




Deep-Learning-Based MRI Microbleeds Detection for Cerebral Small Vessel Disease on Quantitative Susceptibility Mapping

Peng Xia, PhD,¹ Edward S. Hui, PhD,²  Bryan J. Chua, MD,³ Fan Huang, PhD,¹
Zuojun Wang, PhD,¹ Huiqin Zhang, PhD,¹ Han Yu, PhD,¹ Kui Kai Lau, MRCP,^{3,4}
Henry K.F. Mak, MD, FRCR,¹  and Peng Cao, PhD^{1*} 

Background: Cerebral microbleeds (CMB) are indicators of severe cerebral small vessel disease (CSVD) that can be identified through hemosiderin-sensitive sequences in MRI. Specifically, quantitative susceptibility mapping (QSM) and deep learning were applied to detect CMBs in MRI.

Purpose: To automatically detect CMB on QSM, we proposed a two-stage deep learning pipeline.

Study Type: Retrospective.

Subjects: A total number of 1843 CMBs from 393 patients (69 ± 12) with cerebral small vessel disease were included in this study. Seventy-eight subjects (70 ± 13) were used as external testing.

Field Strength/Sequence: 3 T/QSM.

Assessment: The proposed pipeline consisted of two stages. In stage I, 2.5D fast radial symmetry transform (FRST) algorithm along with a one-layer convolutional network was used to identify CMB candidate regions in QSM images. In stage II, the V-Net was utilized to reduce false positives. The V-Net was trained using CMB and non CMB labels, which allowed for high-level feature extraction and differentiation between CMBs and CMB mimics like vessels. The location of CMB was assessed according to the microbleeds anatomical rating scale (MARS) system.

Statistical Tests: The sensitivity and positive predicative value (PPV) were reported to evaluate the performance of the model. The number of false positive per subject was presented.

Results: Our pipeline demonstrated high sensitivities of up to 94.9% at stage I and 93.5% at stage II. The overall sensitivity was 88.9%, and the false positive rate per subject was 2.87. With respect to MARS, sensitivities of above 85% were observed for nine different brain regions.

Data Conclusion: We have presented a deep learning pipeline for detecting CMB in the CSVD cohort, along with a semi-automated MARS scoring system using the proposed method. Our results demonstrated the successful application of deep learning for CMB detection on QSM and outperformed previous handcrafted methods.

Level of Evidence: 2

Technical Efficacy: Stage 2

J. MAGN. RESON. IMAGING 2023.

Cerebral small vessel disease (CSVD) is a common cause of stroke, cognitive impairment, and gait disturbances.^{1,2} MRI features of CSVD include recent small subcortical infarct, lacunes, perivascular space, white matter hyperintensities, and cerebral microbleeds.³ Patients with a high total CSVD burden have a higher risk of recurrent stroke.^{4,5}

View this article online at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1002/jmri.29198). DOI: 10.1002/jmri.29198

Received Sep 17, 2023, Accepted for publication Dec 5, 2023.

*Address reprint requests to: P.C., LG3, 5 Sassoon Road, Pok Fu Lam, Hong Kong, Hong Kong SAR, China. E-mail: caopeng1@hku.hk

From the ¹Department of Diagnostic Radiology, The University of Hong Kong, Hong Kong, China; ²Imaging and Interventional Radiology, The Chinese University of Hong Kong, Hong Kong, China; ³Division of Neurology, Department of Medicine, The University of Hong Kong, Hong Kong, China; and ⁴The State Key Laboratory of Brain and Cognitive Sciences, The University of Hong Kong, Hong Kong, China

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

Cerebral microbleeds (CMBs) are mainly caused by intravascular hemosiderin deposit.⁶ They appear as small, round foci on T2*-weighted MRI, susceptibility-weighted imaging (SWI),⁷ and quantitative susceptibility mapping (QSM).⁸ CMBs are frequently found in stroke patients⁹ and have been linked to executive dysfunction, memory decline, and other cognitive impairment.⁶ CMBs in the lobar regions are a possible biomarker for cerebral amyloid angiopathy (CAA).¹⁰ On the other hand, CMBs in the deep brain areas are thought to be caused by hypertension.³ Therefore, the number and location of CMBs can provide valuable diagnostic or prognostic information for CSVD,¹¹ which could be systematically evaluated using the microbleeds anatomical rating scale (MARS) system.¹²

Detecting CMBs manually on SWI or QSM images is a tedious and time-consuming process. Deep learning is an ideal alternative to manual detection. Several studies have demonstrated the feasibility of using deep-learning-based feature extraction, along with the selection of handcrafted features, such as intensity and shape of CMBs, for the detection of CMB.^{13–15} Seghier et al achieved 50% sensitivity using an automated segmentation method.¹⁴ While Kuijff et al achieved 71.2% sensitivity using radial symmetry transform (RST)¹³ on SWI. Bian et al further improved the sensitivity to 88.4% with 44.9 false positives per patient¹⁵ using blood vessel masks obtained from 2D fast radial symmetrical transform (FRST).¹⁶ Chen et al incorporated manual feature extraction, such as area and circularity properties, to extend Bian's work.¹⁷ Handcrafted features refer to features related to shape, size, volume, length–volume ratio feature, and fractional anisotropy feature. These features are focused on the geometry of the object. In contrast, the features learned by deep learning networks are different from the geometry features. They are referred to as latent features, and their learning mechanism is not easily explainable. Many research papers have discussed the latest developments in this area.^{18,19} Qi et al applied a deep convolutional neural network (CNN) to extract high-level latent feature for detecting CMB.¹⁸ However, the major drawback of these studies was that a manual check had to be performed on each slice due to the 2D FRST algorithm. To mitigate this limitation, a 3D deep CNN can be employed, which would enable automated CMB detection and eliminate the need for manual checking.¹⁹

Previous studies on detecting CMBs have used a two-stage approach.^{13–15,20} The first stage involved identifying CMB candidates, while the second stage aimed to reduce false positives. In stage II, various deep neural networks were used, such as V-Net²¹ and ResNet.²² However, the algorithm for stage I still needs improvement in terms of balancing sensitivity and false positive tolerance. In stage I, the traditional handcrafted feature selections might not be general enough for clinical applications, considering variations in clinical datasets such as slice thickness and in-plane resolution.

Hybrid approaches that combine FRST and deep learning methods^{18,23} have demonstrated potential for stage I, with relatively high sensitivity achieved at the expense of high false-positive rates. Another study employed 3D FRST for stage I, incorporating deliberately selected handcrafted features such as isolated pixel connections and fractional anisotropy (FA) analysis.²⁰ However, this method relies on numerous parameters, which might limit its generalizability. Automatic detection of CMBs using parameter-free and the deep-learning-based pipeline is thus valuable for routine clinical applications.

