (Wardlaw et al., 2013). Some classical machine learning studies applied semi-automated methods followed by manual correction to accelerate the annotation of true PVS (Park et al., 2016; Zhang et al., 2016a, 2017). Unlike labelling PVS on MRI scans, it is challenging to create an image with improved PVS visibility, reduced noise, and more distinguishable contrasts from the surrounding brain tissue, to serve as the gold standard in supervised image enhancement tasks. A deep learning study (Jung et al., 2019) used established classical image processing methods (Hou et al., 2017) to generate gold standard data and validate its proposed model.

3.2.3. Algorithms
Classical image processing.

Intensity. Wang et al. (2016) applied linear intensity adjustment to a basal ganglia slice on the T2-weighted scan. The voxel-by-voxel product of the original T2 scan with this intensity adjusted T2 image was then calculated. The basal ganglia region was then segmented. After interactive adjustment of minimum and maximum sizes of hyperintense objects, one of three intensity thresholds was applied. The selection of threshold was based on the severity of PVS, background intensity, and other visible lesions.

In another study, voxel-level intensity inhomogeneity was used to segment PVS candidates (Schwartz et al., 2019). Specifically, the median intensity of surrounding voxels within a search filed of radius of 3.5 mm was defined as the median score for the voxel i. The mean intensity differences between the voxel i and its neighbouring voxels within a search field of radius of 5.5 mm was defined as the difference score for voxel i. PVS voxels were then defined as those within the white matter mask, with an intensity no more than 90% of the median score, and a difference score larger than 5% of their intensities. The thresholding was done on T1-weighted images, taking into account T2-weighted FLAIR data to avoid WMH regions. Geometric and size constraints were then applied to reduce FPs.

Geometry. Provided that the pattern of PVS follows the course of blood vessels, filters sensitive to vessel-like tubular structures will be helpful in enhancing the presence of PVS in images (Fig. 2(D)). Liu et al.

(2020a) and a body of work by Ballerini et al. (2020, 2016, 2018) have applied 2D or 3D Frangi filter to detect vessel-like objects (Frangi et al., 1998). The Frangi vesselness is based on the eigenvalues of the Hessian matrix consisting of second-order derivatives of Gaussian filter responses. It describes the similarity of one structure to an ideal tube. Ballerini et al. (2020, 2016, 2018) optimised the parameters in the Frangi vesselness function by applying an ordered logit model which took visual rating results into account. Given the various sizes of PVS, multi-scale vesselness was often calculated (Ballerini et al., 2020, 2016, 2018; Sepehrband et al., 2019).

Boespflug et al. (2018) used the linearity and width of candidate clusters to remove FPs. Specifically, a line representing the eigenvector associated with the largest eigenvalue of the candidate cluster, was derived by using singular value decomposition. Norms were then fit from each voxel in the candidate cluster to the line. The correlation between the distance from each voxel to the cluster centre, and the distance from the point on the line defined by the voxel's norm to the cluster centre, was considered as a measure of the cluster linearity. The width of cluster was described as the summed distance of the two farthest points from the line and 1.7 (the corner-to-corner distance of a 1 mm unit cube). A vessel-like, linear object should have high linearity (i.e., high correlation coefficients) and relatively small width. Similar approach was also applied to another study (Schwartz et al., 2019) to describe linearity, width and length of candidate clusters.

Classical machine learning.

Intensity features. Bag of words (BoW) is an approach to represent data, usually used in natural language processing. González-Castro et al. (2016) applied BoW to the task of PVS segmentation. Image patches around key points extracted from the basal ganglia region were first created through dense sampling. An intensity distribution descriptor was calculated to characterise the image patch. The k prototype vectors from the k-means clustering were considered as 'visual words', and a dictionary was created with these k visual words. All images in the dataset were used to build the dictionary of visual words. For each patch of an image in the dataset, the intensity distribution descriptor was calculated (patch-level descriptor). A visual word was considered as the representation of the patch when it had the shortest Euclidean distance from the patch-level descriptor. An image-level descriptor was obtained through considering the histogram of visual words representative of all patches. An SVM model was built with these image-level descriptors/features to predict image-level PVS visual rating scores (70% of all images as training set, and 30% as test set).

Geometric and orientation features. In addition to the Frangi vesselness, similar to studies extracting CMB, other shape and orientation features have been derived from the Hessian matrix. These included the features derived from all 3 eigenvalues ($\frac{|\lambda_1|}{|\lambda_2|}$, $\frac{|\lambda_1|}{\sqrt{|\lambda_2 \lambda_3|}}$, $\sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}$, where $\lambda_1 > \lambda_2 > \lambda_3$), and those resulting from the eigenvector (x_1, y_1, z_1) corresponding to the largest eigenvalue λ_1 ($\arccos(\frac{|z_1|}{\sqrt{x_1^2 + y_1^2 + z_1^2}})$, $\arctan(\frac{y_1}{x_1})$) (Zhang et al., 2016a, 2017).

In addition to the Hessian matrix-based features, Zhang et al. (2016a, 2017) applied 2 other filters to describe the orientation PVS, namely the steerable filter and optimally oriented flux. The steerable filters are based on the fact that any arbitrarily oriented responses to the first- and second-order Gaussian derivative filters can be derived by linear combinations of some basis filter responses. In these two studies (Zhang et al., 2016a, 2017), responses to 19 such basis filters (1 Gaussian filter, 9 oriented first-order Gaussian derivative filters, 9 oriented second-order Gaussian derivative filters) were used as features for the following RF classifier. Optimally oriented flux emphasises curvilinear structures by showing the amount of image gradient flowing in or out of a local sphere. Two sets of features were used: (1) features based on the first 2 eigenvalues of the 3 × 3 optimally oriented flux matrix (λ_1, λ_2, $\sqrt{\max(0, \lambda_1 \lambda_2)}$, $\lambda_1 + \lambda_2$), and (2) features derived from the

eigenvector corresponding to the largest eigenvalue (similar to the features extracted from Hessian matrix mentioned above).

Park et al. (2016) first identified candidates based on anatomical information and vesselness. Instead of specifying discriminative features, 3D Haar features were calculated in randomly defined image patches. Since the majority of informative features were located in the centre line, more features were extracted from the main direction of the vessel-like tube.

Deep learning. Lian et al. (2018a, 2018b) employed multi-channel inputs, including the original T2-weighted scan and the T2-weighted images enhanced for thin tubular structures, to provide both the information of regional intensities and the emphasis of tubular objects. In addition, the output from the initial run of the model was considered to provide contextual guidance for locating PVS, and therefore, recursively used as an additional input channel to refine the performance of the fully convolutional network. Sudre et al. (2019) used T1- and T2-weighted and T2-weighted FLAIR images as input, to detect and segment PVS and lacunes. This study considered the inconsistency between multiple raters, and used, instead of a hard classification, a soft probability label averaging among all raters as the gold standard.

