A Computer Aided Detection System for Cerebral Microbleeds in Brain MRI

Babak Ghafary Asl

Submitted for the Degree of Master of Science
in Electrical Engineering
Department of Technology,
Kalmar University, Sweden

Supervisor: **Marleen de Bruijne**
Associate Professor in Department of Computer Science (DIKU),
University of Copenhagen, Denmark
Assistance Professor in Biomedical Imaging Group,
Erasmus Rotterdam Medical Center, Netherlands

Examiner: **Wlodek J. Kulesza**
Professor in School of Engineering,
Blekinge Technical University(BTH), Karlskrona, Sweden

Abstract

Advances in MR technology have improved the potential for visualization of small lesions in brain images. This has resulted in the opportunity to detect cerebral microbleeds (CMBs), small hemorrhages in the brain that are known to be associated with risk of ischemic stroke and intracerebral bleeding. Currently, no computerized method is available for fully- or semi-automated detection of CMBs.

In this paper, we propose a CAD system for the detection of CMBs to speed up visual analysis in population-based studies. Our method consists of three steps: (i) skull-stripping (ii) initial candidate selection (iii) reduction of false-positives using a two layer classification and (iv) determining the anatomical location of CMBs. The training and test sets consist of 156 subjects (448 CMBs) and 81 subjects (183 CMBs), respectively. The geometrical, intensity-based and local image descriptor features were used in the classification steps. The training and test sets consist of 156 subjects (448 CMBs) and 81 subjects (183 CMBs), respectively. The sensitivity for CMB detection was 90% with, on average, 4 false-positives per subject.

Key words: Brain MRI, cerebral microbleeds, classification, computer aided diagnosis.

Acknowledgments

First and foremost I would like to sincerely acknowledge my supervisor Dr. Marleen de Bruijne , for providing me an opportunity to conduct my master's thesis and her support throughout the work. I would like to thanks Prof.Wlodek J. Kulesza for all his support and comments for this thesis. An special thanks to neuroimaging group in Erasmus Rotterdam and their patient to share their ideas especially group members: Fedde van der Lijn, Renske de Boer, Stefan Klein and Henri Vrooman. Also special thanks to Radiology group in Erasmus MC for their support and following the thesis process: Meike Vernooij, Marrielle Peols and Aad van der Lugt. Finally, I would like to thank each and everyone who helped me to finish this thesis.

Contents

List of figures

1 Introduction

Cerebral microbleeds (CMBs) are hemosiderin deposits in the brain that are presumed to be caused by leakage of red blood cells from small blood vessels, which can be image using MRI [12]. They are known to be highly prevalent in patients with ischemic stroke or intracerebral bleeding and are possibly related to worse prognoses in these diseases [4, 16]. Furthermore, it has been recently shown that microbleeds are also common in the general elderly population, and that their location in the brain may indicate their underlying etiology [26].

The MRI protocols can contain different sequences including: T1-weighted, proton density-weighted (PD), Flair (fluid-attenuated inversion recovery) and T2*-weighted GRE (gradient-recalled echo). However only some of these sequences are used for visual detection of CMBs.

For visual detection of the CMBs different criteria are applied [12] :

- hypointense spherical lesions on T2*-weighted MRI;

- absence of signal hyperintensity on T1-weighted or T2-weighted sequence;

- location in the brain parenchyma (grey or white matter) and not in cerebrospinal fluid (CSF);

Visual rating of CMBs is a tedious task, as it requires from the observer to scroll through hundreds of thin slices in order to detect small foci of hypointensity and to distinguish them from linear shaped vessels. Also, there are other structures such as, blood vessels, calcium and iron deposits which may mimic CMB and make the task even more difficult [12]. Therefore, fully- or semi-automated methods for CMB detection are of large interest, especially within the context of large imaging studies. Due to the importance of CMBs and a lack of system used for developed for detecting these lesions, the aim of this thesis is detecting the CMB structures in the brain using different image sequences of MRI.