Considering that CMB was readily delineated on SWI, some studies used both SWI and QSM as inputs to deep learning models.²⁰ Liu et al²⁰ and Almasni et al²³ proved that QSM was the best choice for one channel input with high sensitivity and positive predictive value (PPV) on their local datasets. The superior performance may be attributable to several properties of CMB contrast in QSM.²⁴ QSM could enhance the contrast of brain tissue around hemorrhage.^{7,8} Blooming artifacts, typically observed on SWI, are substantially reduced on QSM. More importantly, QSM could differentiate paramagnetic CMBs from diamagnetic calcification, which is a prevalent CMB mimic.^{25,26}

This study aimed to develop a two-stage parameter-free and deep-learning-based pipeline to automatically detect CMB on QSM from a relatively large CSVD cohort (N = 393). The performance of our newly proposed pipeline was compared with existing methods, including the handcrafted features selection and the ResNet deep learning model.²²

Methods

Subjects and Data Acquisition

Approval from the institutional review board was obtained for this retrospective study. A waiver for informed consent was obtained as well. SWI images from a total of 393 patients were collected between 2008 and 2020 from HKU MRI Unit.⁴ Half of the data from the HKU-MRI cohort (from 2008 to 2014) was previously published.¹¹ The other half of the data were collected between 2015 and 2020. The data were randomly divided into training, validation and testing sets.

MRI was performed using a 3.0-T MRI (Philips Healthcare, Achieva) equipped with an 8-channel SENSE head coil. For QSM scan, parameters included 23 msec echo time, 28 msec repetition time, flip angle (FA) = 15°, field of view (FOV) = 225 × 225 × 135 mm³, matrix size = 704 × 704 × 135, reconstructed resolution = 0.33 × 0.33 × 1.0 mm³, and number of excitations (NEX) = 1.

Visual CMB Label

The numbers and locations of CMBs were first labeled on the SWI images. The CMBs were labeled according to the

MARS system, which divided the brain into three categories and 13 regions, namely the infratentorial (including brainstem and cerebellum), deep region (including basal ganglia, thalamus, internal capsule, external capsule, corpus callosum, deep and periventricular white matter), and lobar region (including frontal, parietal, temporal, occipital, and insula). The datasets were labeled by three raters. Firstly, a radiology resident with more than 6 years of experience labeled all the cases. Secondly, a radiology scientist with 5 years of experience labeled all the cases independently. Finally, an experienced neurologist labeled all cases for the third time and reviewed all the labels. Therefore, each case was labeled three times. The CMBs agreed upon by at least two raters were considered definite. In cases of discrepancy, the reference from the neurologist and radiology resident was considered. Datasets with differing judgments across three raters were excluded. Only definite true CMBs were used for training the deep neural network. After the location of CMB was obtained, the labels were subsequently used for creating the volumetric CMB mask using a 3D region growing algorithm.²⁷ The region growing algorithm compared the intensity value of neighboring pixels and expanded the segmentation if the intensity of pixels were similar.

Brain Edge Removal, FRST CMB Detection, and Frangi Vesselness

Around the brain edges, blood vessels can resemble CMBs, leading to a large number of false positive detections. In this study, we performed a 3D erosion on QSM images to reduce the CMB mimics near the edges of the brain. The edge erosion for 5 pixels distance reduced 13% of the undesired CMB candidates, and the true candidates were not missed.

FRST could detect circular patterns in images. In this study, we adapted the FRST algorithm from the open-source MATLAB script.²⁸ To identify cerebral microbleeds (CMBs), we look for small areas of hypointensity or hyperintensity with a blooming effect on T2* weighted image or QSM. The FRST algorithm is an effective tool for this purpose. However, to adapt it specifically for identifying CMBs, we have extended the traditional FRST to a new version called the 2.5D FRST. In our study, we made some modifications to the image gradient (g_x , g_y , g_z) by rescaling it according to the image resolution along the three axes (res_x , res_y , res_z). This resulted in a rescaled gradient ($res_x \times g_x$, $res_y \times g_y$, $res_z \times g_z$). We then multiplied the 2D FRST results from three planes (x - y , y - z , x - z planes) to produce the 2.5D FRST. Both the 2.5D FRST and 3D FRST methods were tested on our patient data, and we used the 2.5D FRST algorithm in stage I of our pipeline. Next, we cropped the results obtained from the 2.5D FRST algorithm into a $7 \times 7 \times 7$ size and thresholded them at 0.5. These cropped FRST cubics were then passed through a convolutional layer to classify CMBs or vessels, as shown in Fig. 1.

Following this classification step, the Frangi vesselness filter²⁹ was applied on our cropped QSM images. This filter is intended to capture the vessel-like shape within a specified region. The MATLAB script for performing the Frangi vesselness filter was obtained from Ref. 27. After detecting potential cerebral microbleeds (CMBs) using a 2.5D FRST, we applied a convolutional layer to further reduce the number of false positives. Only those with positive labels classified by the convolutional layer were used as inputs to the V-Net in stage II. To obtain a CMB candidate mask, we applied 2.5D FRST on the quantitative susceptibility mapping (QSM) map. The candidate mask contained both true CMBs and false positives, which were subsequently reduced after applying stage II.

QSM Reconstruction and Segmentation

QSM reconstruction converted the raw phase images to tissue susceptibility maps using the following steps. Firstly, the brain masks were generated using Brain Extraction Tool (BET) from FSL.³⁰ Secondly, the raw phase was unwrapped with Laplacian-based algorithm,³¹ and the residual background field was removed using Variable-kernel Sophisticated Harmonic Artifact Reduction for Phase data (V-SHARP) algorithm.³² In the end, susceptibility maps were created using the algorithm of streaking artifact reduction for QSM (STAR-QSM).³³ The FSL toolbox (<https://fsl.fmrib.ox.ac.uk/fsl/fslwiki>) was used to perform the segmentation of brain regions. The QSM template³⁴ was first registered to the individual QSM images. Then the brain atlas³⁴ was transformed to the image space of the same subject by applying the transformation matrix.

