

Computer-aided detection of cerebral microbleeds in susceptibility-weighted imaging



Amir Fazlollahi^{a,b,*}, Fabrice Meriaudeau^b, Luca Giancardo^c, Victor L. Villemagne^d, Christopher C. Rowe^d, Paul Yates^d, Olivier Salvado^a, Pierrick Bourgeat^a, on behalf of the AIBL Research Group¹

^a CSIRO Digital Productivity Flagship, The Australian e-Health Research Centre, Herston, QLD, Australia

^b Le2I, University of Burgundy, Le Creusot, France

^c RLE, Massachusetts Institute of Technology, MA, USA

^d Department of Nuclear Medicine and Centre for PET, Austin Hospital, Melbourne, VIC, Australia

ARTICLE INFO

Article history:

Received 30 December 2014

Received in revised form 3 August 2015

Accepted 2 October 2015

Keywords:

Cerebral microbleed

Susceptibility-weighted imaging

Radon transform

Multi-scale Laplacian of Gaussian

Random forests

ABSTRACT

Susceptibility-weighted imaging (SWI) is recognized as the preferred MRI technique for visualizing cerebral vasculature and related pathologies such as cerebral microbleeds (CMBs). Manual identification of CMBs is time-consuming, has limited reliability and reproducibility, and is prone to misinterpretation. In this paper, a novel computer-aided microbleed detection technique based on machine learning is presented: First, spherical-like objects (potential CMB candidates) with their corresponding bounding boxes were detected using a novel multi-scale Laplacian of Gaussian technique. A set of robust 3-dimensional Radon- and Hessian-based shape descriptors within each bounding box were then extracted to train a cascade of binary random forests (RF). The cascade consists of consecutive independent RF classifiers with low to high posterior probability constraints to handle imbalanced training sets (CMBs and non-CMBs), and to progressively improve detection rates. The proposed method was validated on 66 subjects whose CMBs were manually stratified into “possible” and “definite” by two medical experts. The proposed technique achieved a sensitivity of 87% and an average false detection rate of 27.1 CMBs per subject on the “possible and definite” set. A sensitivity of 93% and false detection rate of 10 CMBs per subject was also achieved on the “definite” set. The proposed automated approach outperforms state of the art methods, and promises to enhance manual expert screening. Benefits include improved reliability, minimization of intra-rater variability and a reduction in assessment time.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Cerebral microbleeds (CMBs) represent small foci hemosiderin deposition in normal brain tissue, corresponding to prior microscopic hemorrhages [1]. CMBs are relatively small in size (few millimeters), appear spherical and hypointense on magnetic resonance images (MRI), and are becoming recognized as important imaging findings in cognitive impairment and dementia [2], stroke and intracerebral haemorrhages [3] and cerebral amyloid angiopathy [4].

Manual localization and counting is the gold standard technique to assess the incidence of CMBs. This is a challenging task, as it is extremely time-consuming, subjective, and has high inter-expert variability [5]. Moreover, visual screening is prone to errors as small CMBs can be easily missed, or mistaken for vessel cross-sections, especially in MR susceptibility-weighted imaging (SWI). SWI is a 3D flow-compensated T2* gradient echo which maximizes the sensitivity of susceptibility effects by combining filtered phase information with magnitude information [6]. SWI has several advantages over T2* gradient-recall echo for detecting CMBs, such as 2 to 6-folds increased sensitivity in CMB detection [7,8], and enhanced the resolution of smaller lesions (i.e. less than 1 mm in diameter) [7]. Furthermore, to improve inter-rater agreement, established rating scales have separated CMBs into two categories: *definite* and *possible* [9]. Definite CMBs, which are defined spherical with enough hypointensity, have been widely considered in clinical studies and assessments [10], since they increase detection

* Corresponding author at: Level 5, UQ Health Sciences Building, 901/16 Royal Brisbane and Womens Hospital, Herston 4029, Queensland, Australia. Tel.: +61 732533618; mobile: +61 405292473.

E-mail address: fazlollahi@gmail.com (A. Fazlollahi).

¹ <http://www.aibl.csiro.au/>.

Table 1
Table below summarizes the performance of the proposed method and other published computer-aided microbleed detection approaches on both Possible and Definite CMBs.

Method	MRI modality	Mag. field	Pathology	# of subjects	# of CMBs	Sensitivity	Ave. FPs
Ref. [11]	T2*	1.5 T	Stroke	30	114	50%	–
Ref. [12]	SWI	1.5 T	AD/MCI	6	126	81.7%	107
Ref. [13]	T2*	7.0 T	CerbVa ^a	18	66	71.2%	17.2
Ref. [14]	SWI	3.0 T	Tumor	10	304	86.5%	44.9
Ref. [5]	T2*	3.0 T	AD/MCI	38 ^b	148	78%	49
Proposed method	SWI	3.0 T	AD/MCI	66	231	87.0%	27.1

AD, Alzheimer's disease; MCI, mild cognitive impairment; CMB, cerebral microbleeds; SWI, susceptibility-weighted imaging; FP, false-positive.

^a CerbVa: cerebrovascular disease.

^b From a total of 72 subjects just 38 were confirmed with CMBs.

confidence [9]. For instance, it has been suggested to exclude individuals with 2 or more definite CMBs (using SWI) from amyloid-modifying therapeutic trials due to the potential of adverse drug events [10].

To aid manual CMB assessments and eventually supplant them with automatic approaches, advanced computer vision and image analysis approaches have been investigated [11–14,5]. These studies implemented a two-step method consisting of an automated pre-screening to identify all CMB-like regions, followed by a high-level detection model to refine falsely detected regions. However, these methods suffered from a high false detection rate and low sensitivity as summarized in Table 1. The pre-screening is designed to achieve a high sensitivity at the cost of specificity due to the diverse appearance of both CMBs and surrounding structures seen on MRI. In the second step, machine learning techniques have not been well explored compared to the rule-based or statistical approaches. However, such a technique may face an imbalanced classification problem since the low-level pre-screening identifies large number of non-CMB regions (majority class) compared to true CMBs (minority class). In such cases, conventional classifiers tend to perform poorly [15], though, it has been demonstrated that random sampling with bagging or boosting strategies have the highest performance in large and imbalanced datasets [16].

