Published in final edited form as:
Magn Reson Imaging. 2011 July ; 29(6): 844–852. doi:10.1016/j.mri.2011.02.028.

Semi-Automated Detection of Cerebral Microbleeds in Magnetic Resonance Images

Samuel R.S. Barnes[a,b,c], **E. Mark Haacke**[a,b,c,d], **Muhammad Ayaz**[e], **Alexander S. Boikov**[c,f], **Wolff Kirsch**[g], and **Dan Kido**[b]

[a]Biomedical Engineering, Wayne State University, Detroit, MI, USA 818 W. Hancock Detroit, Michigan 48201

[b]Radiology, Loma Linda University Medical Center, Loma Linda, CA, USA 11234 Anderson Street, Room B623 Loma Linda, CA 92350

[c]The Magnetic Resonance Imaging Institute for Biomedical Research, Detroit, MI, USA 440 E. Ferry St. Detroit, MI 48202

[d]Department of Radiology, Wayne State University, Detroit, MI, USA HUH—MR Research G030/ Radiology 3990 John R Road Detroit, MI 48201

[e]Hemorrhagic Stroke Research Center, Department of Neurology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA 175 Cambridge Street, Suite-300 Boston, MA - 02114

[f]Wayne State University School of Medicine, Detroit, MI, USA 540 East Canfield Street Detroit, MI 48201-1998

[g]Neurosurgery Center for Research, Training and Education, Loma Linda University, Loma Linda, CA, USA 11175 Campus St. Ste. 11113 Loma Linda, CA 92350

Abstract

Cerebral microbleeds (CMB) are increasingly being recognized as an important biomarker for neurovascular diseases. So far all attempts to count and quantify them have relied on manual methods that are time consuming and can be inconsistent. A technique is presented that semi-automatically identifies CMBs in susceptibility weighted images (SWI). This will both reduce the processing time and increase the consistency over manual methods. This technique relies on a statistical thresholding algorithm to identify hypointensities within the image. A support vector machine (SVM) supervised learning classifier is then used to separate true CMB from other marked hypointensities. The classifier relies on identifying features such as shape and signal intensity to identify true CMBs. The results from the automated section are then subject to manual review to remove false positives. This technique is able to achieve a sensitivity of 81.7% compared to the gold standard of manual review and consensus by multiple reviewers. In subjects with many CMBs this presents a faster alternative to current manual techniques at the cost of some lost sensitivity.

Corresponding Author: Semi-Automated Detection of Cerebral Microbleeds Samuel Barnes Loma Linda University 11234 Anderson Street, Room B623 Loma Linda, CA 92350 Phone: (909) 558-7114 Fax: (909) 558-0335 sbarnes@llu.edu.

Keywords

susceptibility weighted imaging; segmentation; support vector machine; cerebral microbleed

1. Introduction

Cerebral microbleeds (CMB) are generally defined as asymptomatic small bleeds (diameter <5–10 mm) seen primarily with a magnetic resonance imaging (MRI) T2*-weighted scan (1). It has been shown that MR-visible CMBs consist mostly of hemosiderin (2,3), whose large susceptibility effect accounts for good visibility in T2*-weighted scans. Individuals with cerebrovascular disease related dementia and even a subset of the healthy aged population develop CMBs. While CMBs are usually asymptomatic, they can be important clinical indicators of different diseases. The presence of CMBs can suggest an increased chance of lacunar infarction, intracerebral hemorrhage, or hemorrhagic stroke in patients already suffering from these pathologies (4–6). Increasing numbers of CMBs indicate a worsening of dementia upon follow-up after primary lobar intracerebral hemorrhage (4). In another longitudinal study, a temporal increase in the number of CMBs was correlated with worsening dementia (7). CMB detection is becoming an important tool for non-invasive detection of diseases like cerebral amyloid angiopathy (CAA), CADASIL (cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy) and hypertension (1). For more details on CMBs please refer to these excellent reviews (1,8).

CMB contrast and detection is sensitive to many imaging variables such as field strength, echo time and resolution (1,9). Modern imaging protocols such as susceptibility weighted imaging (SWI)(10,11), that are routinely run at high resolution ($\leq 1 mm^3$), long echo time, and use the phase image to enhance contrast, are much more sensitive in detecting small bleeds than traditional protocols (7). Recent publications have shown that when SWI is compared with standard gradient echo imaging there is a three to six-fold increase in the number of CMBs seen (9,12).

CMBs are, by definition, small and are easy to confuse with other structures (so-called CMB mimics (1)). The small size and large numbers of CMBs and CMB mimics have hampered efforts to quantify their number and volumes. So far, methods have centered on manually identifying and drawing the lesions, or manually defining local thresholds for a small region of interest (ROI) (7,13). These methods suffer from a high amount of inter- and intra-observer error (14) and become extremely time-consuming as the number of patients and CMBs grows. In manual identification, a set of rules can be defined to determine whether any small hypointensity is a CMB (7) to help improve consistency. Recent publications have shown improvements by using standardized rating scales (14) but further improvement in consistency, speed of identification and quantification are desirable.