V-Net and ResNet Models

As depicted in Fig. 1, the V-Net encoder employed each convolutional block with a residual or bypass connection to apply $5 \times 5 \times 5$ zero-padded convolutions with $1 \times 1 \times 1$ stride w . For downsampling, a $2 \times 2 \times 2$ convolution with $2 \times 2 \times 2$ stride was utilized as a pooling operation. The network architecture consisted of four levels with varying numbers of convolutional layers: 1, 2, 3, and 3, starting from top to bottom. The first level had eight feature channels for its convolutions. On each lower level, the number of channels for convolutions doubled after each pooling. The Rectified Linear Unit (ReLU) was applied after each convolutional layer. On the decoder side (expansive path), a $5 \times 5 \times 5$ deconvolution was employed for up-sampling followed by a concatenation operation that combined the feature map on the same level from the encoder side and the upsampled feature map from the lower level. The convolutional block with the residual connection was identical to the one on the same level on the encoder side. The final layer was a 1×1 convolution that mapped each 128-channel feature vector to the segmentation output, as shown in Fig. 1. Additionally, a max-pooling layer was applied to the segmentation output of

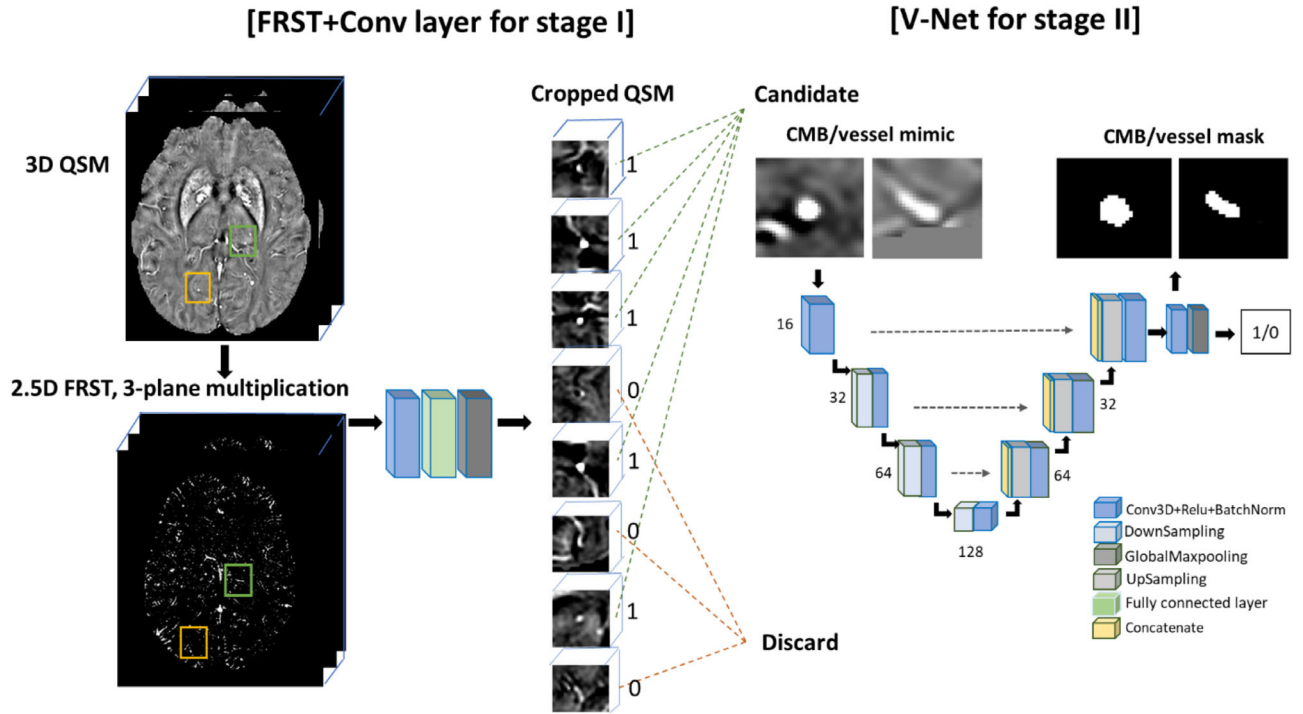


FIGURE 1: The flow chart of our two-stage CMB detection pipeline. On the left side, 3D features of QSM were extracted by the 2.5D FRST. The yellow rectangle denoted the missing one CMB by FRST. The green indicated the detected CMB. After the 2.5D FRST + convolution layer at stage I, candidates (with FPs remaining) were fed into the V-Net for stage II. The V-Net model was trained with CMB labels and non CMB labels. The probability value was from the global max-pooling of the segmentation outputs from the final output.

the V-Net model, creating CMB/vessel classification. In total, the network had 18 convolutional layers. The number of total parameters is 611,265, among them, 610,097 are trainable parameters, and 1168 are non-trainable parameters. The learning rate is 0.0001.

Our ResNet architecture utilized a series of blocks, each comprising of a $5 \times 5 \times 5$ convolutional layer with zero-padding and a residual or by-pass connection. For down-sampling, we employed $2 \times 2 \times 2$ convolutional layers with a $2 \times 2 \times 2$ stride as a pooling operation. Through this approach, we were able to incorporate a total of nine convolutional layers in the ResNet, culminating in a classification output. The number of total parameters is 452,529, among them, 154,649 are trainable parameters, and 880 are non-trainable parameters. The learning rate is 0.0001.

The loss functions used for training the V-Net and ResNet were as follows: we used weighted cross-entropy for the classification output, and for the segmentation output, we used a combination of weighted cross-entropy and Dice loss. We trained the models using a batch size of 100 and over the course of 150 epochs.

Data Augmentation Using CycleGAN

To address the challenge of small sample size, specifically the lack of low-intensity and small-sized CMB in the training set,

we utilized a CycleGAN model for data augmentation.³⁵ Figure 2 demonstrates that we utilized the domain A inputs for the generator from A to B (G_{A2B}), which represents an actual CMB QSM patch. On the other hand, the domain B inputs for the generator from B to A (G_{B2A}) were the concatenated normal QSM patch and target CMB mask. The target CMB mask indicated the shape of CMB. G_{A2B} generated synthetic normal QSM patches and the CMB masks, while G_{B2A} produced synthetic CMB QSM patches. To ensure that these synthetic CMB QSM patches (i.e., fake A), normal tissue patch, and CMB mask (i.e., fake B) are similar to real A and real B, two discriminators (D_A and D_B) were used.

The encoder-decoder structure was used to implement G_{A2B} and G_{B2A} , with the number of features increasing from 24 to 384 for each feature extraction level. D_A and D_B were adapted from patchGAN and employed five convolution layers.³⁶ The generator losses were computed using the mean square error, except for the losses between the real A and synthetic B, and between real B and synthetic A1, where L1-norm (also known as the identity loss) was used. The L1-norm encouraged the similarity between domains A/B and prevented the reversal of contrast in the synthetic A and B. The discriminators were trained using the cross-entropy loss, over 200 epochs. We used the CycleGAN model