We propose a framework for semi-automated detection of CMBs in T2*- and PD-weighted MR images to decrease the workload for the radiologist. We approached the problem of distinguishing CMBs from the other structures in the brain by using a multistage classifier which uses different features. Our approach is similar to computer aided detection system in other applications [21, 31]. In each classification step we remove the candidates which are not CMBs and are detected as Non-CMBs structures (false-positives) and feed the rest of the candidates as the input to the next step. The candidates which are remained at final are considered as true candidates. We validated our system using the labeled and segmented CMBs to find the true-positives and false-positives and base on these measurements we evaluated our system.

The radiologists' desired system is the one with sensitivity over 90%, with on average, less than five false positives per subject as CMBs candidate. Our system succeeded to fulfill this desire by having 90% sensitivity and four number of false positive detected as CMBs using an multistage detection algorithm.

An overview of the medical background is presented in section 2 where the cerebral microbleeds and their disease-associated risk factors are described. In section 3 the related works are discussed. The problem statement is described in section 4 and our method is discussed in section 5 starting with preprocessing steps in sections 5.1 and 5.2. Then the feature extraction is described in section 5.3 and the applied classifiers are discussed in section 5.4. The image rigid registration technique is described briefly in section 5.5. The anatomical location of the microbleeds candidates which are extracted to help the researchers to find the relation among the location of CMBs and its effect on brain function are also described in details in section 5.6. In section 6, the results are presented and then discussed in section 7. We close our thesis with the conclusion part in section 8.

2 Medical background

Cerebral microbleeds, mimics and detection criteria

Cererbral microbleeds (CMBs), also called hemosiderin deposits, are iron-storage complex that are caused by leakage of red blood cells from small blood vessels. The CMBs are important prediction factors for dementia and cerebrovascular disease [12]. They can be detected in T*2-weighted gradient-recalled-echo (GRE) magnetic resonance (MR) imaging sequences as round hypo-intense structures within a certain size interval [12]. The hemosiderin has paramagnetic properties which means it can get the magnetic property in the presence of an external magnetic field. The list of the identification criteria can be found in the Panel 1. Despite of various features of CMBs as MRI lesions (abnormal tissues), there is an accepted agreement which indicates CMBs as small black round lesions with blooming affect on T2*-weighted MRI. In Figure 1(a), one can see an example of a microbleed in T2*-weighted MRI sequences. The size of the CMBs in diameter is classified in the interval of 2-10 mm [12]. A certain size of the microbleeds cannot be defined since in different studies, this size varies because of various MR imaging systems with different imaging parameters e.g. filed strength. On the other hand the microbleeds and marcobleeds define a cut-off point as 5.7 mm which is located in the size interval of 5 mm and 10 mm. The presence of the CMBs can be detected by expert radiologists by looking at continuous slices of the MRI sequences to be able to visualize the spherical structure of CMBs.

There could be also other structures in T2*-weighted MRI which mimics the CMBs and are shown in Figure 1(b). Calcium and iron deposits which are hypointense small foci in T2*-weighted MRI mimics the CMBs because of their spherical structures. These deposits usually are located in the basal ganglia. Calcification may be located in the choroid plexus, the pineal gland and the lobar locations in the brain. The computer tomography (CT) modality can be used for identification of clacification [12].

Flow voids are the other structures which mimics the CMBs. They have linear structures mostly located in cortical sulci which can be also visualized on T2-weighted spin-echo and GRE sequences. These properties can make them distinguishable from the CMBs [12]. Partial volume artifacts, from bones (caused by the air in the sinuses), are the other artifacts which especially in the orbit and mastoid mimics CMBs. Cavernous malformations which look like ringing artifact in T2*-weighted MRI are other CMBs mimics. Since they are present in both T1- and T2- weighted MRI we can distinguish them by observing the both sequences [30]. Metastatic melanoma which have the same features as CMBs, can also be distinguished by observing them in T1-weighted MRI as high intensity structures.

Panel 1: Recommended criteria for detection of cerebral microbleeds [12]

-Black lesions on T2*-weighted MRI,
-Round or ovid lessions (rather than linear or tubular),
-Blooming effect on T2*-weighted MRI,
-Absence of signal hyperintensity in other sequences as T1-weighted or
T2-weighted and PD-weighted sequences,
-Surrounded by brain parenchyma (at least half of the lesion),
-Distinguishable from other potential mimics as iron
or calcium deposits, bone or vessels,
-Clinical history excluding traumatic diffuse axonal injury.