Automatically identifying the CMB is problematic as there are many dark structures on SWI images that can easily be confused with CMBs. Most notably, SWI's excellent venous contrast makes veins also appear dark. Any method that seeks to identify CMBs will need a way to differentiate CMBs from venous structures. We propose that after marking all dark structures, characteristic features (notably shape) can be calculated and used to separate CMBs from veins with the use of a support vector machine (SVM) classifier.

In this work we propose a semi-automated method of identifying and quantifying CMBs seen on high resolution SWI scans. This method has four steps: 1) a pre-processing step that consists of image interpolation and brain extraction to remove the skull and background, 2) statistical thresholding that marks all hypointensities, 3) an SVM classifier that eliminates

hypointense noise and veins, and finally, 4) a manual review of results to eliminate the remaining false positives (Figure 1). The manual intervention is limited to making yes or no decisions on the automated suggestions which reduces processing time and observer variability.

2. Methods

2.1 Data Collection

All SWI images were acquired as part of a longitudinal study on individuals with mild cognitive impairment or early Alzheimer's disease (15). Permission was obtained from our institutions IRB and data were collected at 1.5T using a fully flow compensated 3D gradient recalled echo (GRE) sequence with a resolution of $0.5\times1.0\times2.0$ mm^3 and a matrix of $512\times320\times48$. Imaging parameters were: TR=57ms, TE=40ms, FA=20°. Standard SWI processing with high pass filtered phase images was performed (11). A total of six datasets (containing 126 CMBs) were processed with this method.

A manual review of the data was performed independently by three reviewers using a standard set of criteria (7). All three reviewers then compared results and reached consensus when differences in counts were encountered. All results were reviewed by a senior neuroradiologist (DK). More details on the manual review can be found in (7). This manual review was used as the gold standard that the computer algorithm was compared to.

2.2 Pre-Processing: Brain Extraction and Interpolation

Brain extraction was accomplished with a thresholding technique based on the magnitude and phase data. Regions of no signal were identified as those having low intensities in the magnitude image and uniform noise distribution over all values in the phase image. By using both the magnitude and phase, the accuracy of the extraction is improved compared to using either independently (16). Isolated islands of points (<5000 voxels) were removed and isolated holes (<1000 voxels) in large regions were restored to their original values to remove remnant skull and restore any holes inside the brain, respectively. Finally, an erosion of three voxels was performed to remove partial-volume brain tissue that has a lower intensity and hence creates a rim artifact that is picked up in the thresholding step.

After the brain extraction, the images were interpolated by two in the X and Y directions using zero filling in k-space with a Hanning filter to reduce Gibbs ringing. This interpolation allowed for better shape definition of small structures and leads to better classification. A given marked structure of three uninterpolated voxels can assume only two different shapes (a line and an `L' shape), making it hard to get meaningful shape information. Upon interpolation, this same structure will be 12 interpolated voxels, which can form a wide variety of shapes—making shape a more robust identifying feature. While interpolation does not create new information, as scanning at a higher native resolution would, it more easily allows the information of the surrounding voxels (which are not thresholded) to be included in the shape analysis.

2.3 Local Statistical Thresholding

Local statistical thresholding is used to separate the Gaussian distribution of background tissue from the low outlier CMBs. Our implementation of local statistical thresholding calculates the threshold based on the mean and standard deviation of the surrounding tissue. We assume a white noise distribution in the parenchyma for a roughly Gaussian distribution with a small number of low-intensity outliers (potential microbleeds, vessels, other objects of low signal intensity) that do not significantly alter the mean or standard deviation of the region. To identify the CMBs, we seek to model the Gaussian distribution by calculating the

mean and standard deviation, then setting a threshold to be a certain number of standard deviations below the mean to separate the parenchyma from the outliers. This method will fail if the number of low outliers is too large, making the distribution non-Gaussian, lowering the mean and increasing the standard deviation. To help prevent this, two restrictions are placed on the mean and standard deviation calculations. First, before local thresholds are calculated, an initial whole-brain global threshold is performed to identify and exclude large dark structures that would skew the calculation. Second, the local thresholding is iterated with previously marked voxels from the global and local thresholds being excluded from subsequent calculations to improve the accuracy of the calculated mean and standard deviation.

The global threshold intensity is preset to a conservative value where only very dark structures (usually the large ones) are marked (Figure 1b and 2a). This step aims to remove large structures, such as large veins, that could skew the mean and standard deviation values that are calculated in the local statistical thresholding. The voxels marked in the global threshold step are not used for the mean and standard deviation calculations in the next step but are marked as low outliers and included in the local thresholding results of potential CMBs.