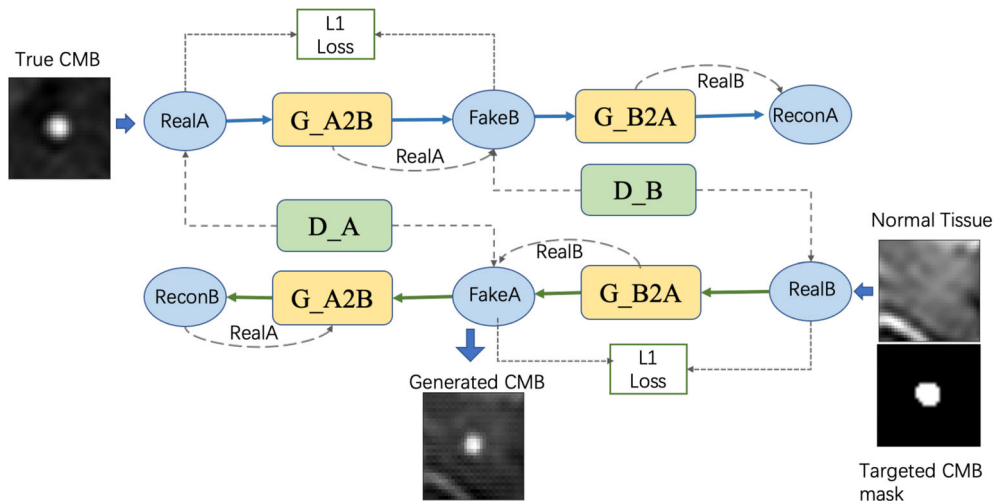


FIGURE 2: The flowchart of CycleGAN model for data augmentation. The RealA input was a true CMB patch on QSM. The RealB input was a normal brain tissue patch without CMB and a targeted CMB mask. We added regression loss and L1 loss to restrain the learning outcome. The generated CMB and normal tissue QSM patches had similar background anatomical features, except the CMB was created based on the input from the target CMB mask. ReconA = reconstructed A; ReconB = reconstructed B; G_A2B = generator A to B; G_B2A = generator B to A; D_A = discriminator A; D_B = discriminator B.

to synthesize $N = 1258$ CMBs. The normal tissue QSM patches were randomly selected from 50 QSM datasets from volunteers without CMB in the brain, while avoiding vessel pixels in the center of the patches. All data were partitioned randomly into training, validation, and testing sets. We implemented all neural networks using TensorFlow v2.2.0 (<https://www.tensorflow.org/>) and Keras v2.4.3. Training and testing were performed on a GPU server with four NVIDIA 2080Ti GPUs, 32 GB RAM, and an 8-core Intel Xeon CPU. It took approximately 13 hours to train a V-Net model on one GPU.

Statistical Analysis

We evaluated the effectiveness of various methods and the proposed deep learning pipeline based on their detection sensitivity, which was the ratio of correctly identified CMBs to all labeled CMBs, and PPV, which was the proportion of correctly identified CMBs out of all the detected CMB labels. Additionally, we calculated the number of false positives per subject. All statistical analyses were performed using IBM SPSS Statistics for Macintosh (Version 27, Armonk, NY: IBM Corp, USA).

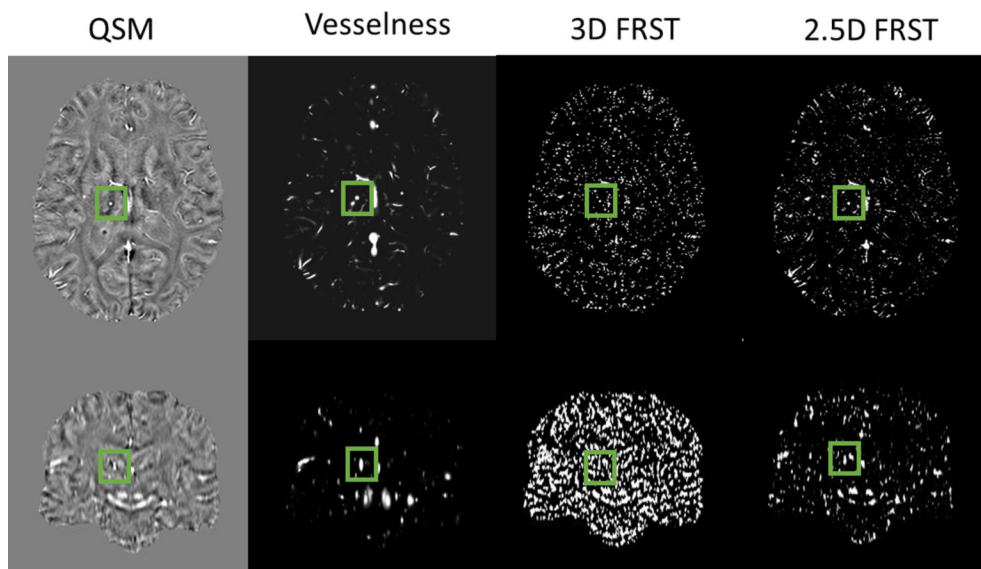


FIGURE 3: CMB candidate detection (at stage I) using Frangi vesselness and 3D and 2.5D FRST. Plots from left to right are QSM, vesselness map, 3D FRST, and 2.5D FRST. The green box indicates a true CMB. The first row is the axial view, and the second row is the coronal view.

Results

The flowchart of our CMB detection pipeline was shown in Fig. 1. After detecting potential cerebral microbleeds (CMBs) using the 2.5D FRST, we applied a convolutional layer to further reduce the number of false positives. Only those with positive labels classified by the convolutional layer were used as inputs to the V-Net in stage II. The model structure of CycleGAN used for data augmentation was illustrated in Fig. 2. The RealA input was a true CMB patch on QSM. The RealB input was a normal brain tissue patch without CMB and a targeted CMB mask. We added regression loss and L1 loss to restrain the learning outcome. Those synthetic CMBs were

used in training the V-Net in stage II in Fig. 1. CMB candidate detection (at stage I) using Frangi vesselness, 3D and 2.5D FRST was presented in Fig. 3. The true CMBs (annotated by green rectangles) were hyperintense on the QSM map (Fig. 3). The 2.5D FRST generally provided a lower false-positive rate compared with that of 3D FRST. The result of synthesized CMB using the CycleGAN model was shown in Fig. 4. The output patch was a combination of true CMB and normal brain tissue with different anatomical features and QSM intensities.

Figure 5 illustrated the performance of detecting true CMB and CMB mimics. The output with a probability of

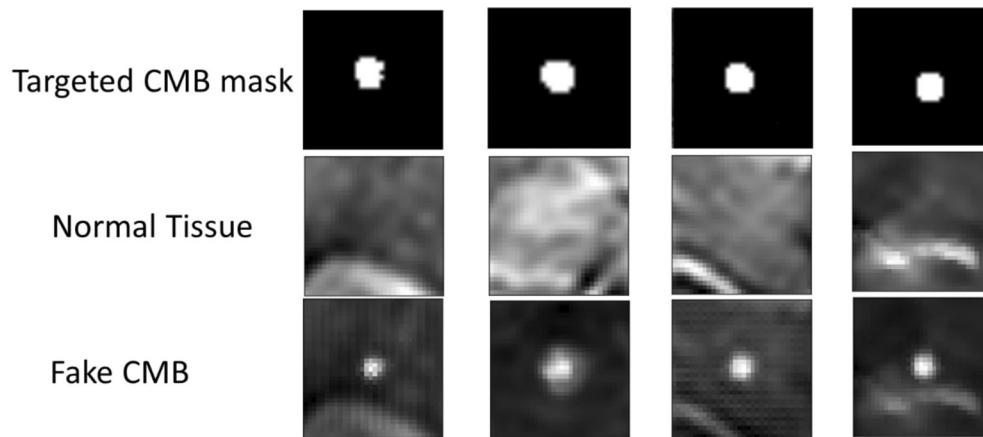


FIGURE 4: CycleGAN model results. The first row is real CMBs, the targeted mask. The second row is the normal tissue patches from 50 QSM datasets without CMBs. After inputting the targeted CMB masks and normal tissue patches, the model generated the corresponding patches with CMBs (the third row).