This paper presents a novel computer-aided technique for detecting cerebral microbleeds in SWI, particularly aimed at achieving a high sensitivity with a low false detection rate. The proposed technique is comprised of three main steps. First, a low-level automated pre-screening scheme based on a multi-scale Laplacian of Gaussian (LoG) is performed to detect potential CMBs (candidates), with the ultimate goal of achieving a high sensitivity leaving almost no true CMBs behind, but at the cost of increased candidates. Second, a 3D bounding box proportional to the size of the central hypointensity is defined for each candidate, and subsequently robust Radon and Hessian shape feature descriptors which are invariant to scale, rotation, translation and intensity shift are computed. To tackle the class imbalance problem for classification, the third step employs a cascade of consecutive binary random forest (RF) classifiers to progressively reduce the training sample size (false detections) throughout the cascade and reach a high detection accuracy. The overview of the proposed method is shown in Fig. 1.

The novelties of the proposed approach are in the following areas: (i) identifying spherical shapes and their extent through a novel multi-scale Laplacian of Gaussian (LoG) technique, and (ii) addressing the class imbalance issue by incorporating a series of binary RF classifiers with low to high posterior probability constraints. Unlike [17,18], the proposed multi-scale LoG approach is adapted for MR images especially with anisotropic resolutions which cause some CMBs to appear relatively semi-spherical or ovoid in shape. Furthermore, the modified Laplacian operator response is capable of detecting of CMBs adjacent to vessels and sulcus folding. The multi-layer classifier approach in contrast to

[19,20] not only handles class imbalance problem in an efficient way using several RF classifiers, but it also provides a flexible framework to obtain a desirable sensitivity/specificity by extending the number of layers, and updating a single parameter (posterior probability) at each layer.

2. Materials and methods

2.1. Dataset

Data used in the preparation of this paper were obtained from the Australian Imaging, Biomarkers and Lifestyle (AIBL) study [21]. The 66 participants included 11 subjects classified as Alzheimer's Disease (AD), 14 as mild cognitive impairment (MCI) and 41 healthy controls (HC). The subjects included 27 females (age 77.9 ± 7.7 , mean \pm std) and 39 males (age 71.9 ± 6.6).

All subjects underwent an anatomical T1-weighted (T1w) and a SWI acquisitions on a 3T Siemens TRIO scanner, where SWIs were automatically reconstructed online using the scanner system (software VB17). The 3D SWIs were acquired with $0.93 \text{ mm} \times 0.93 \text{ mm}$ in-plane resolution and 1.75 mm slice thickness, repetition time/echo time of 27/20 ms, and flip angle 20° . T1w images were acquired using a standard 3D magnetization-prepared rapid gradient echo sequence with in-plane resolution $1.0 \text{ mm} \times 1.0 \text{ mm}$, slice thickness 1.2 mm, repetition-time/echo-time/T1 = 2.300/2.98/900, flip angle 9° , field of view 240×256 , and 160 slices.

SWIs were manually inspected for CMBs by two clinical experts using the MARS rating method [9], labeling CMBs as *definite* and

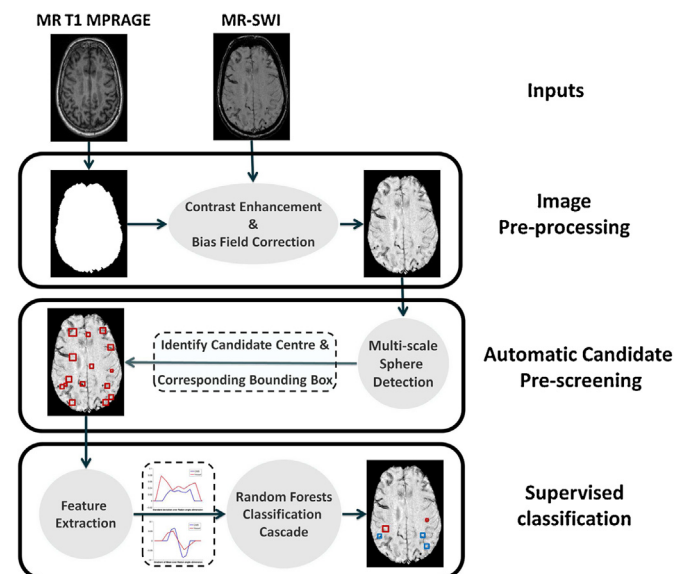


Fig. 1. Overview of the proposed processing pipeline.

possible. CMBs are considered as spherical hypointense lesions smaller than or equal to 10 mm [22]. Suspected lesions are usually tracked through multiple slices to exclude them from blood vessels. After independent reading by the 2 readers, a consensus was reached for presence and number of lesions on each scan. The agreement between readers 1 and 2 was 83% ($\kappa = 0.65$); reader 1 vs. consensus 91% ($\kappa = 0.8$); and reader 2 vs. consensus 91% ($\kappa = 0.83$).

2.2. Image preprocessing

The low frequency intensity non-uniformity present in the SWI was corrected using the N4 bias field correction technique [23]. For each subject, the T1w image was also classified into gray/white matter and CSF using an implementation of the expectation maximization segmentation algorithm [24]. The resulting gray/white matter and CSF segmentations were combined to generate a skull-stripping mask. The T1w was then rigidly aligned to the SWI, and the resulting transformation was applied to propagate the skull-stripped mask in order to exclude non-brain tissues from the analysis. From here on SWI refers to the skull-stripped bias field corrected SWI.

SWIs suffer from signal saturation near the air-tissue interface, such as sinus regions. To enhance the contrast, and also normalize the intensity across subjects, the dynamic intensity range of SWIs was normalized to [0, 1] after trimming the top 1% of intensity values. The image intensity was then inverted so that CMBs and vessels appear hyperintense which simplifies the multi-scale Gaussian analysis. In addition, to locally enhance the contrast and reduce noise, adaptive histogram equalization (window size of 1.5 voxel, $\alpha = 1.5$ and $\beta = 0.75$) and gradient based anisotropic diffusion (20 iterations with step size of 0.04) were also applied on the SWIs.

2.3. Step 1: Automatic candidate pre-screening

A low-level automatic scheme is proposed to identify potential spherical-like CMBs (candidates) and their extents with the objective of achieving a high detection sensitivity leaving no true CMB behind. The scheme considers: (1) CMBs that appear semi-spherical due to anisotropic spatial resolutions and strong partial volume effects (PVE) of SWIs, and (2) CMBs touching/adjacent to other structures.

In the current work, a similar approach as in [25] was used, where we proposed a new multi-scale spherical-like object detection for 3D brain images with a high sensitivity to small objects and a low number of false-positives (FPs). For methodological detail of the multi-scale sphere detection method refer to [Supplementary Material](#).