Association of CMBs with risk factors and disease states

The risk factor of the reoccurrence of ischemic stroke and interacerebral bleeding increases by the patient with cerebral microbleeds [22]. The high blood pressure, age [23] and low serum cholesterol concentrations [26] are common factors for a prevalence of CMBs or their numbers.

The CMBs could also have direct effects on neurological function, cognition, and disability. The presence of the CMBs can lead to damages in brain tissues which can cause to dys-functionalities [8]. In this case the anatomical location of CMBs can play an important role on their effects on brain function. So in this study we also are interested to know the anatomical location of the CMBs. For this reason a brain atlas was used to define anatomical locations in the brain.

(a)

(b)
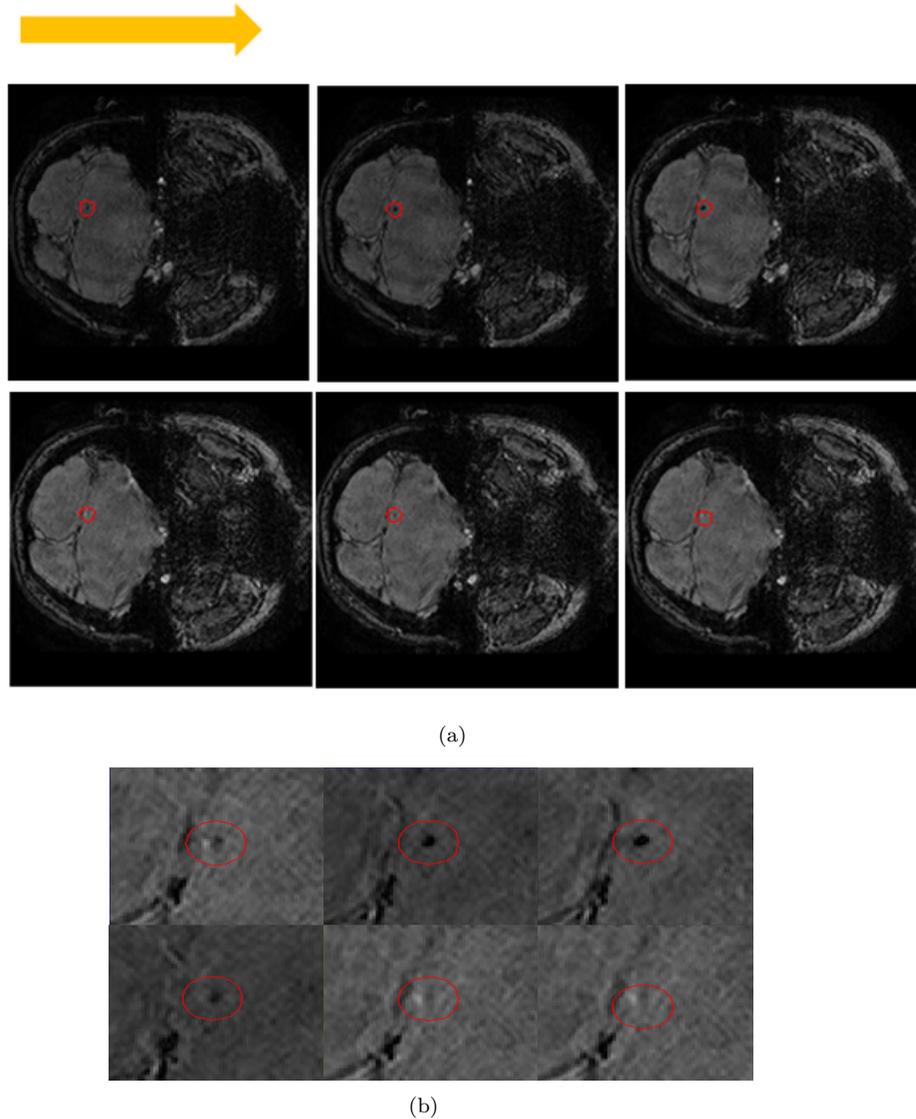
Figure 1: Microbleed (a) the axial view of T2*-weighted brain MRI in 6 continuous slices which shows the CMBs as black regions of the image indicated by red circle. The spherical character of CMBs can be seen by observing these