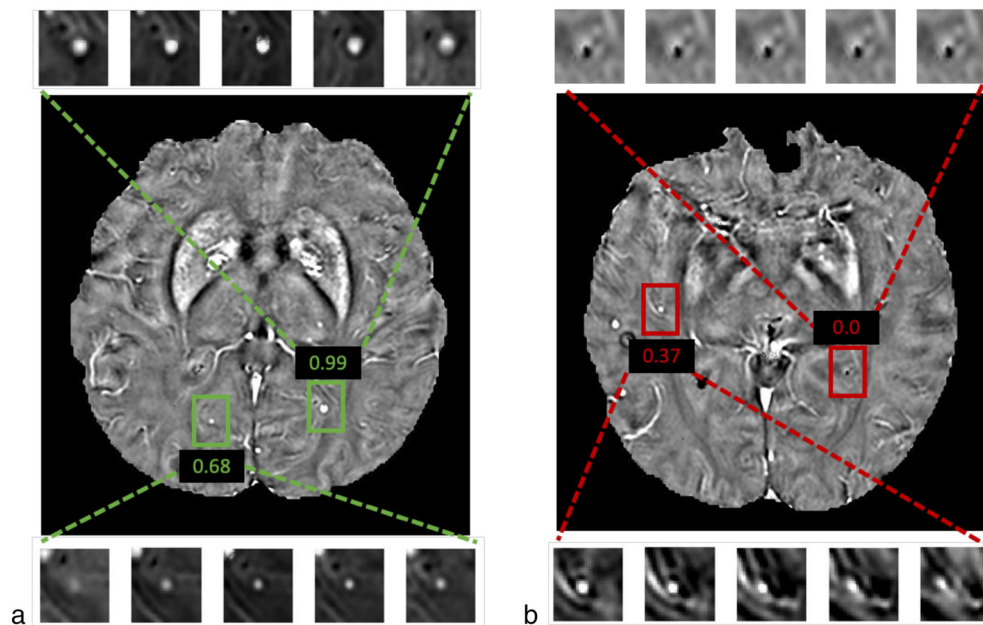


FIGURE 5: The results from stage II in our pipeline. (a) The probability of 0.68 and 0.99 indicated a true CMB. (b) A probability of 0.37 showed a mimic. A calcification mimic (b, top) was distinguished by the network as well. The mosaic plots were the cropped QSM views from different slices. Calcification mimics were discarded at stage I.

0.99 detected a true CMB. On the other hand, the prediction of the probability of 0.37 found instead a vessel-like CMB mimic that appeared to be round and hyperintense on the QSM map. A calcification mimic was distinguished by the network as well. The output from the network also avoided the nearby vessel extension.

The ROC curves of stage I and stage II are presented in Fig. 6. For stage I, using one convolution layer for selecting CMB candidates, the area under the curve (AUC) was 0.94. The convolutional layer reduced the number of false positives from 5266 to 3506, facilitating the V-Net-based stage II. For stage II, V-Net and ResNet were tested and compared. The ROC-AUC score was 0.93 for V-Net

and 0.89 for ResNet. To achieve high sensitivity and a small FP avg, we chose the thresholds as indicated by circles in Fig. 7. For example, the FP avg for 2.5D was 30. Meanwhile, for the same sensitivity level using 3D FRST, the FP avg was over 400.

The average age was 69 ± 12 years in mean \pm SD as shown in Table 1. The result of our two-stage deep-learning-based pipeline on a test dataset that consisted of $N = 78$ CSVD patients was shown in Table 2. Our pipeline successfully detected 290 CMBs out of the 326 candidates from the testing set. The detection sensitivity was 88.9%, with an average of 3.7 CMBs detected per subject and 2.87 false positives per subject. The training model missed 6.1%

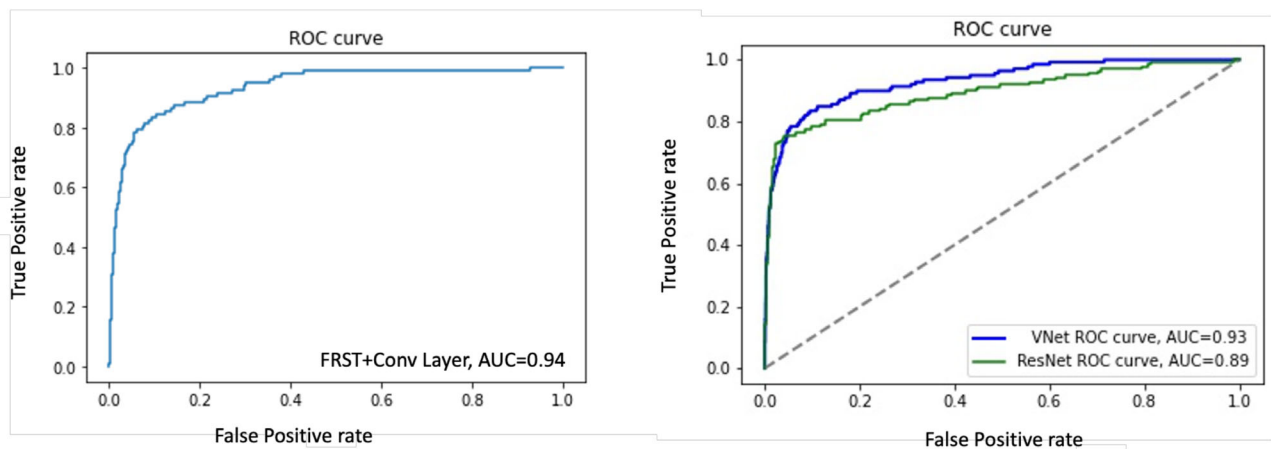


FIGURE 6: The ROC curves of stage I (left part) and stage II (right part). For stage I, using one convolution layer for prediction, the ROC-AUC was 0.94. For stage II, V-Net and ResNet were tested and compared. The ROC-AUC score was 0.93 for V-Net and 0.89 for ResNet.