Local thresholding is performed next (Figure 1b and 2b). A small ROI (size 21×21×3 voxels) centered on a single voxel is considered. In this ROI, the mean and standard deviation are calculated, excluding any points that have been previously marked. A local threshold is then calculated to be $\alpha \cdot \sigma$ below the mean, where σ is the standard deviation. If the voxel of interest is below this value, it is marked.

$$voxel_{marked} \quad < \bar{x}_{local} - \alpha \cdot \sigma_{local} \tag{1}$$

$$voxel_{background} \geq \bar{x}_{local} - \alpha \cdot \sigma_{local} \tag{2}$$

where α sets the degree of thresholding and can be adjusted for different types of data and different signal-to-noise ratios (SNRs). α was carefully chosen to set the threshold at the low edge of normally distributed background tissue for our data (average SNR = 24). A value of $\alpha = 2.5$ was used as this would theoretically mark only 0.62% of normally distributed parenchyma (observed values were <1%, see Figure 3). This provides for maximum sensitivity while keeping the number of false positives in the parenchyma low enough that they can be eliminated in subsequent steps (they tended to be isolated voxels and were eliminated with by the cluster size filter discussed below). Values lower than 2.5 gave excessive parenchyma contamination that was not easily removed and values larger than 2.5 reduced sensitivity. The area of interest is then moved and the process is repeated until a threshold has been calculated for every voxel and the entire image has been processed. In practice, calculating a new threshold for every single voxel proved too computationally expensive, so a single calculated threshold was applied to the center group of 5x5x3 voxels, rather than to the single center voxel. This provided a decrease in computation time by a factor of 75 with little effect on the end results.

This local thresholding process was then iterated. After the entire image has been processed, all thresholds are recalculated with previously marked voxels excluded from the mean and standard deviation calculations. Each successive mean and standard deviation calculation therefore includes fewer low outliers and a more accurate threshold is obtained. This improves the identification of edge voxels and small structures that might be located next to

a large dark structure (a large vein for example). Practically, three iterations were run since little change was noted for higher numbers of iterations.

The size of the 21×21×3 region used to calculate the mean and standard deviation can also be adjusted to help ensure a Gaussian distribution. A smaller region gives more localized results but runs the risk of having its mean dominated by a large bleed or vein in the area. A larger region is less likely to be non-Gaussian but is also less sensitive to local information. A balance must be struck between the two and a relatively small area of 21×21×3 voxels was found to be a good compromise. The volume of the CMBs is significantly smaller than this so they do not significantly skew the distribution. In this way, the technique is optimized to detect small structures; larger structures will only be detected if they are dark enough to be picked up by the initial global threshold.

2.4 Support Vector Machines

Support vector machines (SVMs) are a type machine learning supervised classifier. In our application they are used to do a binary classification between two classes of objects: CMBs and noise. For this type of binary classification, they seek to define a separating hyperplane between each class that both maximizes the distance between each class and the hyperplane (a maximum margin classifier) and minimizes the number of incorrectly classified objects. This makes the SVM a very general and robust type of classifier (17).

SVMs take advantage of the so-called kernel trick to transform the original features into a higher dimensional Hilbert space. The maximum margin hyperplane is derived in this high dimensional space. The hyperplane is then transformed back from the high dimensional Hilbert space into the lower dimensional feature space. This is a nonlinear process that transforms the linear hyperplane into a nonlinear separation.

2.4.1 Classification – Feature Selection—Applying a classifier to the thresholding results is a necessary step to reduce the large number of false positives that are generated (see Table 1). The false positives are largely generated by dark veins, although some noise and susceptibility artifacts (such as those near the sinuses) also are marked. To remove these false positives a SVM classifier is used.

Before the classifier can be used, all marked voxels are sorted into connected clusters, where a cluster is defined as all marked voxels that are adjacent to one another. Each cluster is treated as a single object and is kept or discarded in its entirety. Clusters less than or equal to 10 interpolated voxels (equivalent to 2.5 voxels on the original datasets) are automatically removed as they are assumed to be noise (Figure 2b and 2c). Relevant features about each cluster (such as shape and intensity) are calculated and used by the classifier to make a “keep” or “discard” decision about each cluster (Figure 4). Feature selection focused on shape, size and intensity to distinguish CMBs from all other marked structures.

CMBs have a very characteristic shape—they are nearly spherical, making shape features the best identifying features. To this end, five different shape features were calculated and used: compactness, the three eigenvalues of the covariance matrix and the relative anisotropy (RA) calculated from these eigenvalues. These features were calculated as described below.

Compactness is approximately a ratio of surface area to volume, with spheres having the highest theoretical compactness. CMBs therefore have a very high compactness and veins have a very low compactness. Since we are dealing with discrete voxels, compactness is calculated by counting adjacent voxel faces:

$$C_d = \frac{A_C - A_{C_{\min}}}{A_{C_{\max}} - A_{C_{\min}}} \quad (2)$$

where A_C is the actual number of adjacent faces, A_{Cmin} is the minimum number of adjacent faces possible for that particular number of voxels, and A_{Cmax} is the maximum number of adjacent faces possible for that particular number of voxels. More details can be found in (18).