continuous slices (The slices are sorted from left to right) (b) for a more clear observation the region of interest where the CMBs are located is enlarged. The arrow depicts the direction for following the sequences from top to bottom.
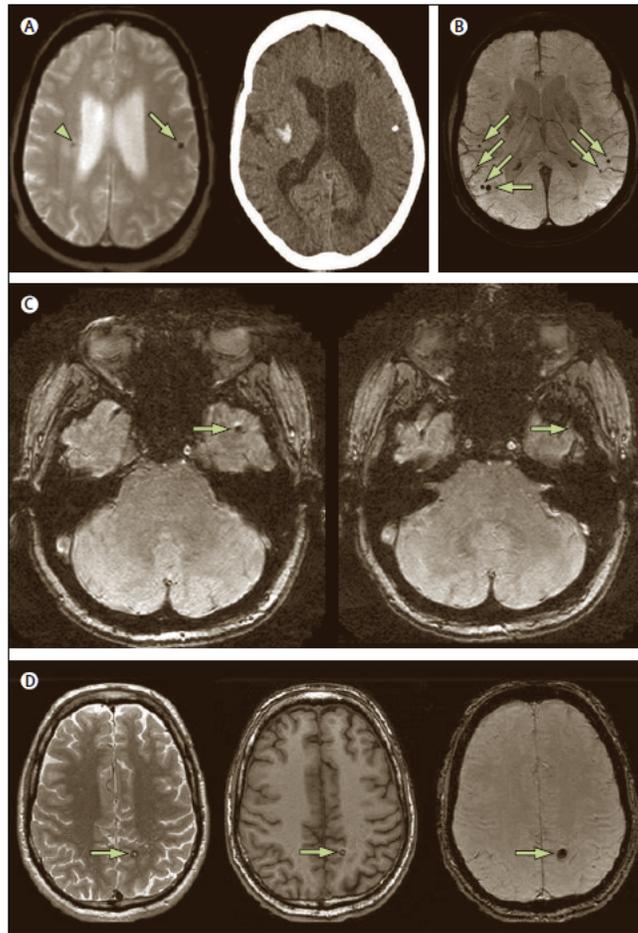
Figure 2: (A) calcification in T2*-weighted MRI (left panel; arrow) mimics CMB. The CT scan shows the calcification as a high density area (right panel). (B) Axial T2*-weighted MRI scan shows CMBs (indicated by arrows) close to blood vessels. The vessels can be distinguished from CMBs by their tubular shape.(C) Partial volume artifact in T2*-weighted MRI scan as a CMB mimic. T*-weighted MRI (left panel) shows a circular region of signal loss (arrow) which can be mistaken as a CMB. left sphenoid bone (righ panel) which shows the adjacent sphenoid bone (arrow).(D) Axial proton density-weighted (left panel), T1-weighted (middle panel), and T2*-weighted (right panel) MRI scans indicate a cavernous malformation (arrows) as a CMB mimic . The proton density or T2-weighted sequences distinguish these lesions from CMBs [12].