Weakly supervised learning. Since PVS are small lesions and highly prevalent within human brains, visual rating is more practical compared to manual annotation when generating the ground truth for PVS. Manual annotation of PVS was often only available in small samples. Therefore, a few studies had proposed strategies to make use of weakly labelled data (i.e., visual ratings) in designing deep learning models for PVS segmentation. Boutinaud et al. (2021) trained a convolutional autoencoder with visual rating data. The autoencoder had the same architecture as a U-Net model, which was used for the final segmentation, except for the absence of skip connections. Weights learnt from the autoencoder were transferred to the U-Net model for PVS segmentation. Dubost et al. (2020, 2017) proposed a class activation mapping approach. During training a U-Net-like model to predict an image-level label (i.e., the number of PVS), a global pooling layer was added after the last convolutional layer and before a fully connected layer. In order to output the attention map, the global pooling layer was removed. The attention map was calculated as the linear combination of the feature maps (i.e., activation maps) derived from the last convolution layer (i.e., the layer before global pooling) using the weights of the following fully connected layer. Higher values in the attention map were more likely to be PVS.

Image enhancement for PVS segmentation. Studies have amplified the Haar transform coefficients corresponding to the desired signal (i.e., signal in PVS regions) to increase PVS visibility (Yang et al., 2021). However, this approach may amplify both signal and background noise when signal is not sufficiently discriminative. To solve such issue, Hou et al. (2017) proposed a non-local strategy where the Haar transform was applied to a stack of neighbouring cubes of the reference cube. A nonlinear mapping function was then applied to amplify the signal and suppress the noise. This was followed by inverse Haar transform to generate enhanced neighbouring cubes. Enhanced neighbouring cubes at the same location corresponding to different reference cubes were finally averaged to form the entire enhanced MRI image. The residual noise was further reduced by applying a block-matching and 4D filtering algorithm which is also based on the non-local strategy. Jung et al. (2019) proposed a deep 3D CNN to address the image enhancement task, and used the enhanced images according to Hou's approach (Hou et al., 2017) as ground truth. The proposed architecture included densely connected blocks with skip connections to reduce the number of parameters and mitigate the vanishing gradient problem in deep layers.

Sepehrband et al. (2019) applied an adaptive non-local mean filtering algorithm taking into account the surrounding voxels weighted by the Euclidean distance. The algorithm included a regularisation term to account for the Rician noise in MRI data. High-frequency noise was removed by applying a filtering patch with a radius of 1 voxel. The algorithm was applied to T1- and T2-weighted images separately, and the final enhanced PVS contrast was calculated as the enhanced T1-weighted image divided by the enhanced T2-weighted data.

3.3. Lacunes

3.3.1. Overview

Ten studies examining lacunes have been identified and summarised in Table 4. We identified 2 published studies applying deep learning algorithms (Al-Masni et al., 2021; Ghafoorian et al., 2017), 2 using image processing techniques (Wang et al., 2012; Yokoyama et al., 2007), and a body of 6 studies by Uchiyama et al. combining image processing with machine learning methods (Uchiyama et al., 2015, 2009, 2012; Uchiyama et al., 2008, 2007a, b), in the extraction of lacunes and separate lacunes from other imaging biomarkers of CSVD.

3.3.2. Materials

Study samples – While some studies did not specify details of included samples, others have used community-based cohorts (Wang et al., 2012), as well as patients with CSVD, transient ischaemic attack, and/or ischaemic/haemorrhagic stroke (Ghafoorian et al., 2017). The median sample size in studies of lacunes was 132 (range, 100–1075), but due to the low prevalence of lacunes in the general population, the number of lacunes is generally low in studies (usually < 100).

MRI modalities and preprocessing – Lacunes appear hypointense on T1-weighted and T2-weighed FLAIR scans, and hyperintense on T2-weighted scans (Wardlaw et al., 2013). Therefore, at least 2 out of these 3 MRI modalities have been used in all included studies. Uchiyama et al. (2009) used magnetic resonance angiography (MRA) data in addition to T1-weighted and T2-weighted FLAIR scans to remove PVS from lacune candidates. MRI preprocessing usually included co-registration of different MRI modalities (Al-Masni et al., 2021; Ghafoorian et al., 2017; Wang et al., 2012), brain extraction (Ghafoorian et al., 2017; Wang et al., 2012), bias field correction (Ghafoorian et al., 2017; Wang et al., 2012), and intensity normalisation (Al-Masni et al., 2021; Ghafoorian et al., 2017).

Defining true lacunes – The ground truth of lacunes was typically labelled by experienced neuroradiologists and neurologists (Al-Masni et al., 2021; Uchiyama et al., 2007a). Ghafoorian et al. (2017) applied the STRIVE protocol, where lacunes were defined as round or ovoid, subcortical, fluid-filled cavities with 3–15 mm diameters (Wardlaw et al., 2013).

3.3.3. Algorithms

Classical image processing and machine learning.

Intensity features. Wang et al. (2012) applied intensity thresholding methods to segment lacunes adjacent to WMH and those located in subcortical structures. During candidate detection, since lacunes showed varying intensities in different stages (acute, subacute, chronic), multiphase thresholding was used in a few studies to remove background (Uchiyama et al., 2007a; Yokoyama et al., 2007). On T2-weighted scans, it is difficult to segment lacunes adjacent to hyperintense regions (e.g., lateral ventricles) because of similar intensities. To address this issue, studies have also applied white top-hat transform to enhance lacunes with a user-defined structuring element (e.g., a square element (Uchiyama et al., 2007a), or circular elements with different radii (Yokoyama et al., 2007)).

Geometric and location features. The largest and smallest eigenvalues of the Hessian matrix were used to describe nodular and linear objects (Uchiyama et al., 2015, 2012, 2007a). Nodular objects should have negative values for the second derivatives of Hessian eigenvalues in all directions. The second derivative of Hessian eigenvalues for linear objects should be close to zero in the direction along the long axis, and negative in the direction perpendicular to the long axis. These geometric features were calculated for multiple scales. The x- and y-coordinates

Table 4
Summary of lacune studies.