The covariance matrix is a three by three matrix that describes mass distribution (19). The eigenvalues and eigenvectors calculated from this matrix describe the principal axis. Thus, a spherical distribution would have three equal eigenvalues while a cylindrical distribution would have one large eigenvalue and two small eigenvalues. The eigenvalues were used individually as features and also combined to create an RA value using the standard formula

$$RA = \sqrt{1/2} \cdot \sqrt{\frac{(a-b)^2 + (b-c)^2 + (c-a)^2}{a+b+c}} \quad (4)$$

where *a*, *b*, and *c* are the eigenvalues of the covariance matrix of the cluster, *a* being the largest and *c* being the smallest. RA is a measure of how anisotropic a shape is, or how non-spherical it is, so CMBs should have a low RA value.

In addition to shape features, intensity and size were also used. A total of eight intensity features were calculated for each cluster: minimum, maximum, mean, and standard deviation for both the magnitude and phase images. The intensity features helped further characterize the CMBs and give the classifier more information to use. Finally, size was also included as a feature. While size alone is not a good criterion for identifying CMBs (1) other characteristic features (such as shape) might change with the size of the CMB. Having the size information thus allows the classifier to select the best set of identifying features for each different size of CMB.

2.4.2 Classification – Training—For the SVM we used the open source library libSVM version 2.88 (20). This was chosen due to its wide use, good documentation, and open source availability so it could be easily integrated with our in-house C++ program (SPIN) in which all the other programming was implemented. A radial basis function was used, all features were scaled from 0 to 1, and the CMB class was weighted by a factor of 100. The radial basis function was chosen as it has been shown to give similar behavior as other commonly used basis functions (linear, and sigmoid) for certain values of model parameters cost and gamma and have less numerical difficulties compared with polynomial kernels (21). Feature scaling was implemented to give all features equal weighting. If feature scaling is not used, the features with larger amplitude generally contribute more to the classification, possibly skewing the results. Finally, a weighting factor of 100 was used to offset the number disparity between the two classes. Approximately 15,000 true negatives and only 120 true positives were used to train the classifier so the most accurate naïve classification is to mark no CMBs, yielding an accuracy of over 99%. To correct for this, the CMB class was weighted in the classifier as more important to prevent this trivial case.

Training the SVM is an important step that can dramatically influence the accuracy of the classifier. Proper parameter selection, feature selection, class weighting, and feature scaling can all have a large influence. The cost and gamma model parameters control how the

classifier behaves and must be selected based on the type of features one is using. To find optimal values a cross validation technique was used. In cross validation, the complete dataset for which the ground truth is known is randomly divided into testing and training portions, the classifier is trained on one portion and tested on another, and the accuracy of the test portion is then recorded. In this way, different values of cost and gamma can be examined and their accuracy determined. A course grid of cost and gamma values was first sampled to find the general area of best performance, a region of good performance was identified and that area was more tightly sampled (Figure 5). The cost and gamma values with the maximum accuracy from the tighter sampling were selected and used for all the proceeding training and accuracy measurements.

The cost and gamma cross validation is processor-intensive, taking multiple days if executed in a serial fashion. However, since each test is independently calculated it easily lends itself to parallelization. A parallelized version was implemented, running on a modern quad core processor with two virtualized cores per physical core (for a total of four physical and eight virtualized cores) resulting in a reduction in processing time by a factor of five. This reduced the processing time down to 1–2 days (depending on the sampling parameters) which is quite reasonable as it only needs to be performed once per cohort of subjects for a particular disease or image type. If the work load were further divided between more cores or more computers it could be reduced further.

Once the cost and gamma parameters were selected, the computation time was minimal. The classifier must be trained using the selected parameters but this takes less than a minute and only needs to be performed once per cohort of subjects. Once the classifier has been setup classifying a single dataset takes less than 10 seconds. The total computation time per dataset for the automated section (brain extraction, thresholding, and SVM classification) is thus around one minute.

2.5 Evaluation and Statistics

Accuracy of the semi-automated technique was evaluated in a binary fashion. If the semi-automated technique identified a cluster of voxels in the final results where a CMB had been identified in the manual review it was considered a true positive, if no cluster of voxels were marked in the final results in that area it was considered a false negative. The issues of edge definition and extraneous surrounding voxels being included in a CMB cluster and contaminating the results practically was only a problem when a nearby structure (such as a vein) would merge with a CMB. These types of clusters were generally eliminated by the classifier as the non-CMB voxels tended to spoil their CMB features (namely shape). All true CMBs that were selected by the classifier had a relatively clean and logical boundary, with a minimal amount of extraneous voxels (Figure 6).

The sensitivity and specificity of each step were evaluated with manual review and consensus by multiple reviewers used as the ground truth. The specificity of the thresholding step was not calculated, as true negatives were undefined at that stage. For subsequent steps true negatives were defined as clusters that were removed (by the SVM or manually) that had not been marked in the ground truth review (so are implicitly defined as negative). The sensitivity and specificity were calculated for each step individually and of the entire process was calculated.