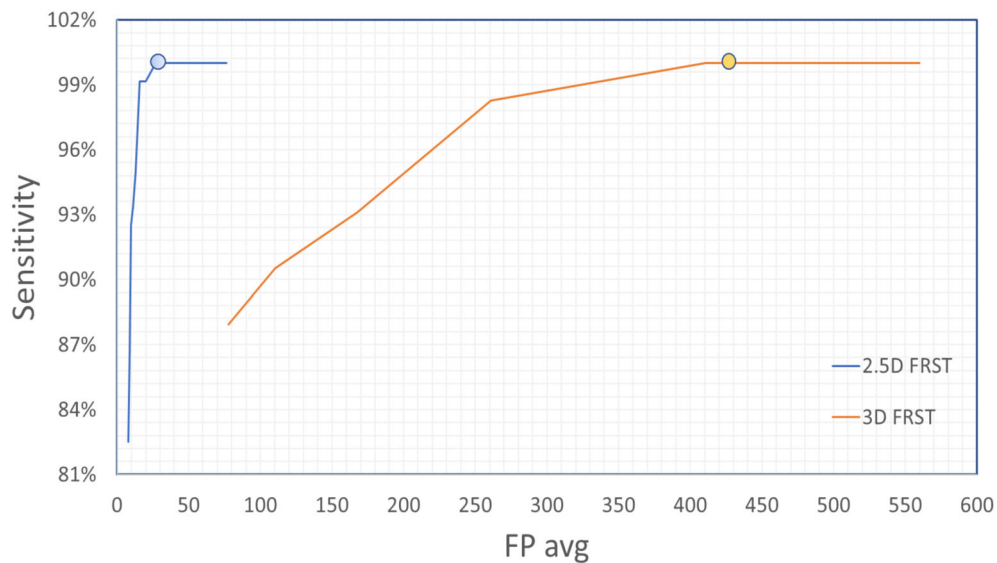


FIGURE 7: Sensitivity and averaged false positive (FP avg) plot for 2.5D FRST and 3D FRST. To achieve high sensitivity and a small FP avg, we chose the thresholds as indicated by circles. For example, the FP avg for 2.5D was 30. Meanwhile, for the same sensitivity level using 3D FRST, the FP avg was over 400.

(20 out of 326) true CMBs. The performance of the handcrafted features selection is shown in the middle column in Table 2 for comparison. First, using handcrafted features selection with volume threshold, 319 true CMBs were detected. After applying a length threshold, 285 true CMBs were detected. Finally, a total of 259 true CMBs remained with the use of the FA threshold. The final sensitivity for handcrafted features selection was 79.4%, with an average of 3.3 CMBs detected per subject and 20.4 false positives per subject.

The confusion matrix of V-Net and ResNet on the testing set was presented in Table 3. For the V-Net model, 276 true CMBs were detected, with 279 false positives. In contrast, the ResNet model detected 253 true CMBs and 390 false positives. After applying the V-Net model with

CycleGAN for data augmentation, 290 true CMBs were detected, with 224 remaining false positives. The sensitivities for ResNet, V-Net, and V-Net + CycleGAN were 81.6% to 89.0% and 93.5%, respectively, while their PPVs were 39.3% to 49.7% and 56.4%, respectively. After combining synthesized data from CycleGAN, the training sensitivity increased 4.5% compared with V-Net only.

The comparison between this study and other published study was illustrated in Table 4. As to the prospect of sequence, QSM was applied in this study. Compared with the handcraft feature detection in early years, the deep learning based methods achieved higher performance. Chen et al achieved a high PPV of 71.9% in 2019. However, 7 T MRI was used in their study which was not common in clinical scenario. Liu et al applied a two-channel input of phase image plus SWI image and achieved a high sensitivity of 95.8%. Our proposed study required one-channel input only, and the problem of calcification was well issued.

As summarized in the MARS table, our work also provided the CMB regional distribution. The distribution of CMBs from the testing set was shown in Table 5. According to MARS, distribution was estimated and divided into three major regions: the infratentorial, deep brain, and lobar regions. Across all the brain regions, more than 80% accuracy was obtained. The highest sensitivity was 89.3% for the temporal lobe region (except for the corpus callosum region where no CMBs were detected) and the lowest sensitivity

TABLE 1. Subjects Information and Data Splitting

	Subjects	Age (Mean)	Male	CMB
Training	280	68 ± 11	111 (39.6%)	1322
Validation	35	67 ± 9	13 (37.1%)	195
Test	78	70 ± 13	33 (42.3%)	326
Total	393	69 ± 12	157 (39.9%)	1843

TABLE 2. Stage I: Sensitivity of CMB Detection Using One Convolution (Conv) Layer Network and Handcrafted Features Selection on the Testing Set (N = 78)

	Total	Fast Radial Symmetry Transform	Network	Handcraft Feature		
			Conv	Volume	Length	Fractional Anisotropy
TP	326	321	310	319	285	259
Sensitivity	100%	98.4%	95.1%	97.8%	87.4%	79.4%
TP + FP		5266	3506	4027	3251	1590

TP = true positive; FP = false positive.

TABLE 3. Stage II: Confusion Matrices of V-Net and ResNet on the Testing Set (N = 78)

	TP	FN	FP	TN	Sensitivity	PPV
V-Net	276	34	279	1153	89.0%	49.7%
V-Net + CycleGAN	290	20	224	1208	93.5%	56.4%
ResNet	253	57	390	1042	81.6%	39.3%

TP = true positive; FP = false positive; FN = false negative; TN = true negative; PPV = positive predictive value.

TABLE 4. Comparison Between Different Studies

Study	Patients	Sequence	Methods	Sensitivity (%)	PPV
Kuijf et al ¹³	18	SWI	3D RST	71.2	Not reported
Bian et al ¹⁵	15	SWI	2D FRST	86.5	Not reported
Qi et al ¹⁸	320	SWI	3D FCN + CNN	93.16	44.31%
Chen et al ¹⁷	73	7 T SWI	3D ResNet	94.7	71.9%
Liu et al ²⁰	220	3 T SWI	3D FRST + ResNet 2 channel Phase + SWI input	95.8	70.9%
This study	393	3 T QSM	2.5D FRST + V-Net 1 channel QSM Input	93.5	56.2%

TABLE 5. MARS Table (N = 78) Presented the CMB Number Per Region

		True	Detected	Sensitivity (%)
Infratentorial	Brainstem	20	15	75.0
	Cerebellum	22	18	81.2
Deep brain	Basal ganglia	24	21	87.5
	Thalamus	61	53	86.9
	Internal capsule	2	1	50
	External capsule	5	3	30.0
	Corpus callosum	0	0	100
	Periventricular/deep WM	31	27	87.1
Lobes	Frontal	44	39	88.6
	Parietal	21	18	85.8
	Temporal	75	67	89.3
	Occipital	17	15	88.2
	Insula	4	4	100
Total		326	281	86.2

was 30% for the external capsule region. More than 85% sensitivity was achieved across nine brain regions. For the lobar regions, the detecting accuracy was considerably higher than the other two categories.