2.4. Step 2: Shape feature extraction

The pre-screening step extracts the centroid of potential CMBs. Subsequently, various descriptive shape features invariant to scale, rotation, translation and intensity are computed by defining a precise 3D bounding box per candidate.

2.4.1. Candidate bounding box detection

Prior to feature extraction, a 3D bounding box (or ROI) has to be defined per candidate proportional to its size. This is particularly crucial for shape features as they achieve a high response once the object of interest is placed at the center of the ROI with sufficient background margin.

An initial patch size of $24 \times 24 \times 20$ voxels was defined around each centroid (R_{fix}) and then was up-sampled by a factor of 2 in every direction through linear interpolation (R_{up}). The multi-scale sphere detection that was proposed for pre-screening was performed again on the R_{up} to estimate the size of the central structure.

The estimated size corresponds to the variance of the maximum response σ_{opt} ([Supplementary Material Sec. 1.1](#)).

2.4.2. Radon transform

The Radon transform is capable of representing shape characteristics through an object's directional property. The 3D Radon transform, based on a parallel plane-integral approach, can effectively transform an image volume by accumulating voxel intensities on several planes where each is oriented along a given set of angles.

The Radon transform can be written mathematically as mapping an image $f(x, y, z)$ to a function $R(\rho, \theta, \phi)$. Let $\vec{n} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\phi)$, $\vec{p} = (x, y, z)^t$ a point on a 3D volume with its center as origin, and $\vec{p} \cdot \vec{n}$ the equation of a plane for a given angle $\alpha = (\theta, \phi)$. As such, the 3D Radon transform of an image volume f is defined as:

$$R(\rho, \alpha) = \sum_{-x}^{+x} \sum_{-y}^{+y} \sum_{-z}^{+z} f(x, y, z) \delta(\vec{p} \cdot \vec{n} - \rho) \quad (1)$$

where δ is the Dirac function and ρ is the distance between the volume center and perpendicular projection planes along the α ([Fig. 2c](#)). This projection is able to highlight properties of geometrical shapes by two parameters (ρ, α) , where $\alpha = (\theta, \phi)$ is the polar projection angle. The 3D Radon implementation maps every image coordinate (x, y, z) to the physical space by defining the center of the image as origin and including the image spacing. Retaining the generality of the Radon projection outputs, the results are organized in 2-dimensional fashion with axes of ρ and α . [Fig. 2a](#) and [b](#) shows the 3D Radon transform of a simulated vessel and CMB.

2.4.3. Radon shape features

In [26,27], a set of Radon-based features was proposed to detect Gaussian-like structures in a sliding window in the presence of noise or other structures. These features are the mean and standard deviation profiles computed across the Radon angle-dimension:

$$F_{mean}(x') = \frac{1}{q+r} \sum_{\alpha=1}^{q+r} R(x', \alpha), \quad 1 \leq x' \leq |\rho| \quad (2)$$

$$F_{std}(x') = \sqrt{\frac{1}{q+r-1} \sum_{\alpha=1}^{q+r} [R(x', \alpha) - F_{mean}(x')]^2} \quad (3)$$

where $|\rho|$ is the number of projection distances and considering image dimension is $\sqrt{x_{dim}^2 + y_{dim}^2 + z_{dim}^2}$, and q and r are the number of projection angles of θ and ϕ , respectively, with $(\theta, \phi) \in [0 : 15 : 165]$. In the current work, standard deviation along the Radon angle-dimension and the global mean are also computed:

$$f_{mean} = \frac{1}{q+r+|\rho|} \sum_{\alpha=1}^{q+r} \sum_{\rho=1}^{|\rho|} R(\rho, \alpha) \quad (4)$$

$$f_{std} = \sqrt{\frac{1}{q+r-1} \sum_{\alpha=1}^{q+r} \left[\frac{1}{|\rho|} \sum_{\rho=1}^{|\rho|} R(\rho, \alpha) - f_{mean} \right]^2} \quad (5)$$

The proposed Radon descriptors are invariant to rotation by nature, and scale invariance is obtained by resizing the descriptor profiles of F_{mean} and F_{std} to a fixed length. Translation invariance is enforced when the candidate microbleeds are placed at the center of the ROI prior to the Radon transform.

2.4.4. Hessian shape features

The Hessian matrix H is a 3×3 matrix which encompasses the second order derivatives. H encodes local high-level shape information and the eigenvalues provide quantitative measures of the

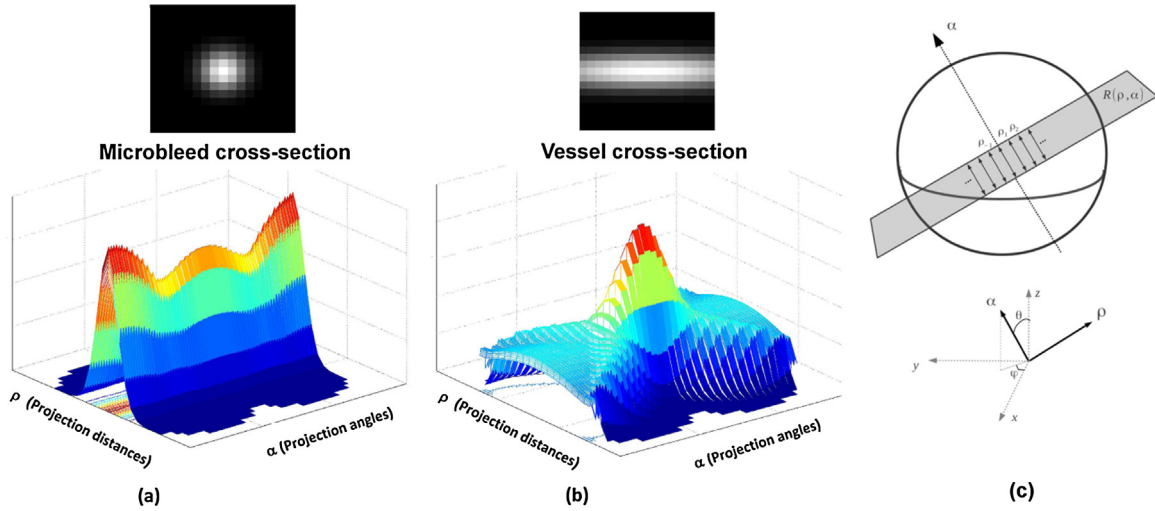


Fig. 2. (a) and (b) The top row shows a cross-section of a synthetic microbleed (left) and a vessel (right), and the bottom row shows normalized 3-dimensional Radon transform of the objects, and (c) the 3-dimensional Radon transform through parallel plane-integral approach is illustrated. ρ_i represents the distances on which the plane integrals are computed.

shape. To extract scale invariant features from H , the γ -normalized Hessian matrix \bar{H}_γ of the center of the ROI is considered:

$$\bar{H}_\gamma = \sigma_{opt}^\gamma \cdot \begin{bmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{xy} & I_{yy} & I_{yz} \\ I_{xz} & I_{yz} & I_{zz} \end{bmatrix} \quad (6)$$

where I_{xx} , I_{yy} , I_{zz} , I_{xy} , I_{xz} , and I_{yz} are the second order derivatives of image intensity convolved with Gaussian kernel of size σ_{opt} (Section 2.4.1). Here, γ is set to 2.5 for the 3D LoG.