3. Results

The thresholding step had a very high sensitivity of 95%, only missing six CMBs out of 126 (see Table 1). However, it had a very large number of false positives as it is designed to mark all hypointense structures in the image. The SVM classifier is able to remove most of

the false positives at the loss of some sensitivity. The automated processing had an overall sensitivity of 81.7% and specificity of 95.9%. After the data has undergone manual review the remaining false positives are removed for a final sensitivity and specificity of 81.7% and 100% respectively.

The subjects had an average of 21 CMBs as determined by manual review, on average 17 were detected by the semi-automated method with 107 false positives. Specific numbers for each subject are given in Table 2.

4. Discussion and Conclusion

Development of the classifier focused on maximizing the sensitivity, even at the cost of specificity. The complicated nature of CMB identification makes a manual review of the automated results necessary so CMB mimics are not included. For the sake of efficiency this review is limited to removal of false positives (as manually searching for missed CMBs would make the automated tools redundant). Since false positives are being manually removed, the only cost of a low sensitivity algorithm is a longer manual review time. False negatives on the other hand can never be recovered. Thus, improving the sensitivity increases the accuracy of the identification while improving the specificity only reduces processing time.

The error rates for each step of the process, shown in Table 1, indicate that the step requiring the most improvement is the classifier; it is the main source of the loss of sensitivity. The classifier could be improved with a number of different techniques. New features that give the classifier more information about each marked cluster could improve the accuracy substantially. First attempts with this classifier focused on shape features alone as those were thought to be the most promising. When the intensity features were included they resulted in a 10% improvement in the classifier sensitivity. The addition of other relevant features could give similar improvements. Some possible features to try would be anatomical location and intensity features from other MR sequences.

Anatomical location could be determined automatically by fitting to a standardized brain atlas. This could significantly improve the classification as anatomic location is used as one of the evaluating criteria in manual classification. Intensities from other MR sequences could be used after co-registration of the datasets. For example, if high resolution MR angiography post contrast was performed, the veins and the arteries could be excluded based on their bright signal. While co-registration has been used in similar lesion segmentation problems (22) and could be used to remove some CMB mimics (1), it probably would not result in a dramatic improvement in sensitivity as most of the CMBs are not seen in other MR sequences. Nevertheless, it could be an important step in removing false positives and CMB mimics.

Further improvements could be made to the classifier itself. The addition of AdaBoost for classifier training could improve the quality of the SVM classification (22,23) and should be an area of future research. Other types of classifiers such as artificial neural networks could also be tried.

While the sensitivity of the technique is a little low, this is probably influenced by the dataset and method of manual review. This dataset was part of a serial study so all patients had at least three time points and some patients had four. For the manual review all four datasets were examined and compared to reach a decision about an individual CMB; this was noted to significantly increase the confidence of identification (7). While this assuredly improves the ground truth accuracy of the manual method this will somewhat unfairly penalize the semi-automated method as it has less information available because it is only

able to use a single time point to make a decision. While these numbers are probably a more accurate measure of its sensitivity for the true number of CMBs, a fair comparison against a manual reviewer who only has access to a single time point would probably show much more favorable numbers. The semi-automated method could also be adapted to take advantage of this additional information to improve results. While possible this would have limited applicability outside of this study as future studies are unlikely to have longitudinal data to analyze.

To get accurate manual counts, multiple people reviewed the data independently and then a consensus was sought on discrepancies between the reviewers. While this does lead to accurate and reliable numbers it is too time consuming for practical application. The average time for one person to identify all CMBs in a dataset manually varied depending on scan quality and number of CMBs but can range from 15 minutes to a few hours with most patients taking approximately 30 minutes. Tracing the CMB, for volume measurements and to ensure there are no double counts (this becomes increasing necessary as the number of CMBs increase), would approximately take an additional 10 minutes (to mark 21 CMBs, the average in this study). This is too time-consuming to be performed clinically—much less two or three times (by multiple reviewers or across multiple time points) to improve consistency. Reviewing a specific list of suspected CMBs generated by semi-automated detection improves consistency and reduces the manual review time to approximately 5–15 minutes, allowing the data to be efficiently assessed even by multiple reviewers.

The time savings we discuss throughout this paper is most applicable to studies with larger numbers of CMBs (>10) and studies where volume measurements are needed. If one is only interested in making a specific diagnostic judgment (e.g. looking for two or more CMBs in the lobar region to diagnose CAA) then manually examining the dataset is sufficient. But identifying and tracing all CMBs on subjects with 20, 40, or more CMBs the tools we describe begin to offer greater time savings as they eliminate the need to look through the entire dataset and identify possible CMBs and manually trace them.

Gains in specificity could likely be obtained by using datasets with higher SNR acquired at higher fields (≥3T). The higher SNR would allow α to be set to higher values, excluding more background voxels and reducing false positives. Gains in specificity for the automated portion could further reduce manual review times as fewer false positives would have to be removed manually.