Study	Method	Sample type	Total N of ppts/lacunes	MRI modalities	Task	MRI preproc	Algorithm	Performance
Yokoyama et al. (2007)	IP, SA		100	T1w, T2w	D		multiple-phase binarization for isolated lacunes; white top-hat transform with circular structuring elements for lacunes adjacent for hyperintense structures; FP reduction using features on T1w	Patch-wise: Sen = 90.1%; Spec = 30.0%; 1.7 FPs per image
Uchiyama et al. (2007b)	IP + ML, SA		132	T1w, T2w	D		same as Uchiyama et al. (2007a), but replacing SVM with a 3-layer neural network followed by modular classifiers to reduce 3 types of FPs	Patch-wise: Sen = 96.80%; 0.30 FP per slice
Uchiyama et al. (2012, 2007a)	IP + ML, SA		132	T1w, T2w	D	brain extraction with region growing	candidate identificatin using white top-hat transform and multi-phase binarisation; FP reduction considering location, intensity and shape.	Patch-wise: Sen = 96.80%; 0.76 FP per slice
Uchiyama et al. (2009, 2008)	IP + ML, SA		109/89	T1w, T2w, (MRA in 2009 study)	D		white top-hat transform; intensity thresholding on transformed images; 3-layer neural network to separate lacunes from PVS (+ MRA to identify and separate PVS from lacunes)	Patch-wise: Sen = 93.3% (83/89); Spec = 75.0% (15/20)
Wang et al. (2012)	IP, SA	community	272/62	T1w, T2w, FLAIR	S	bias field correction, coregistration, brain extraction, tissue segmentation, WMH extraction	lacunes near WMH: dilation + intensity contrasts; lacunes in subcortical structures: intensity thresholding.	Cluster-wise: Sen = 80.6% (detected 50/62 lacunes)
Uchiyama et al. (2015)	IP + ML, SA		132	T1w, T2w	D		same as Uchiyama et al. (2007a). Add template matching in eigenspace for FP reductions	Patch-wise: Sen = 96.80%; 0.47 FP per slice
Ghafoorian et al. (2017)	DL, FA	non-demented older individuals, young stroke patients	1075	T1w, FLAIR	D	coregistration, warp to standard space, bias field correction, intensity normalisation	FCN + 3D CNN (incld. multi-scale contextual info, and 7 explicit location features)	Patch-wise: Sen = 97.4%; 0.13 FPs per slice
Al-Masni et al. (2021)	DL, SA		288/696	T1w, FLAIR	D	coregistration, intensity normalisation	manual identification of candidates, 3D ResNet	Patch-wise: Sen = 96.41%; Spec = 90.92%; Acc = 93.67%; Prec = 91.40%; avg FPs per ppt = 1.32
Sudre et al. (2019)	Sudre et al. (2019) segmented both lacunes and PVS. Refer to Table 3 for more details.							

SA, semi-automated methods; FA, fully automated methods; DL, deep learning; ML, machine learning; IP, image processing; S, segmentation, WMH, white matter hyperintensity; FP, false positive; SVM, support vector machine; MRA, magnetic resonance angiography; PVS, perivascular spaces; FCN, fully convolutional network; CNN, convolutional neural network; Sen, sensitivity; Spec, specificity; Acc, accuracy; Prec, precision.

were used to describe the location of lacune candidates (Uchiyama et al., 2015, 2012, 2007a). Template matching in eigenspace was also used to further reduce FPs (Uchiyama et al., 2015).

Deep learning. Two recent studies have applied deep learning techniques. A fully convolutional network was used in the candidate detection phase to reduce the large amount of time consumed in the sliding window approach (Ghafoorian et al., 2017). For FP reduction, a 3D CNN was employed with multi-scale contextual information and hard-coded location features. Another study applied 3D multi-scale residual convolutional network with T1-weighted and T2-weighted FLAIR data as inputs (Al-Masni et al., 2021).

Separating lacunes from PVS. Lacunes and PVS are both hypointense on T1-weighted and T2-weighted FLAIR scans, and hyperintense on T2-weighted images (Wardlaw et al., 2013). Therefore, a few efforts have been made to separate these two types of lesions. Uchiyama et al. (2009, 2008) applied a 3-layer neural network with 6 features: (1) candidate locations described as x and y coordinates, (2) size of candidates, (3) degree of shape irregularity calculated as $1 - C/L$, where C is the length of circumference of the circle having the same area as the candidate, and L is the boundary length of the candidate, and (4) the intensity contrast between the candidate and its surrounding areas. Moreover, cerebral blood vessels were reconstructed from MRA, and superimposed onto T2-weighted scans, to aid the identification of PVS (Uchiyama et al., 2009).

4. Discussion

In a systematic review, we summarised 70 studies on computer-aided methods for extracting three imaging biomarkers of CSVD (CMB, PVS, lacunes). There is a trend of favouring artificial intelligence techniques (classical machine learning, and more recently, deep learning) in examining these biomarkers. In the following sections, we aimed to discuss promising techniques and challenges, and, wherever possible, provide recommendations.

4.1. Weak supervision and transfer learning

Since CMB, PVS and lacunes are small lesions, manually labelling them on MRI scans can be demanding. In particular, although PVS are highly prevalent even in children and healthy young adults (Groeschel et al., 2006), they tend to be widely distributed in the brain with significantly varying numbers and locations between individuals, making manual labelling all PVS almost impossible. As a result, studies had focused on certain highly vascularised areas (e.g., basal ganglia and centrum semiovale), and considered the PVS in these regions to be a valid representation of the severity of PVS in the whole brain. Despite being restricted to specific areas, many studies still only had access to visual rating data (e.g., the number of PVS, and categories of PVS severity according to the number). This raised the issue of weakly labelled data for PVS.

On the other hand, CMB and lacunes are less common to a significant degree. In the general population, CMB are present in 7–13% of individuals aged 40–69 years (Lu et al., 2021a; Poels et al., 2010). Individuals over 80 years of age (35.7% prevalence Poels et al., 2010) and stroke patients (29.4%, Ibrahim et al., 2019) had a higher prevalence of CMB. About 70–80% of individuals who have CMB have a single CMB (Lu et al., 2021a; Poels et al., 2010). Lacunes were reported in 8–31% of the healthy older individuals (Das et al., 2019). Similar to CMB, most individuals (66% of those with lacunes) had a single lacune. The low prevalence of CMB and lacunes has limited the size of data with labelled ground truth.

A few novel approaches have been applied to accommodate the weakly labelled data and limited data sizes. In image processing tasks using CNN, the early layers typically learn simple features such as edges

and contrasts. Additional layers then follow to learn more abstract features, with the last layers trying to map from learnt features to the task-dependent output (e.g., number of lesions). Since the features learnt in early layers are not expected to be specific to the task, it is possible to learn these features from other large datasets with similar semantics. The parameters in the early layers can then be frozen, and later layers can be adjusted to address the particular task using the gold-standard labels. This representation learning technique is usually referred to as transfer learning. It significantly reduces the required sample sizes for training a CNN model. This has been applied in some studies in this review to address the issue of small training sample sizes (Afzal et al., 2022; Hong et al., 2019; Li et al., 2021). Some of these studies used models pre-trained on the ImageNet dataset (Afzal et al., 2022; Li et al., 2021), a large-scale dataset of natural images. However, natural images differ in many ways from MRI scans. Natural images are often colour images with RGB channels with a few salient objects, whereas MRI scans are grayscale images indicating magnetic features of brain tissue with less distinction between foreground and background. Moreover, MRI scans of the brain are 3D images with roughly consistent locations and contrasts for brain structures, though individual differences exist. Therefore, models pre-trained on large MRI datasets with image-level labels (e.g., age, sex) can be promising solutions (Dinsdale et al., 2021). Moreover, since there have been established pipelines to automatically segment WMH, another biomarker of CSVD, future studies may consider pre-training on the severity of WMH.

For the issue of image-level weak labels, some studies modified the U-Net architecture for segmentation to first resolve a task of predicting these weak labels (i.e., a regression task). The learnt weights were then transferred to the final segmentation model (Boutinaud et al., 2021). Alternatively, the global pooling layer can be removed to generate activation maps indicating the probability of target objects (Dubost et al., 2020). Future studies may also consider having these results checked and modified by domain experts to make gold-standard labels for segmentation. This will save much time compared to labelling lesions from scratch.