Discussion

Deep Neural Networks and Proposed Pipeline

This study developed a two-stage deep-learning-based pipeline to detect CMB on quantitative susceptibility mapping (QSM). Our proposed pipeline achieved sensitivities up to 94.9% at stage I and 93.5% at stage II. The overall sensitivity was 88.9%, and the number of false positives per subject was

2.87. For MARS, the sensitivities were above 85% for nine brain regions. Our study applied 2.5D FRST and a convolution layer to detect the CMB candidates and V-Net to reduce the number of false positives. QSM was used as a single channel input for simplicity. We compared the performance of V-Net and ResNet using our local CSVD cohort (N = 393). The CycleGAN synthesized small size, and low-intensity CMB mimics to augment the training data. V-Net along was difficult to recognize those small sizes and low-intensity CMB due to their small percentage of total data distribution. Our research work proved the possibility of using synthesized data to address the problem of limited availability on the task of CMB detection.

2.5D FRST vs. 3D FRST

Meanwhile, compared with the 3D FRST and handcrafted features selection, the 2.5D FRST balanced high accuracy and low false-positive numbers. For stage II, both V-Net and ResNet performed better than the traditional CMB detection algorithms in the literature. Therefore, our pipeline exceeds the traditional methods, like the handcrafted features selection.²² FRST algorithm remained a suitable choice for the task of CMB detection in recent research work. Our results suggested that 3D FRST could produce too many false positives and some modification based on FRST could help reduce the number of unwanted false positives. When reporting the correctly detected true CMB counts, we suggested reporting the distribution of CMB (including the location in brain regions), which could be a crucial factor considered during clinical CMB diagnosis.³⁷

Calcification Issue and MARS Scoring

The proposed two-stage pipeline to detect CMB on QSM provided more accurate quantification of CMB than SWI. More importantly, QSM could differentiate paramagnetic CMBs from diamagnetic calcification. Calcifications are a prevalent CMB mimic which could be missed by using SWI only.²³ This study provided the location distribution of detected CMBs in the MARS table. The statistics showed our proposed pipeline achieved more than 85% sensitivity across different brain regions divided by MARS, especially with sensitivity above 85% for the lobar regions. Therefore, the output of CMB distribution could facilitate the measurement of MARS scores with potential diagnosis values.

Comparison with Other Studies

There are research groups who successfully demonstrated the automated or semi-automated CMB detection on hemodialysis, traumatic brain injury, stroke, and brain tumor radiation therapy cohorts.^{17,20} Our study's novelty lies in the fact that we specifically focused on the CSVD group. Our findings provide a reliable reference for the CSVD group. Notably, our study had a large sample size of 393 participants, which is higher than other studies. Detecting CMB is crucial in assessing CSVD. CMBs act as an essential diagnostic marker for CSVD and help evaluate the disease's severity based on their location and number. This information can positively impact treatment decisions for individuals with CSVD.

Limitations

While our proposed method was effective in characterizing the burden of CMBs, this study had a few limitations. Due to the sparse distribution of CMBs in CSVD, it was challenging to improve the sensitivity and PPV of detection, especially considering the imbalance between the numbers of CMBs and vessel mimics. The traditional method of extracting handcrafted features resulted in a high number of false

positives. In particular, the sample size of the training data was significant, with over 100 CMB candidates produced for each case after preprocessing. This led to a long training time and a class imbalance issue due to the large number of false positives generated during the preprocessing stage. Finally, the process of labeling CMBs for the creation of training and testing datasets was time-consuming and labor-intensive.

Conclusion

We demonstrated the feasibility of a deep learning pipeline for the CMB detection on QSM CSVD cohort, CycleGAN-based data augmentation, and a semi-automated MARS scoring system using the proposed method.

Acknowledgments

This work was supported by Hong Kong Health and Medical Research Fund (grant numbers 07182706 and 06172916).

References

- Pantoni LMD. Cerebral small vessel disease: From pathogenesis and clinical characteristics to therapeutic challenges. *Lancet Neurol* 2010; 9(7):689-701. [https://doi.org/10.1016/S1474-4422\(10\)70104-6](https://doi.org/10.1016/S1474-4422(10)70104-6).
- Ambler G, Lee K-J, Shiozawa M, Lovelock C, et al. Cerebral microbleeds and stroke risk after ischaemic stroke or transient ischaemic attack: A pooled analysis of individual patient data from cohort studies. *Lancet Neurol* 2019;18(7):653-665. [https://doi.org/10.1016/S1474-4422\(19\)30197-8](https://doi.org/10.1016/S1474-4422(19)30197-8).
- Wardlaw JM, Smith EE, Biessels GJ, et al. Neuroimaging standards for research into small vessel disease and its contribution to ageing and neurodegeneration. *Lancet Neurol* 2013;12(8):822-838. [https://doi.org/10.1016/S1474-4422\(13\)70124-8](https://doi.org/10.1016/S1474-4422(13)70124-8).
- Lau KK, Li L, Schulz U, et al. Total small vessel disease score and risk of recurrent stroke: Validation in 2 large cohorts. *Neurology* 2017;88(24):2260-2267. <https://doi.org/10.1212/WNL.0000000000004042>.
- Ambler G, Lim J-S, Shiozawa M, et al. Development of imaging-based risk scores for prediction of intracranial haemorrhage and ischaemic stroke in patients taking antithrombotic therapy after ischaemic stroke or transient ischaemic attack: A pooled analysis of individual patient data from cohort studies. *Lancet Neurol* 2021;20(4):294-303. [https://doi.org/10.1016/S1474-4422\(21\)00024-7](https://doi.org/10.1016/S1474-4422(21)00024-7).
- Martinez-Ramirez S, Greenberg SM, Viswanathan A. Cerebral microbleeds: Overview and implications in cognitive impairment. *Alzheimers Res Ther* 2014;6(3):33. <https://doi.org/10.1186/alzrt263>.
- Liu C, Li W, Tong KA, Yeom KW, Kuzminski S. Susceptibility-weighted imaging and quantitative susceptibility mapping in the brain. *J Magn Reson Imaging* 2015;42(1):23-41. <https://doi.org/10.1002/jmri.24768>.
- Liu C, Wei, H, Gong, N-J, Cronin, M, Dibb, R, Decker, K. Quantitative susceptibility mapping: Contrast mechanisms and clinical applications. 2015.
- Puy L, Pasi M, Rodrigues M, et al. Cerebral microbleeds: From depiction to interpretation. *J Neurol Neurosurg Psychiatry* 2021;92(6):598-607. <https://doi.org/10.1136/jnnp-2020-323951>.
- Shi Y, Wardlaw JM. Update on cerebral small vessel disease: A dynamic whole-brain disease. *Stroke Vasc Neurol* 2016;1(3):83-92. <https://doi.org/10.1136/svn-2016-000035>.
- Lau KK, Wong YK, Teo KC, et al. Long-term prognostic implications of cerebral microbleeds in Chinese patients with ischemic stroke. *J Am Heart Assoc* 2017;6(12):e007360. <https://doi.org/10.1161/JAHA.117.007360>.