The segmentation techniques and software developed by one of the authors (SB) have also been successfully applied to the related problem of segmenting and quantifying veins in SWI (24). This is a simple extension where the shape filtering/classification is run in the other direction and seeks to eliminate more spherical shapes (unwanted dark structures) and keep cylindrical objects (veins). This earlier work used a more simplistic filter (linear thresholds) with only two features (compactness and relative anisotropy). Extending this to use SVM and all 14 features proposed here could improve results significantly.

The training requirements of the classifier need to be explored in more depth. This is by far the most time intensive part of the operation, both in the manual CMB identification that is required for the SVM classifier training and the computational time for cost and gamma parameter selection. Due to this, it is highly desirable to minimize the amount of training that needs to be done, especially across studies. How well the classifier will perform when trained on data collected at one institution or on one type of patients and applied to data from a different institution or a different type of patients remains to be determined. Despite this, even if training is required for each cohort of patients, manually counting a subset of the subjects for training will still require less time than manually counting all of the subjects.

Processing times for training can be further reduced using parallelization algorithms (as we implemented) and multi-core multi-processor computers which are becoming increasingly common.

In conclusion we have shown a semi-automated technique of identifying and quantifying CMBs can perform reasonably well with some manual intervention. Although this technique still has some shortcomings in sensitivity we believe this is a strong first step in addressing this difficult problem. This technique reduces the processing time in subjects with many CMBs and should lead to more consistency between observers. Further verification on a larger group of patients with more diverse types of CMBs needs to be done to prove the applicability of this method as a general CMB detection algorithm.

Acknowledgments

Role of the funding source This work was supported by the National Institutes of Health (NIH); Contract grant number: 2R0 HL062983-04A2. They had no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; and in the decision to submit the paper for publication.

References

1. Greenberg SM, Vernooij MW, Cordonnier C, Viswanathan A, Al-Shahi Salman R, Warach S, Launer LJ, Van Buchem MA, Breteler MM. Cerebral microbleeds: a guide to detection and interpretation. Lancet Neurol. 2009; 8(2):165–174. [PubMed: 19161908]
2. Fazekas F, Kleinert R, Roob G, Kleinert G, Kapeller P, Schmidt R, Hartung HP. Histopathologic analysis of foci of signal loss on gradient-echo T2*-weighted MR images in patients with spontaneous intracerebral hemorrhage: evidence of microangiopathy-related microbleeds. AJNR Am J Neuroradiol. 1999; 20(4):637–642. [PubMed: 10319975]
3. Schrag M, McAuley G, Pomakian J, Jiffry A, Tung S, Mueller C, Vinters HV, Haacke EM, Holshouser B, Kido D, Kirsch WM. Correlation of hypointensities in susceptibility-weighted images to tissue histology in dementia patients with cerebral amyloid angiopathy: a postmortem MRI study. Acta Neuropathol. 2010; 119(3):291–302.
4. Greenberg SM, Eng JA, Ning M, Smith EE, Rosand J. Hemorrhage burden predicts recurrent intracerebral hemorrhage after lobar hemorrhage. Stroke. 2004; 35(6):1415–1420. [PubMed: 15073385]
5. Fan YH, Zhang L, Lam WW, Mok VC, Wong KS. Cerebral microbleeds as a risk factor for subsequent intracerebral hemorrhages among patients with acute ischemic stroke. Stroke. 2003; 34(10):2459–2462. [PubMed: 12958325]
6. Imaizumi T, Horita Y, Hashimoto Y, Niwa J. Dotlike hemosiderin spots on T2*-weighted magnetic resonance imaging as a predictor of stroke recurrence: a prospective study. J Neurosurg. 2004; 101(6):915–920. [PubMed: 15597750]
7. Ayaz M, Boikov AS, Haacke EM, Kido DK, Kirsch WM. Imaging cerebral microbleeds using susceptibility weighted imaging: One step toward detecting vascular dementia. J Magn Reson Imaging. 2009; 31(1):142–148. [PubMed: 20027582]
8. Cordonnier C, Al-Shahi Salman R, Wardlaw J. Spontaneous brain microbleeds: systematic review, subgroup analyses and standards for study design and reporting. Brain. 2007; 130(Pt 8):1988–2003. [PubMed: 17322562]
9. Nandigam RN, Viswanathan A, Delgado P, Skehan ME, Smith EE, Rosand J, Greenberg SM, Dickerson BC. MR imaging detection of cerebral microbleeds: effect of susceptibility-weighted imaging, section thickness, and field strength. AJNR Am J Neuroradiol. 2009; 30(2):338–343. [PubMed: 19001544]
10. Reichenbach JR, Venkatesan R, Schillinger DJ, Kido DK, Haacke EM. Small vessels in the human brain: MR venography with deoxyhemoglobin as an intrinsic contrast agent. Radiology. 1997; 204(1):272–277. [PubMed: 9205259]
11. Haacke EM, Xu Y, Cheng YC, Reichenbach JR. Susceptibility weighted imaging (SWI). Magn Reson Med. 2004; 52(3):612–618. [PubMed: 15334582]