4.2. Data availability and quality

There are three dimensions regarding training data that need to be considered: *size, diversity*, and *accuracy*. Although a few techniques can be applied to reduce the required amount of data for training, a decent sample size is necessary to tune the network parameters for a specific task, and to avoid overfitting issues. However, including inaccurately labelled data will compromise the model performance. Due to the low prevalence of some CSVD lesions in the general population, including MRI data acquired from patients in clinical settings can help increase the size of data for training and validation. However, clinical MRI data are usually limited by the short acquisition time and therefore low spatial resolution. CMB, PVS and lacunes are small lesions, usually with diameters less than 5 mm. Clinical MRI scans with thick slices (3–5 mm) may miss some lesions due to the significant partial volume effects. In the case of PVS, since the number of PVS is usually large in the brain, studies usually examine representative regions (e.g., centrum semiovale, basal ganglia). A slightly thicker slice may be acceptable in such cases if at least one slice is acquired on the desired region. The location of the acquired image slice should be as consistent as possible across individuals through locating landmarks (e.g., anterior and posterior commissure). Having said that, overly thick slices will affect the richness of contextual information from nearby brain regions which is sometimes used in identifying PVS. A general recommendation of slice thickness ≤3 mm would be safe for extracting CMB, PVS and lacunes from clinical scans. In addition, a large number of similar cases contribute less than a good variety of data. The diversity of data can come from different sample types (healthy vs. various diseases) and scanners/scanning parameters. Studies using more than one scanner/cohort for development, validation and testing have been summarized in Supplementary Text 1.

Table 5
Publicly available code/datasets.

Link	Description	Ref.
https://github.com/Yonsei-MILab/Cerebral-Microbleeds-Detection	code for CMB segmentation	Al-Masni et al. (2020)
https://github.com/Yonsei-MILab/Lacunes-Identification	code for lacune segmentation	Al-Masni et al. (2021)
https://appsrv.cse.cuhk.edu.hk/~qdou/cmb-3dcnn/cmb-3dcnn.html	code and labelled data for CMB segmentation	Dou et al. (2016)
https://github.com/hjkuijf/MixLacune	code for lacune segmentation	–
https://valdo.grand-challenge.org/	"Where is VALDO" challenge, including segmenting PVS, CMB and lacunes, with example training data	–
https://github.com/hjkuijf/MixMicrobleed; https://hub.docker.com/r/hjkuijf/mixmicrobleed	Code for a deep learning-based toolbox to detect and segment CMB, participating in 2021 'Where is VALDO' challenge.	Girones Sanguesa et al. (2021)
https://github.com/hjkuijf/MixMicrobleedNet; https://hub.docker.com/r/hjkuijf/mixmicrobleednet	Code for a deep learning-based toolbox to detect and segment CMB, participating in 2021 'Where is VALDO' challenge.	Kuijf (2021)

On most occasions, these three factors (size, accuracy, diversity) cannot be all satisfied. Labelling small lesions, like CMB, PVS and lacunes, at the voxel level is time-consuming and subjective. Usually, multiple raters are assigned with the same set of data for labelling. Results are then compared, and for inconsistent labels, a consensus is reached by the majority vote, or being checked by a more experienced clinician. As a result, individual studies are unlikely to obtain sufficient high-accuracy voxel-level labels for training. Collaborations are a promising approach in this regard, as pooling small data samples with labels increases both the size and diversity of data. Public resources identified during the systemic review have been summarised in Table 5. Available studies have applied data augmentation with basic translations, rotations and mirroring to generate more data (Rashid et al., 2021). Introducing artefacts such as bias field and noise to the images may be helpful for the model to generalise to different samples/scanners (Billot et al., 2020). In addition, techniques such as uncertainty modelling (Karimi et al., 2020) and domain adaption (Guan and Liu, 2022) can be incorporated to address the issues due to inaccurate or subjective labels and excessive data heterogeneity due to the data acquisition process. Such methods have been developed for biomedical image analysis in general, but not particularly for the domain of CSVD lesions.

The MRI modalities to input into models are also worth considering carefully. Although there have been consensus MRI modalities for each CSVD lesion, a small number of studies included additional modalities (e.g., QSM (Rashid et al., 2021) and PD (Ghafaryasl et al., 2012)), that are not commonly available, to improve the performance. This may not be optimal when considering the generalisability of the model. Some techniques, such as knowledge distillation (Guan et al., 2021), could be considered to resolve this issue.

4.3. Performance evaluation

The performance of models would be more comparable if there is a reference dataset with a good diversity of data where the performance metrics can be calculated and compared. To this end, the 'Where is VALDO – Vascular Lesions Detection Challenge' (https://valdo.grand-challenge.org/Description/) is a good initiative. Training data with labels are provided, and a hidden dataset is used to compare performance. The selection of performance metrics should consider the clinical importance of the measures. For example, studies showed stronger associations of the volume, length, width and size of PVS with WMH burden (Ballerini et al., 2020), a well-established biomarker for CSVD, compared to the number of PVS. This emphasised the importance of segmenting PVS, and therefore voxel-wise performance metrics should be used. On the other hand, due to the low prevalence, studies of CMB (Lu et al., 2021a) and lacunes (Ghaznawi et al., 2019) usually examined the number or presence of these lesions. As a result, cluster- or image-level performance metrics may be more clinically relevant. Moreover, we propose that clinical validations (correlating with demographic and/or clinical measures) should supplement image processing performance metrics when assessing the models. Most available studies validated the algorithms through calculating performance metrics (e.g., sensitivity, specificity, etc.) with few investigating the relationship with clinical findings (Ballerini et al., 2020, 2016, 2018). Although our ultimate goal is to accurately segment CSVD lesions, missing a small proportion of lesions or lesion voxels usually do not significantly alter the correlation with clinical measures when studying large cohorts. Therefore, as the first step, it makes sense to aim at models with reasonable performance measurements which produce meaningful clinical associations in large datasets (e.g., UK Biobank (Miller et al., 2016), OASIS (Marcus et al., 2010)). Clinical measures to correlate with can include: (1) *vascular risk factors*, e.g., hypertension, hypercholesterolaemia, diabetes, atrial fibrillation, obesity, and smoking. (2) *performance in relevant cognitive domains*, e.g., executive function, processing speed, and 3) *other imaging biomarkers of small vessel disease and global atrophy*, e.g., white matter hyperintensities, cortical thickness, visual rating of small vessel disease lesions.

4.4. Factors to consider when selecting algorithms

There has been a clear increase in the number of CMB, PVS and lacunes studies using deep learning methods in recent years. However, the performance metrics do not seem to differ significantly between classical image processing, classical machine learning, and deep learning methods, although, strictly speaking, the metrics are not directly comparable as they were generated from different datasets. Classical image processing approaches require the least computational resources, but extensive expertise in image processing algorithms/filters, careful design, and sometimes manual work to reduce FPs (e.g., manually set thresholds from experience). Deep learning methods, on the other hand, need the least knowledge on image processing algorithms, but a large number of labelled data for training. Training deep learning models is also computationally expensive.

4.5. Conclusion

Advancements in artificial intelligence have resulted in promising computer-aided algorithms for the examination of CSVD lesions. Future studies could consider pooling data from different sources, techniques like weak supervision and transfer learning, and assessing the model performance with both image processing metrics and clinical correlations.