The eigenvalues of the \bar{H}_γ were used individually as features, and three dissimilarity measures were also calculated considering $|\lambda_1| \leq |\lambda_2| \leq |\lambda_3|$: sphericalness (f_{sphere}), largest cross-section (f_{LC}) and fractional anisotropy (f_{FA}). These features are rotation, scale and shift invariant. Sphericalness, which is volume to largest cross-section area, attains a maximum of 1 for a perfect sphere-like structure and a minimum of 0 for other types of structures:

$$f_{sphere} = \frac{|\lambda_1|}{\sqrt{|\lambda_2\lambda_3|}} \quad (7)$$

The second feature accounts for the aspect ratio of the two largest eigenvalues. This ratio helps to further distinguish between CMBs and vessels where it attains a maximum of 1 for CMBs:

$$f_{LC} = \frac{|\lambda_2|}{|\lambda_3|} \quad (8)$$

And finally FA is a measure of how non-spherical an object is and for CMBs the value is relatively low:

$$f_{FA} = \sqrt{\frac{1}{2} \frac{\sqrt{(\lambda_1 - \lambda_2)^2 + (\lambda_2 - \lambda_3)^2 + (\lambda_1 - \lambda_3)^2}}{\sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}}} \quad (9)$$

Because of partial voluming and voxel anisotropy of SWI, CMBs can appear elongated along the z direction. Therefore, a structure orientation property (f_o), defined as the angle between the eigenvector associated with the largest eigenvalue and the xy -plane normal, was also computed as a descriptor to differentiate these elongated structures from vessels.

2.5. Step 3: Cascaded classification

The ability of the proposed Radon- and Hessian-based features for CMB detection was tested using a binary multi-layer classifier.

A binary classification scheme was considered by dividing the pre-screened candidates into non-CMB (negative) and CMB (positive) classes based on the available manual ratings. As a consequence of the pre-screening step which is characterized by a low specificity (Section 2.3), the negative class sample size is several times larger than the positive one (imbalanced training set), containing many irrelevant candidates without distinct geometric information. To overcome the machine learning bias toward the majority class, a binary random forest classifier (RF), which combines bagging and random sub-sampling techniques, was used [28]. However, the high presence of non-informative candidates in any random subset will adversely influence the RF performance. Therefore, we proposed a cascade of sequential RF classifiers that progressively rules out obvious candidates by imposing low to high posterior probability thresholds throughout the cascade (Fig. 3). In short, FP lesions are removed through each layer.

In the first layer, the classifier RF_1 was trained using all the pre-screened candidates (TS_1), and evaluated with a low probability threshold TH_1 . Any negative outcome leads to an immediate rejection while positively detected samples will compose a new training set TS_2 . This set will be subsequently split into CMB and non-CMB classes knowing the true CMBs provided by the manual ratings. As such, a similar procedure was repeated to build RF_2 with $TH_2 \geq TH_1$ for evaluation. More layers can be added until either the detection sensitivity or false-positive rate (FPR) become constant compared to the previous layer:

$$Sensitivity = \frac{TP}{TP + FN}, \quad FPR = \frac{FP}{FP + TN} \quad (10)$$

Consequently, the final layer is trained on a relatively balanced dataset leading to a high sensitivity and specificity (low FPs).

2.6. Experimental evaluation

Definite microbleeds have been found to have increased inter-rater agreement as they have a more distinct appearance compared to possible ones [29,9]. Therefore, for the current validation, two subsets comprised of *Definite* (D) and *Possible and Definite* (PD) CMBs were considered. It is acceptable to have a greater number of pre-screened candidates and a lower detection sensitivity when PD CMBs are examined due to the fact that *possible* ones are less hypointense and are not strictly spherical. As such, two different set

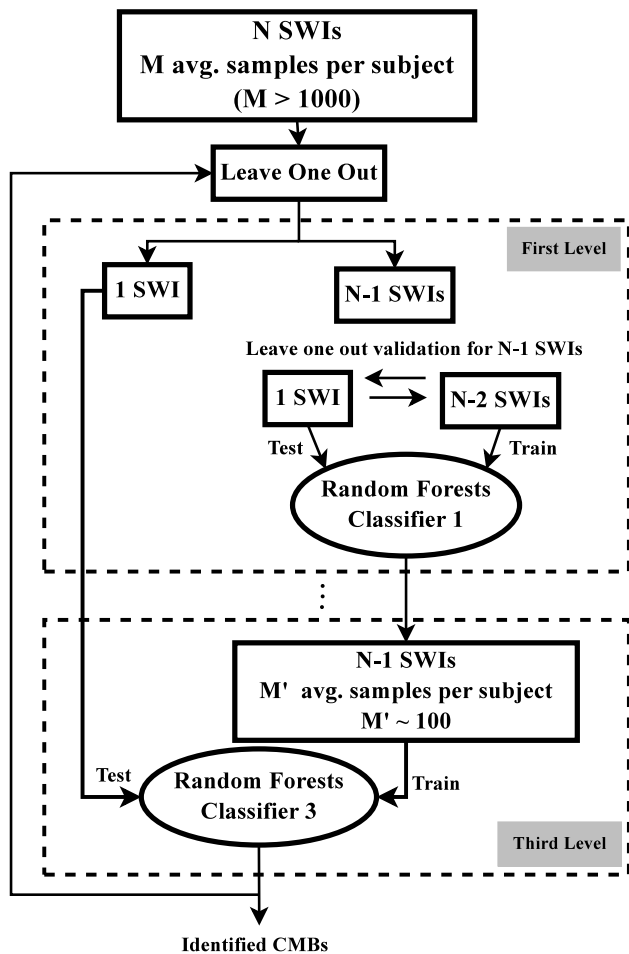


Fig. 3. A two-level classification pipeline which can be extended by adding more layers similar to the first level.

of parameter values were considered based on D and PD datasets for the automatic pre-screening and supervised detection steps.