12. Tong KA, Ashwal S, Holshouser BA, Shutter LA, Herigault G, Haacke EM, Kido DK. Hemorrhagic shearing lesions in children and adolescents with posttraumatic diffuse axonal injury: improved detection and initial results. Radiology. 2003; 227(2):332–339. [PubMed: 12732694]

13. Greenberg SM, Nandigam RN, Delgado P, Betensky RA, Rosand J, Viswanathan A, Frosch MP, Smith EE. Microbleeds versus macrobleeds: evidence for distinct entities. Stroke. 2009; 40(7): 2382–2386. [PubMed: 19443797]

14. Cordonnier C, Potter GM, Jackson CA, Doubal F, Keir S, Sudlow CL, Wardlaw JM, Al-Shahi Salman R. improving interrater agreement about brain microbleeds: development of the Brain Observer MicroBleed Scale (BOMBS). Stroke. 2009; 40(1):94–99. [PubMed: 19008468]

15. Kirsch W, McAuley G, Holshouser B, Petersen F, Ayaz M, Vinters HV, Dickson C, Haacke EM, Britt W Iii, Larsen J, Kim I, Mueller C, Schrag M, Kido D. Serial Susceptibility Weighted MRI Measures Brain Iron and Microbleeds in Dementia. J Alzheimers Dis. 2009; 17(3):599–609. [PubMed: 19433895]

16. Pandian DS, Ciulla C, Haacke EM, Jiang J, Ayaz M. Complex threshold method for identifying pixels that contain predominantly noise in magnetic resonance images. J Magn Reson Imaging. 2008; 28(3):727–735. [PubMed: 18777533]

17. Vapnik, V. Statistical Learning Theory. Wiley-Interscience; New York: 1998.

18. Bribiesca E. A Measure of Compactness for 3D Shapes. Computers and Mathematics with Applications. 2000; 40:1275–1284.

19. Wasserman, L. All of statistics : a concise course in statistical inference. Springer; New York: 2004. p. xixp. 442

20. Chang C-C, Lin C-J. LIBSVM : a library for support vector machines. Software. 2001 available at http://wwwcsientuedutw/~cjlin/libsvm.

21. Hsu, C-W.; Chang, C-C.; Lin, C-J. Technical Report. National Taiwan University; 2003. A practical guide to support vector classification.

22. Lao Z, Shen D, Liu D, Jawad AF, Melhem ER, Launer LJ, Bryan RN, Davatzikos C. Computer-assisted segmentation of white matter lesions in 3D MR images using support vector machine. Acad Radiol. 2008; 15(3):300–313. [PubMed: 18280928]

23. Morra JH, Tu Z, Apostolova LG, Green AE, Toga AW, Thompson PM. Comparison of AdaBoost and support vector machines for detecting Alzheimer's disease through automated hippocampal segmentation. IEEE Trans Med Imaging. 2010; 29(1):30–43. [PubMed: 19457748]

24. Ge Y, Zohrabian VM, Osa EO, Xu J, Jaggi H, Herbert J, Haacke EM, Grossman RI. Diminished visibility of cerebral venous vasculature in multiple sclerosis by susceptibility-weighted imaging at 3.0 Tesla. J Magn Reson Imaging. 2009; 29(5):1190–1194. [PubMed: 19388109]

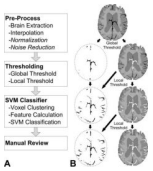

Figure 1.

A) A flowchart of processing steps, the two pre-processing steps in italics were not used in these datasets but might be useful for other types of data, and B) a schematic showing thresholding step in more detail are shown. Each thresholding iteration marks additional voxels that are added to the final results and excluded from the next iteration.

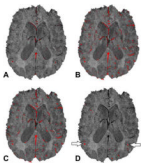

Figure 2.
The intermediate results after different thresholding steps on a single slice are shown. A) Results after the whole brain global threshold. B) Results of the local statistical threshold before removing small clusters (assumed to be noise). C) Results of local statistical threshold after removal of small clusters. D) Results of the SVM classifier which removes most false positives (veins and other dark structures).

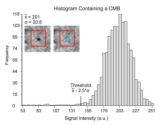

Figure 3.

A histogram of a 21×21×3 ROI that is used to calculate local thresholds is shown containing a typical CMB. The choice of α=2.5 adequately excludes the parenchyma while accurately marking the CMB. The CMB is small enough relative to the size of the ROI that it does not significantly affect the mean or standard deviation. The cutout shows the ROI and CMB and the thresholding results.

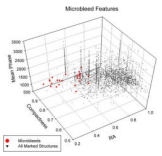

Figure 4.