Data availability

No data was used for the research described in the article.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2022.119528.

References

Afzal, S., Khan, I.U., Lee, J.W., 2022. A transfer learning-based approach to detect cerebral microbleeds. Comput. Mater. Contin. 71 (1), 1903–1923.

Al-Masni, M.A., Kim, W.R., Kim, E.Y., Noh, Y., Kim, D.H., 2020. Automated detection of cerebral microbleeds in MR images: a two-stage deep learning approach. Neuroimage Clin. 28, 102464.

Al-Masni, M.A., Kim, W.R., Kim, E.Y., Noh, Y., Kim, D.H., 2021. 3D multi-scale residual network toward lacunar infarcts identification from MR images with minimal user intervention. IEEE Access 9, 11787–11797.

Balakrishnan, R., Valdes Hernandez, M.D.C., Farrall, A.J., 2021. Automatic segmentation of white matter hyperintensities from brain magnetic resonance images in the era of deep learning and big data - a systematic review. Comput. Med. Imaging Graph. 88, 101867.

Ballerini, L., Booth, T., Valdés Hernández, M.d.C., Wiseman, S., Lovreglio, R., Muñoz Maniega, S., Morris, Z., Pattie, A., Corley, J., Gow, A., Bastin, M.E., Deary, I.J., Wardlaw, J., 2020. Computational quantification of brain perivascular space morphologies: associations with vascular risk factors and white matter hyperintensities. A study in the Lothian birth cohort 1936. NeuroImage Clin. 25, 102120.

Ballerini, L., Lovreglio, R., Hernández, M.d.C.V., Gonzalez-Castro, V., Maniega, S.M., Pellegrini, L., Bastin, M.E., Deary, I.J., Wardlaw, J., 2016. Application of the ordered logit model to optimising frangi filter parameters for segmentation of perivascular spaces. Procedia Comput. Sci. 90, 61–67.

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.

Barnes, S.R., Haacke, E.M., Ayaz, M., Boikov, A.S., Kirsch, W., Kido, D., 2011. Semiautomated detection of cerebral microbleeds in magnetic resonance images. Magn. Reson. Imaging 29, 844–852.

Bian, W., Hess, C.P., Chang, S.M., Nelson, S.J., Lupo, J.M., 2013. Computer-aided detection of radiation-induced cerebral microbleeds on susceptibility-weighted MR images. NeuroImage Clin. 2, 282–290.

Billot, B., Bocchetta, M., Todd, E., Dalca, A.V., Rohrer, J.D., Iglesias, J.E., 2020. Automated segmentation of the hypothalamus and associated subunits in brain MRI. Neuroimage 223, 117287.

Boespflug, E.L., Schwartz, D.L., Lahna, D., Pollock, J., Iliff, J.J., Kaye, J.A., Rooney, W., Silbert, L.C., 2018. MR Imaging-based multimodal autoidentification of perivascular spaces (mMAPS): automated morphologic segmentation of enlarged perivascular spaces at clinical field strength. Radiology 286, 632–642.

Boutinaud, P., Tsuchida, A., Laurent, A., Adonias, F., Hanifehlou, Z., Nozais, V., Verrecchia, V., Lampe, L., Zhang, J., Zhu, Y.C., Tzourio, C., Mazoyer, B., Joliot, M., 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.

Brown, R., Low, A., Markus, H.S., 2021. Rate of, and risk factors for, white matter hyperintensity growth: a systematic review and meta-analysis with implications for clinical trial design. J. Neurol. Neurosurg. Psychiatry 92, 1271.

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

Chen, H., Yu, L., Dou, Q., Shi, L., Mok, V.C.T., Heng, P.A., 2015. Automatic detection of cerebral microbleeds via deep learning based 3D feature representation. In: Proceedings of the IEEE 12th International Symposium on Biomedical Imaging (ISBI), pp. 764–767.

Chen, Y., Villanueva-Meyer, J.E., Morrison, M.A., Lupo, J.M., 2019. Toward automatic detection of radiation-induced cerebral microbleeds using a 3D deep residual network. J. Digit. Imaging 32, 766–772.

Cheng, A.L., Batool, S., McCreary, C.R., Lauzon, M.L., Frayne, R., Goyal, M., Smith, E.E., 2013. Susceptibility-weighted imaging is more reliable than T2*-weighted gradient-recalled echo MRI for detecting microbleeds. Stroke 44, 2782–2786.

Chesebro, A.G., Amarante, E., Lao, P.J., Meier, I.B., Mayeux, R., Brickman, A.M., 2021. Automated detection of cerebral microbleeds on T2*-weighted MRI. Sci. Rep. 11, 4004.

Das, A.S., Regenhardt, R.W., Vernooij, M.W., Blacker, D., Charidimou, A., Viswanathan, A., 2019. Asymptomatic cerebral small vessel disease: insights from population-based studies. J. Stroke 21, 121–138.

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

Dinsdale, N.K., Bluemke, E., Sundaresan, V., Jenkinson, M., Smith, S., Namburete, A.IL. 2021. Challenges for machine learning in clinical translation of big data imaging studies. arXiv:2107.05630. doi:10.48550/arXiv.2107.05630.

deSouza, N.M., Achten, E., Alberich-Bayarri, A., Bamberg, F., Boellaard, R., Clément, O., Fournier, L., Gallagher, F., Golay, X., Heussel, C.P., Jackson, E.F., Manniesing, R.,

Mayerhofer, M.E., Neri, E., O'Connor, J., Oguz, K.K., Persson, A., Smits, M., van Beek, E.J.R., Zech, C.J.European Society of, R., 2019. Validated imaging biomarkers as decision-making tools in clinical trials and routine practice: current status and recommendations from the EIBALL* subcommittee of the European society of radiology (ESR). Insights Imaging 10, 87.

Dou, Q., Chen, H., Yu, L., Shi, L., Wang, D., Mok, V.C., Heng, P.A., 2015. Automatic cerebral microbleeds detection from MR images via Independent Subspace Analysis based hierarchical features. In: Proceedings of the 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 7933–7936.

Dou, Q., Chen, H., Yu, L., Zhao, L., Qin, J., Wang, D., Mok, V.C., Shi, L., Heng, P., 2016. Automatic detection of cerebral microbleeds from MR images via 3D convolutional neural networks. IEEE Trans. Med. Imaging 35, 1182–1195.

Dubost, F., Adams, H., Bortsova, G., Ikram, M.A., Niessen, W., Vernooij, M., de Bruijne, M., 2019a. 3D regression neural network for the quantification of enlarged perivascular spaces in brain MRI. Med. Image Anal. 51, 89–100.

Dubost, F., Adams, H., Yilmaz, P., Bortsova, G., Tulder, G.v., Ikram, M.A., Niessen, W., Vernooij, M.W., Bruijne, M.d., 2020. Weakly supervised object detection with 2D and 3D regression neural networks. Med. Image Anal. 65, 101767.

Dubost, F., Bortsova, G., Adams, H.H.H., Ikram, M.A., Niessen, W.J., Vernooij, M.W., de Bruijne, M., 2017. GP-Unet: Lesion Detection from Weak Labels With a 3D Regression Network. MICCAI.