Definite set In the pre-screening, the thresholds on minimum CMB intensity $I(c)$ and its sphericalness $S(c)$ were both empirically set to 0.55. The multi-scale LoG smoothing kernels were assigned to $[0.8:5.0]$ with step size of 0.1, assuming a maximum *definite* CMB diameter size of 10 mm. For the multi-layer classification, a three-layer cascade with low to high posterior probability thresholds of $Th_1 = 0.1$, $Th_2 = 0.3$, and $Th_3 = 0.4$ was considered. For this experiment, all the *possible* CMBs were excluded from the pre-screened candidate.

Possible and definite set The thresholds $I(c)$ and $S(c)$ were both empirically set at 0.35 and 0.5 to ensure CMBs with low contrast and elongated shape were captured. In the manual pre-screening, CMBs as small as one voxel were also considered *possible*. Therefore, the smoothing kernels were assigned to $[0.4:5.0]$ with step size of 0.1. For the classification, a three-layer cascade with probability thresholds of $Th_1 = 0.05$, $Th_2 = 0.2$, and $Th_3 = 0.4$ was considered.

These probabilities were chosen with enough margin to avoid over-fitting. The random forest was trained with 100 trees to a maximum depth of 5 restricting the number of samples per leaf node to be 5 samples. The validation for random forest was done

using a leave-one-out validation scheme throughout the cascade. The C++/OpenCV implementation of RF was used here.

3. Results

Using the established MARS rating scale [9] a total of 234 CMBs were identified by two clinical experts inspecting 66 subjects. Among them, 3 CMBs were excluded due to their irregular shapes (as shown in Fig. 4), 123 CMBs were labeled as “possible” with prevalence of 1.86 per subject, and 108 as “definite” with prevalence of 2.57 per subject.

The overall detection performance of the proposed method along with its pre-screening and classification results evaluated on both D and PD sets are listed in Table 2. The initial detection achieved a high sensitivity of 97% leaving just a few true CMBs behind. This stage, however, was less specific to true CMBs, and produced 669 and 1373 average candidates per subject of D and PD sets, respectively, leading to an imbalance class ratio ([average FP]/[average TP]) of 1 positive to 260 negative samples for D set, and a ratio of 1:406 for PD set. The full process evaluated on D set had a sensitivity of 92.6%, average FP of 10.28 CMBs per subject, and produced 4.3 FPs per true *definite* CMB (average FP/[average TP]). A corresponding evaluation on the PD set showed an overall sensitivity of 87%, average FP rate of 27.1 CMBs, and producing 8.9 FPs per true CMB. Despite the high specificity, it is not a suitable validation measure due to the high true-negative (TN) samples which correspond to the majority of the image volume (in pre-screening) and imbalanced CMB set (in classification).

The performance associated with each layer of the cascaded-classifier shows that each layer progressively obtained a lower FP rate for a given sensitivity as shown in Fig. 5, where sensitivity is expressed as a function of average FPs per subject.

The anatomical occurrences of FPs (yellow), TPs (blue) and FNs (red) CMBs detected by the proposed method on the *definite* set is illustrated in Fig. 6(a), and an example output with few TP and FP CMBs is shown in Fig. 6(b).

The average incidence of *definite* CMBs across AD, MCI and HC subjects shows that the detection rate of the proposed method is proportional to the number of manually identified CMBs. These results are shown in Table 3.

The efficiency and study characteristics of our technique compared to the previously published methods are summarized in Table 1, which indicate that the current method achieved a higher sensitivity and a lower false detection rate on a relatively larger cohort size.

Table 2

Table below summarizes the step-specific and overall CMB detection performance. Note that TP is presented as identified-TP/[total-TP]. The sensitivity and specificity are computed as TP/(TP + FN) and TN/(TN + FP), respectively.

Definite (42 subjects)				
	Manual reads	Pre-screening	Classification	Overall
TP	108	105/[108]	100/[105]	100/[108]
FN	–	3	5	8
AFPs	–	668.19	10.28	10.28
Sen.	–	97.2%	95.2%	92.6%
Spc.	–	–	99%	99%
Possible and definite (66 subjects)				
TP	231	223/[231]	201/[223]	201/[231]
FN	–	8	22	30
AFPs	–	1372.3	27.1	27.1
Sen.	–	96.5%	90.1%	87.0%
Spc.	–	–	98%	98%

TP, true-positive; FN, false-negative; AFPs, average false-positives per subject; Sen., sensitivity; Spc., specificity.

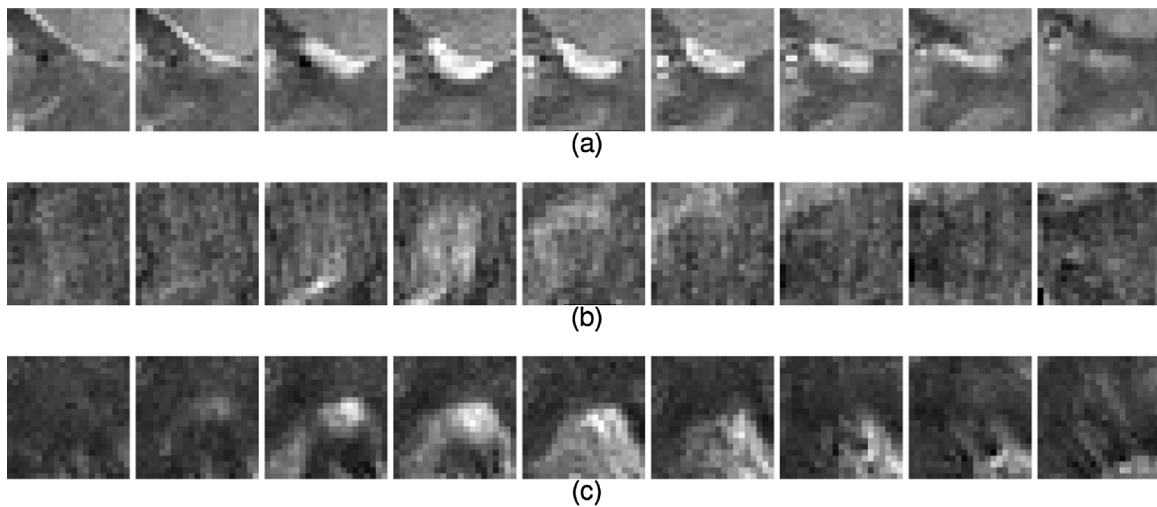


Fig. 4. Figures above illustrate the excluded cerebral microbleeds due to irregular shapes. Note that for visualization purposes, each 3-dimensional bounding box has been decomposed to nine in-plane slices and are shown side by side. The intensity has also been inverted so that CMBs appear bright.