A plot of three features (out of the 14 used) that can be used to identify CMBs is shown. The triangles show all marked voxel clusters as generated by a statistical thresholding algorithm, the circles are the clusters which correspond to actual CMBs as determined by manual inspection of the dataset. Using these features, it is possible to obtain a subset of a few hundred suspected CMBs which can be quickly sorted manually. The circles are drawn larger to make them more visible.

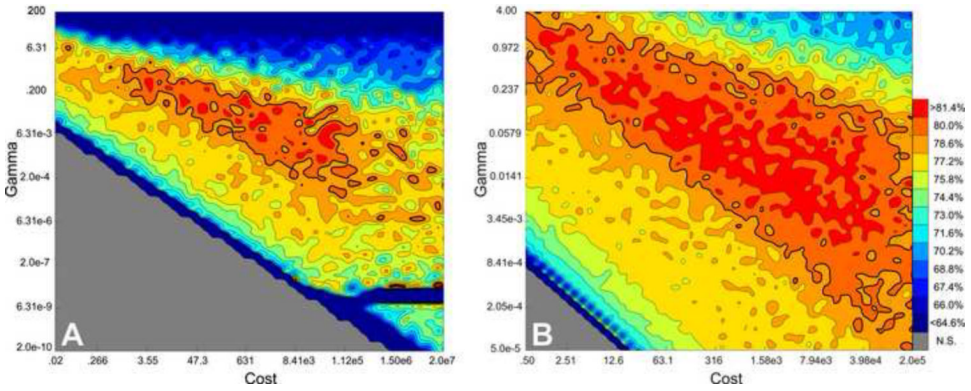

Figure 5.

A plot of parameter selection of cost and gamma using cross validation is shown. A) A large parameter space was originally examined with nFold of 2 (half dataset used for training half for testing) to identify the best region for a more detailed examination. B) A more detailed examination of a smaller parameter space with nFold of 5 (4/5 of the dataset used for training 1/5 for testing repeated 5 times). The gray area represents parameters that weren't tested due to their consistently low values.

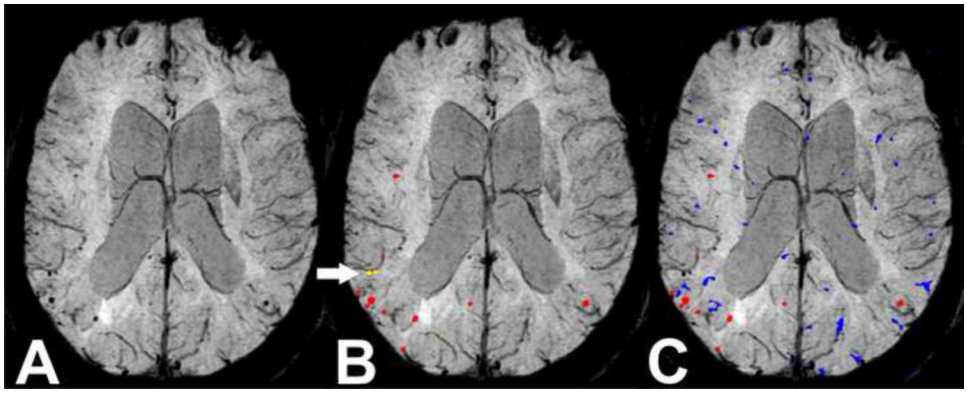

Figure 6.

A) Minimum intensity projection (mIP) over 26mm of SWI data. B) Manually marked true CMBs identified using magnitude, phase, SW, and mIP images. C) Automatically marked suspected CMBs. The yellow CMBs (marked by arrows) indicate the ones that were missed by the automated methods, the red indicate bleeds that were identified in both methods, and blue are the false positives from the automated methods. The bleeds that were missed by the automated methods (yellow) were merged with the vein they lie adjacent to causing them to have more vein like features.

Table 1

Summary of Accuracy

The accuracy and error rates of each step individually are listed above using the consensus manual counts as the ground truth. Note the high number of false positive connected clusters from the thresholding step requires the use of an SVM classifier. The classifier eliminates most of the false positive but leaves a significant number (645). These are spread across six datasets meaning there are approximately 100 false positives per dataset that must be removed in the manual review.

	Threshold	SVM Classifier	Manual Review	Final Results
True Positives	120	103	103	*103*
True Negatives	NA	15,162	645	*15,807*
False Positives	15,807	645	0	*0*
False Negatives	6	17	0	*23*
Sensitivity	95.2%	85.8%	100%	*81.7%*
Specificity	NA	95.9%	100%	*100%*

Table 2

Summary of Subjects

The number of true CMBs and CMBs found with the semi-automated method along with false positives are listed for each of the six subjects.

	True CMB	True Positives	False Positive
Subject 1	40	37	145
Subject 2	6	2	76
Subject 3	10	9	167
Subject 4	37	27	48
Subject 5	6	6	102
Subject 6	27	22	107
Average	*21*	*17.2*	*107.5*