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

Fazlollahi, A., Meriaudeau, F., Giancardo, L., Villemagne, V.L., Rowe, C.C., Yates, P., Salvado, O., Bourgeat, P., 2015. Computer-aided detection of cerebral microbleeds in susceptibility-weighted imaging. Comput. Med. Imaging Graph. 46 (Pt 3), 269–276.

Fazlollahi, A., Meriaudeau, F., Villemagne, V.L., Rowe, C.C., Yates, P., Salvado, O., Bourgeat, P., 2014. Efficient machine learning framework for computer-aided detection of cerebral microbleeds using the Radon transform. In: Proceedings of the IEEE 11th International Symposium on Biomedical Imaging (ISBI), pp. 113–116.

Frangi, A.F., Niessen, W.J., Vincken, K.L., Viergever, M.A., Wells, W.M., Colchester, A., Delp, S., 1998. Multiscale vessel enhancement filtering. In: Medical Image Computing and Computer-Assisted Intervention — MICCAI'98. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 130–137.

Ghafaryasl, B., Lijn, F.v.d., Poels, M., Vrooman, H., Ikram, M.A., Niessen, W.J., Lugt, A.v.d., Vernooij, M., Bruijne, M.d., 2012. A computer aided detection system for cerebral microbleeds in brain MRI. In: Proceedings of the 9th IEEE International Symposium on Biomedical Imaging (ISBI), pp. 138–141.

Ghafoorian, M., Karssemeijer, N., Heskes, T., Bergkamp, M., Wissink, J., Obels, J., Keizer, K., de Leeuw, F.E., Ginneken, B.V., Marchiori, E., Platel, B., 2017. Deep multi-scale location-aware 3D convolutional neural networks for automated detection of lacunes of presumed vascular origin. Neuroimage Clin. 14, 391–399.

Ghaznawi, R., Geerlings, M.I., Jaarsma-Coes, M.G., Zwartbol, M.H., Kuijf, H.J., van der Graaf, Y., Witkamp, T.D., Hendrikse, J., de Bresser, J., 2019. The association between lacunes and white matter hyperintensity features on MRI: the SMART-MR study. J. Cereb. Blood Flow Metab. 39, 2486–2496.

Girones Sanguesa, M., Kutnar, D., van der Velden, B.H.M., Kuijf, H.J., 2021. MixMicrobleed: Multi-stage detection and segmentation of cerebral microbleeds. arXiv:2108.02482. doi:10.48550/arXiv.2108.02482.

González-Castro, V., Valdés Hernández, M.d.C., Armitage, P.A., Wardlaw, J.M., Campilho, A., Karray, F., 2016. Automatic rating of perivascular spaces in brain MRI using bag of visual words. In: Image Analysis and Recognition. Springer International Publishing, Cham, pp. 642–649.

González-Castro, V., Valdés Hernández, M.D.C., Chappell, F.M., Armitage, P.A., Makin, S., Wardlaw, J.M., 2017. Reliability of an automatic classifier for brain enlarged perivascular spaces burden and comparison with human performance. Clin. Sci. 131, 1465–1481 (Lond.).

Greenberg, S.M., Vernooij, M.W., Cordonnier, C., Viswanathan, A., Al-Shahi Salman, R., Warach, S., Launer, L.J., Van Buchem, M.A., Breteler, M.M., 2009. Cerebral microbleeds: a guide to detection and interpretation. Lancet Neurol. 8, 165–174.

Gregoire, S.M., Chaudhary, U.J., Brown, M.M., Yousry, T.A., Kallis, C., Jager, H.R., Werring, D.J., 2009. The microbleed anatomical rating scale (MARS): reliability of a tool to map brain microbleeds. Neurology 73, 1759–1766.

Groeschel, S., Chong, W.K., Surtees, R., Hanefeld, F., 2006. Virchow-Robin spaces on magnetic resonance images: normative data, their dilatation, and a review of the literature. Neuroradiology 48, 745–754.

Guan, H., Liu, M., 2022. Domain adaptation for medical image analysis: a survey. IEEE Trans. Biomed. Eng. 69, 1173–1185.

Guan, H., Wang, C., Tao, D., 2021. MRI-based Alzheimer's disease prediction via distilling the knowledge in multi-modal data. Neuroimage 244, 118586.

Guerrero, R., Qin, C., Oktay, O., Bowles, C., Chen, L., Joules, R., Wolz, R., Valdés-Hernández, M.C., Dickie, D.A., Wardlaw, J., Rueckert, D., 2018. White matter hyperintensity and stroke lesion segmentation and differentiation using convolutional neural networks. Neuroimage Clin. 17, 918–934.

Hernandez Mdel, C., Piper, R.J., Wang, X., Deary, I.J., Wardlaw, J.M., 2013. Towards the automatic computational assessment of enlarged perivascular spaces on brain magnetic resonance images: a systematic review. J. Magn. Reson. Imaging 38, 774–785.

Hong, J., Cheng, H., Zhang, Y.D., Liu, J., 2019. Detecting cerebral microbleeds with transfer learning. Mach. Vis. Appl. 30, 1123–1133.

Hong, J., Wang, S.H., Cheng, H., Liu, J., 2020. Classification of cerebral microbleeds based on fully-optimized convolutional neural network. Multimed. Tools Appl. 79, 15151–15169.

Hou, Y., Park, S.H., 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. Sci. Rep. 7, 8569 -8569.

Ibrahim, A.A., Ibrahim, Y.A., Darwish, E.A., Khater, N.H., 2019. Prevalence of cerebral microbleeds and other cardiovascular risk factors in elderly patients with acute ischemic stroke. Egypt. J. Radiol. Nucl. Med. 50, 38.

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.

Karimi, D., Dou, H., Warfield, S.K., Gholipour, A., 2020. Deep learning with noisy labels: exploring techniques and remedies in medical image analysis. Med. Image Anal. 65, 101759.

Kuijf, H.J., 2021. MixMicrobleedNet: segmentation of cerebral microbleeds using nnU-Net. arXiv:2108.01389. doi:10.48550/arXiv.2108.01389.

Kuijf, H.J., Bresser, J.d., Biessels, G.J., Viergever, M.A., Vincken, K.L., 2011. Detecting cerebral microbleeds in 7.0 T MR images using the radial symmetry transform. In: Proceedings of the IEEE International Symposium on Biomedical Imaging: From Nano to Macro, pp. 758–761.

Kuijf, H.J., Brundel, M., de Bresser, J., van Veluw, S.J., Heringa, S.M., Viergever, M.A., Biessels, G.J., Vincken, K.L., 2013. Semi-automated detection of cerebral microbleeds on 3.0 T MR images. PLoS One 8, e66610.

Kuijf, H.J., de Bresser, J., Geerlings, M.I., Conijn, M.M.A., Viergever, M.A., Biessels, G.J., Vincken, K.L., 2012. Efficient detection of cerebral microbleeds on 7.0T MR images using the radial symmetry transform. Neuroimage 59, 2266–2273.

Li, T., Zou, Y., Bai, P., Li, S., Wang, H., Chen, X., Meng, Z., Kang, Z., Zhou, G., 2021. Detecting cerebral microbleeds via deep learning with features enhancement by reusing ground truth. Comput. Methods Progr. Biomed. 204, 106051.