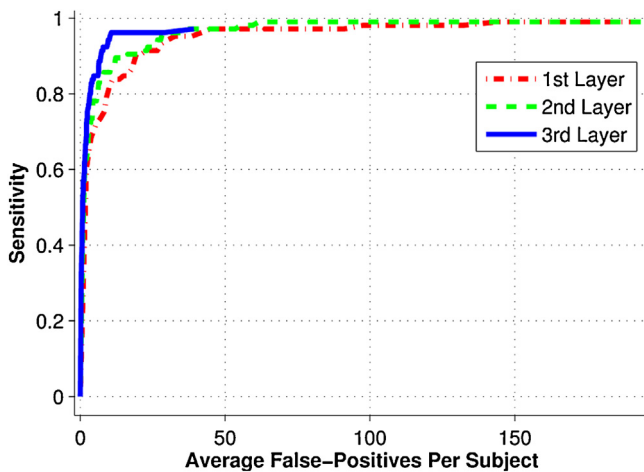


Fig. 5. The free-response receiver operating characteristic (FROC) curve comparing each layer's performance of the proposed three-layer classifier performed on *definite* CMB set.

4. Discussion

This paper presents a new computer-aided technique for CMB detection on SWI with the overall objective of achieving a high

Table 3

Group-wise average incidence of *definite* CMBs per subject.

Avg.	True CMBs	Falsely detected (FPs)	Total detected (FPs + TPs)
AD	3	14	16.7
MCI	2.2	8	9.9
HC	2.6	11.1	13.6

AD, Alzheimer's disease; MCI, mild cognitive impairment; HC, healthy control; CMB, cerebral microbleeds; FP, false-positive; TP, true-positive.

detection rate without increasing FPs to facilitate and speed up manual screening. The high sensitivity is also vital for clinical pre-screenings, where the aim is not to identify all CMBs, but to find enough of them to assess disease developments [30], to diagnose (CAA with ≥ 2 CMBs) [31], and to utilize as exclusion criteria (anti-amyloid therapies) [30].

In the present work, a pre-screening based on a Laplacian of Gaussian operator along with a multi-scale analysis shows a promising performance to mark potential CMB candidates and true CMBs. This low-level approach is not only computationally efficient, but it also reduces the overall computation time by limiting the high-level analysis to the selected candidates rather than all image voxels. The proposed pre-screening identified 97% of all CMBs in the PD set, and 98.5% of CMBs in the D set with just two

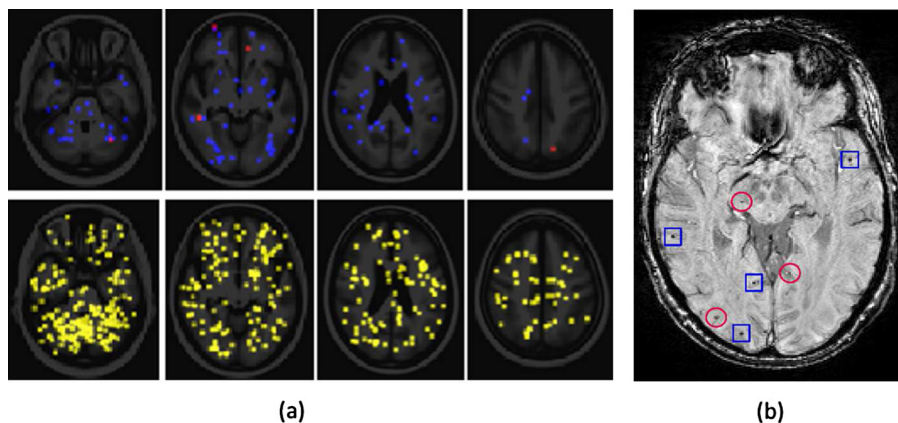


Fig. 6. (a) Minimum intensity projection of *definite* microbleeds (CMBs) on 4 axial cross-sections of brain. False-positive (yellow), true-positive (blue) and false-negative (red) CMBs of the proposed method are shown on the different regions; (b) an example of axial susceptibility-weighted image with overlay of correctly identified microbleeds (blue squares), and few false detections (red circle). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

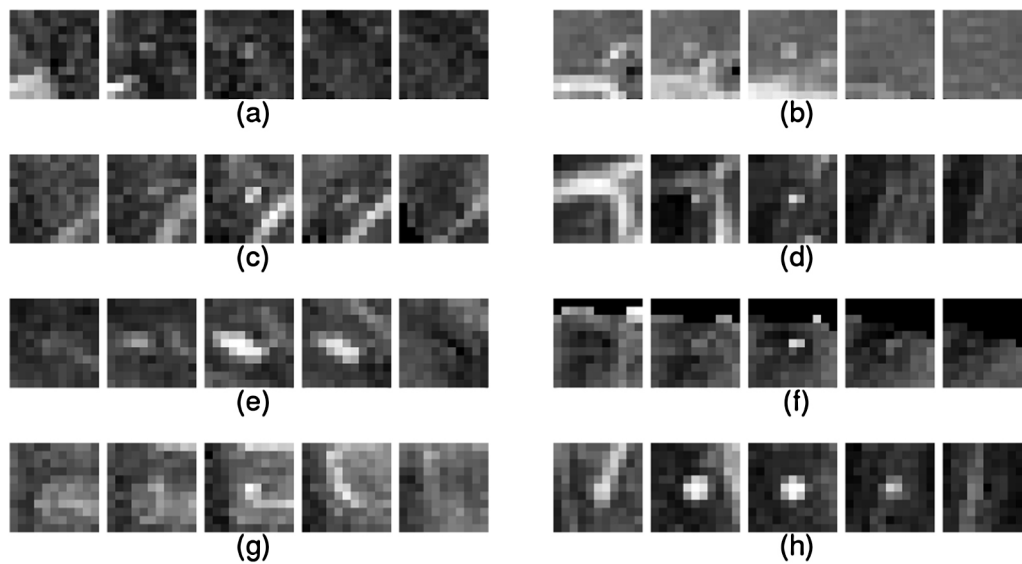


Fig. 7. Figures above show the miss-classified true microbleeds (CMBs) in automatic pre-screening (a–c), and the false-negatives from the classification step (d–h) validated on *definite* training set. (a–c) Have non-uniform and low contrast, (d, g and h) are connected to bottom and top structures, (f) is next to the skull and (e) has irregular shape. Note that for visualization purposes, each 3-dimensional bounding box has been decomposed to five in-plane slices and are shown side by side. The intensity has also been inverted so that CMBs appear bright.

parameters to control the lesion sphericalness and contrast. Such a high sensitivity was achieved by combining three multi-scale 1D line detection for detecting spherical lesions. In addition, the response normalization term has the advantage of detecting small CMBs adjacent to other structures, and allows a precise estimation of candidates extent which is essential for feature extraction.