.

3 Related works

3.1 Semi-utomated detection of pulmonary nodules in helical CT images based on an improved template-matching technique

Y. Lee et al [18] proposed a computer aid diagnosis system to detect lung nodules in chest computed tomography (CT) images. In this method a template-matching technique based on genetic algorithm (GA), using Gaussian model of nodules as reference template, was used as initial candidate detection. GA was used to find the location of the nodules, which have spherical shapes, in the lung area. It is also used in choosing among different reference patterns for the proper template image.

To detect nodules located along the lung wall, a conventional template matching was employed. Because of the spherical shapes of lung nodules, semi-circular templates were used as reference patterns.

The threshold values which were determined experimentally for each feature were used for eliminating the FPs. By selecting the initial candidates based on two template matching methods, thirteen features were used consisting of nine and four features in GA and in the conventional template matching respectively, to reduce the number of false positives. Some of these features can be listed as:

- Inverse Difference Moment (IDM) and Entropy;

- Mean and Standard Deviation;

- Area, Circularity and Irregularity;

- Contrast and Max Mean CT Value;

- Directional Variance and Directional Cross-Correlation of the Pixel Gradient.

3.2 Semi-automatic detection of small lung nodules on CT utilizing a local density maximum algorithm

Binsheng Zhao [32] presented a method for semi-automatic detection of small lung nodules in CT images. In this method a three steps algorithm was used to solve the problem which consists of:

- segmentation of lungs by finding a threshold in the density histogram of CT chest images combining with a morphological operation;

- initial candidate detection using a local density maximum (LDM) algorithm for detecting higher density structures in the lungs;

- reducing the number of FPs using the pre-knowledge of lung nodules features.

Five parameters are used in the algorithm are: the threshold step, threshold stop value, minimal density peak of local maximum, minimal size of local maximum, and the ratio specifying the change of object's volume to its surrounding box's volume. The threshold values for parameters are determined experimentally.

Regards to prior knowledge of lung nodules, different features are calculated to reduce the number of FPs. These features consist of (1) candidate volume (2) the ratio of length of candidate in z direction to the maximum length in x and y direction (3) the ratio of maximum length of candidate to the minimum length in x and y direction.

This work used three steps of preprocessing, initial candidate selection and classification stage based on the nodule features to approach the problem. The presented model is similar to our model with some difference in the methods they used.

3.3 Semi-automated detection of pulmonary nodules from low-dose computed tomography scans using a two-stage classification system based on local image features

Murphy et al [21] proposed a method for semi-automatic lung nodule detection based on local image features. Their method consists of a preprocessing step including down-sampling and lung segmentation which is followed by a multi-step classification for lung nodule detection. The shape index (SI) and curvedness (CV) features which are three-dimensional local image features are used to detect initial nodule candidates. The seed points which are selected by thresholding of CV and SI values are grown by hysteresis thresholding to create the connected components. These connected components are filtered by size to keept the desired components in a certain size interval.

For reducing the numbers of false-positives 8 simple geometrical- and intensity-based features are used for feeding to a kNN-classifier. In the next step 22 features relating to gradient orientation and magnitude, grey value intensities, SI and CV values and geometric cluster properties are calculated and feed to a second kNN-classifier.

4 Problem statement and main contribution

Our conclusion from section 3 shows, up to this date, there is no system for semi-automatic detection of CMB lesions. However the researchers developed different computer aided detection (CAD) systems for other kind of lesions in medical images. In section 3 we mentioned some of these works which are related to semi-automatic detection of lung nodule lesions in CT images. These nodules have the same spherical structures as CMBs. However our research is different from the works mentioned in section 3 since it is focused on brain instead of lung organ. In Figure 3 a CMB in brain MR and a nodule in lung CT are shown. In Figure 3 (a) the air filled in the lungs can be seen as dark area in the CT image. In this image the airways and vessels are shown as branch-like white structures and the array indicates a lung nodule. In Figure 3 (b) the CMBs is shown with an array as a dark circle and the blood vessels are branch-like dark structures. The background (white matter) of the vessels in Brain MRI is bright while the background (air) of the vesels and airways in lung CT is dark.

(a)   (b)
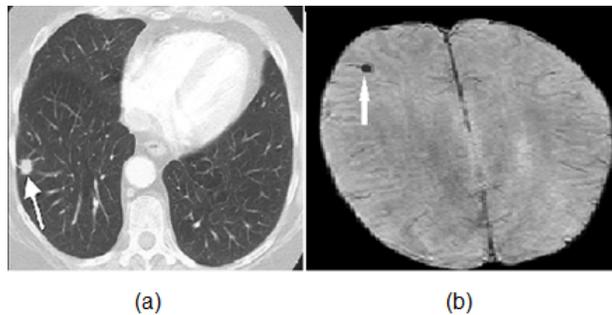
Figure 3: (a) In a lung CT slice of chest, the lung nodule is pointed out by a white array. The heart, spinal cord and the other organs are shown in the center of the image where the lungs are separated. (b) In a brain MR image a CMB is indicated by an array. The intensity difference of the CMB and vessels as dark structures in MR and the lung nodule and vessels as white structures in CT can be seen in these images.