Lian, C., Liu, M., Zhang, J., Zong, X., Lin, W., Shen, D., 2018a. Automatic Segmentation of 3D Perivascular Spaces in 7T MR images using multi-channel fully convolutional network. In: Proceedings of the International Society for Magnetic Resonance in Medicine. Scientific Meeting and Exhibition, 2018.

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

Liu, C., Habib, T., Salimeen, M., Pradhan, A., Singh, M., Wang, M., Wu, F., Zhang, Y., Gao, L., Yang, G., Li, X., Yang, J., 2020a. Quantification of visible Virchow-Robin spaces for detecting the functional status of the glymphatic system in children with newly diagnosed idiopathic generalized epilepsy. Seizure 78, 12–17.

Liu, H., Rashid, T., Habes, M., 2020b. Cerebral microbleed detection via fourier descriptor with dual domain distribution modeling. In: Proceedings of the IEEE 17th International Symposium on Biomedical Imaging Workshops (ISBI Workshops), pp. 1–4.

Liu, S., Utriainen, D., Chai, C., Chen, Y., Wang, L., Sethi, S.K., Xia, S., Haacke, E.M., 2019. Cerebral microbleed detection using susceptibility weighted imaging and deep learning. Neuroimage 198, 271–282.

Lu, D., Liu, J., MacKinnon, A.D., Tozer, D.J., Markus, H.S., 2021a. Prevalence and risk factors of cerebral microbleeds. Neurology 97, e1493.

Lu, S., Liu, S., Wang, S.H., Zhang, Y.D., 2021b. Cerebral microbleed detection via convolutional neural network and extreme learning machine. Front. Comput. Neurosci. 15, 738885.

Lu, S., Lu, Z., Hou, X., Cheng, H., Wang, S., 2017. Detection of cerebral microbleeding based on deep convolutional neural network. In: Proceedings of the 14th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP), pp. 93–96.

Madan, C.R., 2021. Scan once, analyse many: using large open-access neuroimaging datasets to understand the brain. Neuroinformatics 1–29. doi:10.1007/s12021-021-09519-6.

Marcus, D.S., Fotenos, A.F., Csernansky, J.G., Morris, J.C., Buckner, R.L., 2010. Open access series of imaging studies: longitudinal MRI data in nondemented and demented older adults. J. Cogn. Neurosci. 22, 2677–2684.

Miller, K.L., Alfaro-Almagro, F., Bangerter, N.K., Thomas, D.L., Yacoub, E., Xu, J., Bartsch, A.J., Jbabdi, S., Sotiropoulos, S.N., Andersson, J.L.R., Griffanti, L., Douaud, G., Okell, T.W., Weale, P., Dragonu, I., Garratt, S., Hudson, S., Collins, R., Jenkinson, M., Matthews, P.M., Smith, S.M., 2016. Multimodal population brain imaging in the UK Biobank prospective epidemiology study. Nat. Neurosci. 19, 1523–1536.

Mittal, S., Wu, Z., Neelavalli, J., Haacke, E.M., 2009. Susceptibility-weighted imaging: technical aspects and clinical applications, part 2. AJNR Am. J. Neuroradiol. 30, 232–252.

Morrison, M.A., Payabvash, S., Chen, Y., Avadiappan, S., Shah, M., Zou, X., Hess, C.P., Lupo, J.M., 2018. A user-guided tool for semi-automated cerebral microbleed detection and volume segmentation: evaluating vascular injury and data labelling for machine learning. NeuroImage Clin. 20, 498–505.

Myung, J., Lee, K.M., Kim, H.G., Oh, J., Lee, J.Y., Shin, I., Kim, E.J., Lee, J.S., 2021. Novel approaches to detection of cerebral microbleeds: single deep learning model to achieve a balanced performance. J. Stroke Cerebrovasc. Dis. 30, 105886.

Park, S.H., Zong, X., Gao, Y., Lin, W., Shen, D., 2016. Segmentation of perivascular spaces in 7T MR image using auto-context model with orientation-normalized features. Neuroimage 134, 223–235.

Poels, M.M., Vernooij, M.W., Ikram, M.A., Hofman, A., Krestin, G.P., van der Lugt, A., Breteler, M.M., 2010. Prevalence and risk factors of cerebral microbleeds: an update of the Rotterdam scan study. Stroke 41, S103–S106.

Rashid, T., Abdulkadir, A., Nasrallah, I.M., Ware, J.B., Liu, H., Spincemaille, P., Romero, J.R., Bryan, R.N., Heckbert, S.R., Habes, M., 2021. DEEPMIR: a deep neural network for differential detection of cerebral microbleeds and iron deposits in MRI. Sci. Rep. 11, 14124.

Roerdink, J.B.T.M., Meijster, A., 2000. The watershed transform: definitions, algorithms and parallelization strategies. Fundam. Inform. 41, 187–228.

Roy, S., Jog, A., Magrath, E., Butman, J., Pham, D., 2015. Cerebral microbleed segmentation from susceptibility weighted images. Preceedings of the SPIE.

Ruetten, P.P.R., Gillard, J.H., Graves, M.J., 2019. Introduction to quantitative susceptibility mapping and susceptibility weighted imaging. Br. J. Radiol. 92, 20181016.

Sarmento, R.M., Vasconcelos, F.F.X., Filho, P.P.R., Wu, W., de Albuquerque, V.H.C., 2020. Automatic neuroimage processing and analysis in stroke-A systematic review. IEEE Rev. Biomed. Eng. 13, 130–155.

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 T(1) and FLAIR MR imaging. Neuroimage 202, 116126.

Seghier, M.L., Kolanko, M.A., Leff, A.P., Jager, H.R., Gregoire, S.M., Werring, D.J., 2011. Microbleed detection using automated segmentation (MIDAS): a new method applicable to standard clinical MR images. PLoS One 6, e17547.

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.

Shams, S., Martola, J., Cavallin, L., Granberg, T., Shams, M., Aspelin, P., Wahlund, L.O., Kristoffersen-Wiberg, M., 2015. SWI or T2*: which MRI sequence to use in the detection of cerebral microbleeds? The Karolinska imaging dementia study. AJNR Am. J. Neuroradiol. 36, 1089–1095.

Sudre, C.H., Anson, B.G., Ingala, S., Lane, C.D., Jimenez, D., Haider, L., Varsavsky, T., Smith, L., Jäger, H.R., Cardoso, M.J., 2019. 3D multirater RCNN for multimodal multiclass detection and characterisation of extremely small objects. In: Proceedings of the MIDL, pp. 447–456.

Tajudin, A.S., Sulaiman, S.N., Isa, I.S., Soh, Z.H.C., Karim, N.K.A., Shuaib, I.L., 2017a. Microbleeds detection using watershed-driven active contour. In: Proceedings of the 7th IEEE International Conference on Control System, Computing and Engineering (ICCSCE), pp. 320–324.