On average, the pre-screening took less than 3 min, which can be further reduced by increasing the step size of the 1D Gaussian kernels (i.e. from 0.1 to 0.3).

Upon visual inspection of the pre-screening results, the undetected *definite* CMBs (3 out of 108 of D set) have a low contrast with varying intensity and all belong to the same subject (Fig. 7(a)–(c)). The missing true CMBs of the PD set include two low contrast *definite* CMBs, few *possible* CMBs with a size of one voxel with low contrast, and the rest are directly connected to neighboring structures from top or bottom.

On the other hand, the high sensitivity of the pre-screening is associated with a large number of FP CMBs, which leads to a class imbalance ratio of almost four-hundreds. The majority of pre-screened FPs associated with the discontinuous foldings of the sulcus and gyrus across the brain, and sharp local intensity changes in the sinus. Unlike other studies [19,32], a well-established RF classifier was chosen as an independent weak classifier to handle imbalanced datasets. This is mainly because the RF algorithm incorporates iterative learning [20], and bagging or boosting ensembles of classifiers [16] which have been shown to be effective in skewed classification problems. The proposed classification cascade focuses on gradual elimination of non-CMB like candidates to refine the training set, so that each classifier becomes progressively more accurate and has a lower FP rate. For instance, the first layer of D set which was assigned a threshold of 0.1 was able to remove more than 65% of candidates while missing only one true CMBs.

The computation time for training a three-layer cascade on 66 subjects was 3 h and the validation time was 2 min per subject on a single core machine with 8 GB of memory. This computation speed could be reduced by employing 10-fold leave-one-out cross validation.

Upon visual inspection of the classification outputs, a majority of the false-negative CMBs are connected to neighboring vessels, and two cases have irregular oval shapes as shown in Fig. 7(d)–(h). The

resulting FPs, appeared with a fairly uniform distribution throughout the brain except for the cerebellum and sinus regions (Fig. 6(a)). This can be explained by horizontal sulcus foldings of the cerebellum which appear as continuous profile on SWIs. Air channels and sinuses also cause susceptibility artifacts that could lead to higher false detection rate.

A similar trend in terms of prevalence between the manual (true) and automated (FPs) CMBs is found within AD, MCI and HC groups, where a higher true CMB count is associated with a higher automated detection (3 vs. 16.7 in AD, 2.6 vs. 13.6 in HC and 2.2 vs. 9.9 in MCI). This correlation shows the reliability of the proposed technique and that there is no performance bias in the presence of pathology. Further assessments are necessary to determine whether these detected lesions are true CMBs which were missed in the manual inspection.

The current method, achieved a higher sensitivity, a lower false detection rate and was validated on a larger number of subjects than other published methods (Table 1). Unlike [12] which excluded small CMBs < 5 mm, here CMBs as small as one voxel (0.93 mm) were included and classified as *possible*. This was achieved by imposing a parameter to control the minimum contrast in the pre-screening and extracting scale-invariant features.

The large cohort size used in this study offers CMBs with diversity in shape, size, anatomical location and, demonstrating the robustness of the algorithm. The current FP to TP rate on D set (10.28 on average) is much lower than PD (27.1 on average). This is important, as most clinical studies have focused on *definite* CMBs [10], as they have been shown to improve inter-rater reproducibility [9,29]. Moreover, this low FP rate provides an acceptable trade-off for the high sensitivity (92.6%) of the proposed method for studies in which exclusion (anti-amyloid therapies) or diagnosis (diagnose CAA with ≥ 2 CMBs) based on CMB incidence are of interest. From a practical standpoint, the proposed method is designed as an aid to manual labeling of CMBs, and visual inspection to remove approximately 11 FP CMBs would lead to a more efficient screening (reliable and reproducible) of SWI images, and reduce manual review time compared to the inspection of the entire image.

The current research has some limitations. The first is the lack of human expert assessment of the automatically identified CMBs. This assessment is necessary to determine whether the falsely

detected lesions are true CMBs which have been missed in the manual inspection. These extra true CMBs then could be added to the training set in order to improve the reliability of the automated detection. Second, the detection is based on a machine learning model trained on ground truth CMBs provided by local experts. The high inter-rater variability due to experience and training may cause a certain level of uncertainty in the final output. Third, the current method has been validated on a SWI sequence with specific TR, TE and software version. Therefore, a different SWI sequence or MRI modality such as T2* GRE would require a re-training of the algorithm.

In conclusion, this paper has presented a novel computer-aided approach to detect cerebral microbleeds in elderly subjects and patients with dementia using SWI. The proposed approach outperforms related published methods with a higher sensitivity and a lower false detection rate. It can assist manual screening by improving reliability, minimizing intra-rater variability and reducing assessment time. Future work will examine the stability of the method on longitudinal data as well as its robustness on other MRI sequences.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.compmedimag.2015.10.001>.