12. Gregoire SM, Chaudhary UJ, Brown MM, et al. The Microbleed Anatomical Rating Scale (MARS): Reliability of a tool to map brain microbleeds. *Neurology* 2009;73(21):1759-1766. <https://doi.org/10.1212/WNL.0b013e3181c34a7d>.
13. Kuijf HJ, de Bresser J, Geerlings MI, et al. Efficient detection of cerebral microbleeds on 7.0 T MR images using the radial symmetry transform. *Neuroimage* 2012;59(3):2266-2273. <https://doi.org/10.1016/j.neuroimage.2011.09.061>.
14. Seghier ML, Kolanko MA, Leff AP, Jäger HR, Gregoire SM, Werring DJ. Microbleed detection using automated segmentation (MIDAS): A new method applicable to standard clinical MR images. *PLoS One* 2011; 6(3):e17547. <https://doi.org/10.1371/journal.pone.0017547>.
15. 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;2:282-290. <https://doi.org/10.1016/j.nicl.2013.01.012>.
16. Loy G, Zelinsky A. Fast radial symmetry for detecting points of interest. *TPAMI* 2003;25(8):959-973. <https://doi.org/10.1109/TPAMI.2003.1217601>.
17. Chen Y, Villanueva-Meyer JE, Morrison MA, Lupo JM. Toward automatic detection of radiation-induced cerebral microbleeds using a 3D deep residual network. *J Digit Imaging* 2019;32(5):766-772. <https://doi.org/10.1007/s10278-018-0146-z>.
18. Qi D, Hao C, Lequan Y, et al. Automatic detection of cerebral microbleeds from MR images via 3D convolutional neural networks. *IEEE Trans Med Imaging* 2016;35(5):1182-1195. <https://doi.org/10.1109/TMI.2016.2528129>.
19. Duan Y, Shan W, Liu L, et al. Primary categorizing and masking cerebral small vessel disease based on "deep learning system". *Front Neuroinform* 2020;14:17. <https://doi.org/10.3389/fninf.2020.00017>.
20. Liu S, Utriainen D, Chai C, et al. Cerebral microbleed detection using susceptibility weighted imaging and deep learning. *Neuroimage* 2019; 198:271-282. <https://doi.org/10.1016/j.neuroimage.2019.05.046>.
21. Milletari F, Navab N, Ahmadi S-A. V-Net: Fully convolutional neural networks for volumetric medical image segmentation. *arXiv* 2016; 565-571. <https://doi.org/10.48550/arXiv.1606.04797>.
22. Kaiming H, Xiangyu Z, Shaoqing R, Jian S. Deep residual learning for image recognition. *IEEE* 2016;770-778. <https://doi.org/10.1109/CVPR.2016.90>.
23. Al-masni MA, Kim W-R, Kim EY, Noh Y, Kim D-H. Automated detection of cerebral microbleeds in MR images: A two-stage deep learning approach. *Neuroimage Clin* 2020;28:102464. <https://doi.org/10.1016/j.nicl.2020.102464>.
24. Tian LIU, Surapaneni K, Min LOU, Liuquan C, Spincemaille P, Yi W. Cerebral microbleeds: Burden assessment by using quantitative susceptibility mapping. *Radiology* 2012;262(1):269-278. <https://doi.org/10.1148/radiol.11110251>.
25. Chen W, Zhu W, Kovanlikaya I, et al. Intracranial calcifications and hemorrhages: Characterization with quantitative susceptibility mapping. *Radiology* 2014;270(2):496-505. <https://doi.org/10.1148/radiol.13122640>.
26. Straub S, Laun FB, Emmerich J, et al. Potential of quantitative susceptibility mapping for detection of prostatic calcifications: Prostate-QSM: Detection of calcifications. *J Magn Reson Imaging* 2017;45(3):spcone-spcone. <https://doi.org/10.1002/jmri.25674>.
27. DrikJanKroon. HessianbasedFrangiVesselnessfilter. *MATLAB Central File Exchange*. 2021. Available from: <http://www.mathworks.com/matlabcentral/fileexchange/24409-hessian-based-frangi-vesselness-filter>
28. Sandro. Fast Radial Symmetry Transform. *MATLAB Central File Exchange*. 2021. Available from: <https://www.mathworks.com/matlabcentral/fileexchange/45961-fast-radial-symmetry-transform>
29. Frangi AF, Niessen WJ, Vincken KL, Viergever MA. *Multiscale vessel enhancement filtering*. Berlin: Springer Berlin Heidelberg; 2006. p 130-137.
30. Smith SM. Fast robust automated brain extraction. *Hum Brain Mapp* 2002;17(3):143-155. <https://doi.org/10.1002/hbm.10062>.
31. Wu D, Liu S, Buch S, Ye Y, Dai Y, Haacke EM. A fully flow-compensated multiecho susceptibility-weighted imaging sequence: The effects of acceleration and background field on flow compensation. *Magn Reson Med* 2016;76(2):478-489. <https://doi.org/10.1002/mrm.25878>.
32. Klohs J, Deistung A, Schweser F, et al. Detection of cerebral microbleeds with quantitative susceptibility mapping in the ArcAbeta mouse model of cerebral amyloidosis. *J Cereb Blood Flow Metab* 2011;31(12):2282-2292. <https://doi.org/10.1038/jcbfm.2011.118>.
33. Wei H, Dibb R, Zhou Y, et al. Streaking artifact reduction for quantitative susceptibility mapping of sources with large dynamic range. *NMR Biomed* 2015;28(10):1294-1303. <https://doi.org/10.1002/nbm.3383>.
34. Zhang Y, Wei H, Cronin MJ, He N, Yan F, Liu C. Longitudinal atlas for normative human brain development and aging over the lifespan using quantitative susceptibility mapping. *Neuroimage* 2018;171:176-189. <https://doi.org/10.1016/j.neuroimage.2018.01.008>.
35. Zhu J-Y, Park T, Isola P, Efros AA. Unpaired image-to-image translation using cycle-consistent adversarial networks. 2017.
36. Isola P, Jun-Yan Z, Tinghui Z, Efros AA. Image-to-image translation with conditional adversarial networks. *IEEE* 2017;5967-5976. <https://doi.org/10.1109/CVPR.2017.632>.
37. Lee SH, Bae HJ, Ko SB, Kim H, Yoon BW, Roh JK. Comparative analysis of the spatial distribution and severity of cerebral microbleeds and old lacunes. *J Neurol Neurosurg Psychiatry* 2004;75(3):423-427. <https://doi.org/10.1136/jnnp.2003.015990>.