Tajudin, A.S., Sulaiman, S.N., Isa, I.S., Soh, Z.H.C., Tahir, N.M., Karim, N.K.A., Shuaib, I.L., 2017b. An improved watershed segmentation technique for microbleeds detection in MRI images. In: Proceedings of the International Conference on Electrical, Electronics and System Engineering (ICEESE), pp. 11–16.

Uchiyama, Y., Abe, A., Muramatsu, C., Hara, T., Shiraishi, J., Fujita, H., 2015. Eigenspace template matching for detection of lacunar infarcts on MR images. J. Digit. Imaging 28, 116–122.

Uchiyama, Y., Asano, T., Hara, T., Fujita, H., Hoshi, H., Iwama, T., Kinosada, Y., 2009. CAD scheme for differential diagnosis of lacunar infarcts and normal Virchow-Robin spaces on brain MR images. IFMBE 25, 126–128.

Uchiyama, Y., Asano, T., Hara, T., Kanematsu, M., Hoshi, H., Iwama, T., Fujita, H., 2012. Computer-aided diagnosis for detection of lacunar infarcts on MR images: ROC analysis of radiologists' performance. J. Digit. Imaging 25, 497–503.

Uchiyama, Y., Kunieda, T., Asano, T., Kato, H., Hara, T., Kanematsu, M., Iwama, T., Hoshi, H., Kinosada, Y., Fujita, H., 2008. Computer-aided diagnosis scheme for classification of lacunar infarcts and enlarged Virchow-Robin spaces in brain MR images. In: Proceedings of the 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 3908–3911.

Uchiyama, Y., Yokoyama, R., Ando, H., Asano, T., Kato, H., Yamakawa, H., Yamakawa, H., Hara, T., Iwama, T., Hoshi, H., Fujita, H., 2007a. Computer-aided diagnosis scheme for detection of lacunar infarcts on MR images. Acad. Radiol. 14, 1554–1561.

Uchiyama, Y., Yokoyama, R., Ando, H., Asano, T., Kato, H., Yamakawa, H., Yamakawa, H., Hara, T., Iwama, T., Hoshi, H., Fujita, H., 2007b. Improvement of automated detection method of lacunar infarcts in brain MR images. In: Proceedings of the 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 1599–1602.

van den Heuvel, T.L., Ghafoorian, M., van der Eerden, A., Goraj, B., Andriessen, T.M., ter Haar Romeny, B., Platel, B., 2015. Computer aided detection of brain micro-bleeds in traumatic brain injury. In: Proceedings of the SPIE.

van den Heuvel, T.L.A., van der Eerden, A.W., Manniesing, R., Ghafoorian, M., Tan, T., Andriessen, T.M.J.C., Vande Vyvere, T., van den Hauwe, L., ter Haar Romeny, B.M., Goraj, B.M., Platel, B., 2016. Automated detection of cerebral microbleeds in patients with traumatic brain injury. NeuroImage Clin. 12, 241–251.

Verdelho, A., Biessels, G.J., Chabriat, H., Charidimou, A., Duering, M., Godefroy, O., Pantoni, L., Pavlovic, A., Wardlaw, J., 2021. Cerebrovascular disease in patients with cognitive impairment: a white paper from the ESO dementia committee - a practical point of view with suggestions for the management of cerebrovascular diseases in memory clinics. Eur. Stroke J. 6, 111–119.

Wang, S., Jiang, Y., Hou, X., Cheng, H., Du, S., 2017. Cerebral micro-bleed detection based on the convolution neural network with rank based average pooling. IEEE Access 5, 16576–16583.

Wang, S., Tang, C., Sun, J., Zhang, Y., 2019. Cerebral micro-bleeding detection based on densely connected neural network. Front. Neurosci. 13, 1–11 article 422.

Wang, X., Valdés Hernández Mdel, C., Doubal, F., Chappell, F.M., Piper, R.J., Deary, I.J., Wardlaw, J., 2016. Development and initial evaluation of a semi-automatic approach to assess perivascular spaces on conventional magnetic resonance images. J. Neurosci. Methods 257, 34–44.

Wang, Y., Catindig, J.A., Hilal, S., Soon, H.W., Ting, E., Wong, T.Y., Venketasubramanian, N., Chen, C., Qiu, A., 2012. Multi-stage segmentation of white matter hyperintensity, cortical and lacunar infarcts. Neuroimage 60, 2379–2388.

Wardlaw, J.M., Smith, E.E., Biessels, G.J., Cordonnier, C., Fazekas, F., Frayne, R., Lindley, R.I., O'Brien, J.T., Barkhof, F., Benavente, O.R., Black, S.E., Brayne, C., Breteler, M., Chabriat, H., DeCarli, C., de Leeuw, F.E., Doubal, F., Duering, M., Fox, N.C., Greenberg, S., Hachinski, V., Kilimann, I., Mok, V., van Oostenbrugge, R., Pantoni, L., Speck, O., Stephan, B.C.M., Teipel, S., Viswanathan, A., Werring, D., Chen, C., Smith, C., van Buchem, M., Norrving, B., Gorelick, P.B., Dichgans, M., Changes, S.R.V., 2013. Neuroimaging standards for research into small vessel disease and its contribution to ageing and neurodegeneration. Lancet Neurol. 12, 822–838.

Yang, E., Gonuguntla, V., Moon, W.J., Moon, Y., Kim, H.J., Park, M., Kim, J.H., 2021.

Direct rating estimation of enlarged perivascular spaces (EPVS) in brain MRI using deep neural network. Appl. Sci. 11, 9398.

Yokoyama, R., Zhang, X., Uchiyama, Y., Fujita, H., Hara, T., Zhou, X., Kanematsu, M., Asano, T., Kondo, H., Goshima, S., Hoshi, H., Iwama, T., 2007. Development of an automated method for the detection of chronic lacunar infarct regions in brain MR images. IEICE Trans. Inf. Syst. 90-D, 943–954.

Zhang, J., Gao, Y., Park, S.H., Zong, X., Lin, W., Shen, D., 2016a. Segmentation of perivascular spaces using vascular features and structured random forest from 7T MR image. Mach. Learn. Med. Imaging 10019, 61–68.

Zhang, J., Gao, Y., Park, S.H., Zong, X., Lin, W., Shen, D., 2017. Structured learning for 3-D perivascular space segmentation using vascular features. IEEE Trans. Biomed. Eng. 64, 2803–2812.

Zhang, Y.D., Hou, X.X., Chen, Y., Chen, H., Yang, M., Yang, J., Wang, S.H., 2018a. Voxelwise detection of cerebral microbleed in CADASIL patients by leaky rectified linear unit and early stopping. Multimed. Tools Appl. 77, 21825–21845.

Zhang, Y.D., Hou, X.X., Chen, Y., Chen, H., Wang, S.H., 2018b. Seven-layer deep neural network based on sparse autoencoder for voxelwise detection of cerebral microbleed. Multimed. Tools Appl. 77, 10521–10538.

Zhang, Y.D., Hou, X.X., Lv, Y.D., Chen, H., Zhang, Y., Wang, S.H., 2016b. Sparse autoencoder based deep neural network for voxelwise detection of cerebral microbleed. In: Proceedings of the IEEE 22nd International Conference on Parallel and Distributed Systems (ICPADS), pp. 1229–1232.