References

- [1] Greenberg S, Vernooij M, Cordonnier C, Viswanathan A, Al-Shahi Salman R, Warach S, et al. Cerebral microbleeds: a guide to detection and interpretation. *Lancet Neurol* 2009;8(2):165–74.
- [2] Yates P, Sirisriro R, Villemagne V, Farquharson S, Masters C, Rowe C, et al. Cerebral microhemorrhage and brain β -amyloid in aging and Alzheimer disease. *Neurology* 2011;77(1):48–54.
- [3] Fiehler J. Cerebral microbleeds: old leaks and new haemorrhages. *Int J Stroke* 2006;1(3):122–30.
- [4] Nakata-Kudo Y, Mizuno T, Yamada K, Shiga K, Yoshikawa K, Mori S, et al. Microbleeds in Alzheimer disease are more related to cerebral amyloid angiopathy than cerebrovascular disease. *Dement Geriatr Cogn Disord* 2006;22(1):8–14.
- [5] Kuijff HJ, Brundel M, de Bresser J, van Veluw SJ, Heringa SM, Viergever MA, et al. Semi-automated detection of cerebral microbleeds on 3.0 T MR images. *PLOS ONE* 2013;8(6):e66610.
- [6] Haacke E, Xu Y, Cheng Y, Reichenbach J. Susceptibility weighted imaging (SWI). *Magn Reson Med* 2004;52(3):612–8.
- [7] Tong KA, Ashwal S, Holshouser BA, Shutter LA, Herigault G, Haacke EM, et al. Hemorrhagic shearing lesions in children and adolescents with posttraumatic diffuse axonal injury: improved detection and initial results. *Radiology* 2003;227(2):332–9.
- [8] Nandigam R, Viswanathan A, Delgado P, Skehan M, Smith E, Rosand J, et al. MR imaging detection of cerebral microbleeds: effect of susceptibility-weighted imaging, section thickness, and field strength. *Am J Neuroradiol* 2009;30(2):338–43.
- [9] Gregoire S, Chaudhary U, Brown M, Yousry T, Kallis C, Jäger H, et al. The microbleed anatomical rating scale (MARS). *Neurology* 2009;73(21):1759–66.
- [10] Yates P, Sirisriro R, Villemagne V, Farquharson S, Masters C, Rowe C, et al. Incidence of cerebral microbleeds in preclinical Alzheimer disease. *Neurology* 2014;78(5):326–33.
- [11] Seghier M, Kolanko M, Leff A, Jäger H, Gregoire S, Werring D. Microbleed detection using automated segmentation (MIDAS): a new method applicable to standard clinical MR images. *PLoS ONE* 2011;6(3):e17547.
- [12] Barnes SR, Haacke EM, Ayaz M, Boikov AS, Kirsch W, Kido D. Semiautomated detection of cerebral microbleeds in magnetic resonance images. *Magn Reson Imaging* 2011;29(6):844–52.
- [13] Kuijff HJ, de Bresser J, Geerlings MI, Conijn M, Viergever MA, Biessels GJ, et al. Efficient detection of cerebral microbleeds on 7.0 T MR images using the radial symmetry transform. *Neuroimage* 2012;59(3):2266–73.
- [14] Bian W, Hess CP, Chang SM, Nelson SJ, Lupo JM. Computer-aided detection of radiation-induced cerebral microbleeds on susceptibility-weighted MR images. *NeuroImage Clin* 2013;2:282–90.
- [15] Guo X, Yin Y, Dong C, Yang G, Zhou G. On the class imbalance problem. In: Fourth international conference on natural computation (ICNC'08), vol. 4, 2008. IEEE; 2008. p. 192–201.
- [16] Galar M, Fernandez A, Barrenechea E, Bustince H, Herrera F. A review on ensembles for the class imbalance problem: bagging-, boosting-, and hybrid-based approaches. *IEEE Trans Syst Man Cybern C Appl Rev* 2012;42(4):463–84.
- [17] Frangi A, Niessen W, Vincken K, Viergever M. Multiscale vessel enhancement filtering. *Med Image Comput Comput Assist Interv MICCAI* 1998:130–7.
- [18] Liu J, White JM, Summers RM. Automated detection of blob structures by Hessian analysis and object scale. In: 17th IEEE international conference on image processing (ICIP). IEEE; 2010. p. 841–4.
- [19] Bria A, Karssemeijer N, Tortorella F. Learning from unbalanced data: a cascade-based approach for detecting clustered microcalcifications. *Med Image Anal* 2014;18(2):241–52.
- [20] El-Naqa I, Yang Y, Wernick MN, Galatsanos NP, Nishikawa RM. A support vector machine approach for detection of microcalcifications. *IEEE Trans Med Imaging* 2002;21(12):1552–63.
- [21] Ellis KA, Bush AI, Darby D, De Fazio D, Foster J, Hudson P, et al. The Australian Imaging, Biomarkers and Lifestyle (AIBL) study of aging: methodology and baseline characteristics of 1112 individuals recruited for a longitudinal study of Alzheimer's disease. *Int Psychogeriatr* 2009;21(04):672–87.
- [22] Cordonnier C, Salman R, Wardlaw J. Spontaneous brain microbleeds: systematic review, subgroup analyses and standards for study design and reporting. *Brain* 2007;130(8):1988–2003.
- [23] Tustison N, Avants B, Cook P, Zheng Y, Egan A, Yushkevich P, Gee J. N4ITK: improved N3 bias correction. *IEEE Trans Med Imaging* 2010;29(6):1310–20.
- [24] Van Leemput K, Maes F, Vandermeulen D, Suetens P. Automated model-based tissue classification of MR images of the brain. *IEEE Trans Med Imaging* 1999;18(10):897–908.
- [25] Fazlollahi A, Meriaudeau F, Villemagne V, Rowe C, Desmond P, Yates P, et al. Automatic detection of small spherical lesions using multiscale approach in 3D medical images. In: International conference on image processing (ICIP). IEEE; 2013. p. 1158–62.
- [26] Giancardo L, Meriaudeau F, Karnowski T, Li Y, Tobin K, Chaum E. Microaneurysm detection with radon transform-based classification on retina images. In: Annual international conference of the IEEE Engineering in Medicine and Biology Society, EMBC. IEEE; 2011. p. 5939–42.
- [27] Giancardo L, Meriaudeau F, Karnowski T, Tobin K, Lic Y, Chaum E. Microaneurysms detection with the radon cliff operator in retinal fundus images. *Proc SPIE* 2010;7623:76230U–1.
- [28] Breiman L. Random forests. *Mach Learn* 2001;45(1):5–32.
- [29] Cordonnier C, Potter GM, Jackson CA, Doubal F, Keir S, Sudlow CL, et al. Improving interrater agreement about brain microbleeds development of the Brain Observer MicroBleed Scale (BOMBS). *Stroke* 2009;40(1):94–9.
- [30] Sperling RA, Jack CR, Black SE, Frosch MP, Greenberg SM, Hyman BT, et al. Amyloid-related imaging abnormalities in amyloid-modifying therapeutic trials: recommendations from the Alzheimer's association research roundtable workgroup. *Alzheimer's Dement* 2011;7(4):367–85.
- [31] Knudsen KA, Rosand J, Karluk D, Greenberg SM. Clinical diagnosis of cerebral amyloid angiopathy: validation of the Boston criteria. *Neurology* 2001;56(4):537–9.
- [32] Marrocco C, Molinaro M, Delia C, Tortorella F. A computer-aided detection system for clustered microcalcifications. *Artif Intell Med* 2010;50(1):23–32.