.

The characteristics of CMB spherical structures in MR images are different from the lung nodule spherical structures in CT images because of the difference in imaging techniques and the chemical difference of these structures. The works related to lung nodule detection applied only one classifier which is k-nearest neighbor (kNN) classifier. In our work for CMBs detection we analyze different classification methods and compare their performances to find the better solution for our problem.

In chapter 3 we have introduced the different structures that mimics the

CMBs, among others vessels. Therefore a goal is to implement a technique which can distinguish the CMBs from vessels on base of structural analysis in 3D. It is important to mention that even though the vessels are continuous tubular structures, but because of the uncertainty and resolution of imaging techniques, they may look as discontinuous structures in 3D image. The proposed technique overcomes this problem.

Another thesis goal is to distinguish noisy structures in the image. These structures, which can be any part of the brain structure or imaging artifacts, are the same as CMBs in their size and intensity features but their shape features are different from CMBs.

CMBs can be distinguished from vessels by analysis of their structural difference in the 3D image. In two dimensional (2D) MR images, the CMBs look like circular plates but in 3D they have spherical shapes (structures). The vessels in 2D are also circular while in 3D they have tubular structures. The structural difference in 3D can be used for distinguishing the noise structures from CMBs, since the noise structures in the image do not have any tubular or spherical patterns. Most of these noisy structures have plate-like patterns.

The other way to distinguish between CMBs from the other structures, like vessels and noisy structures, can be based on analysis of such features as CMBs size and intensity. The CMBs are low intensity (dark) structures in the T2* MR images while many noisy structures are high intensity regions in the image. According to chapter 2 the size of the CMBs in diameter is in the interval of 2-10 mm. This feature can be used to distinguish the other artifact, like vessels and noise, which are out of this size interval.

As it was mentioned in section 1, the radiologist use T2*-weighted MRI images for distinguishing the CMBs in the brain MR images. However since the blood vessels are visible in Proton Density (PD) sequences as high intensity value structures but CMBs are not visible in these sequences, the PD sequences can be used to distinguish between the CMBs and blood vessels. For this aim we need to find the corresponding points for the T2* and PD sequences.

The structure of the system for detecting the CMBs can be different from a single- to a multi-stage system. All the intensity, geometrical and structural features, can be used in any stage of a multi-stage system depending on our design. Our system should use the most important features, by selecting them, using some machine learning techniques.

This thesis introduces a system for semi-automatic detection of CMBs in the brain MR images. Our system contains three steps which are: preprocessing, classification and postprocessing. Each of these steps can be divided to different sub-steps.

The preprocessing step contains: image registration, soft (brain) tissue segmentation and initial candidate selection. The image registration techniques are used for finding the corresponding points on T2* and PD sequences of MR images. The brain MR images contain brain, skull and the background. The brain should be segmented from the other part of the image since the CMBs are located only in the brain tissues. For initial canidate selection, size and intensity features are used to pre-select the structures which can be CMBs. In this step all parts of the image which are not CMBs should be removed from our system while keeping all the CMBs.

After the preprocessing, two classification steps are used to detect the CMBs from the other structures in the image. Each classification step uses only one classifier in the final design of the system. However we investigated to apply different kind of classifiers in each step and by evaluating their performance, select the most suitable for our system. The simple features are used in the first classification step while the more complicated features are estimated in the second classification step. In each classification step the most important features should be applied for training our classifiers. For this aim different machine learning techniques can be used. The first classification step is used to remove false-positives while keeping the sensitivity high.

In the final stage a post-processing step was applied to determine the anatomical location of the microbleeds candidates. For evaluating our system, we will compare its performance with the manual segmentation of CMBs done by expert radiologists.

5   System Design

Our approach towards detection of CMBs is based on a supervised classification method using geometric, intensity-based and structure-based features. In Figure 4 the general model of our system is shown. T2*-weighted and PD-weighted MR brain images are the input of our system. The preprocessing step which is described in section 5.1 contains three steps. In the first step T2* and PD sequences of MR images are registered to find the correspondence points in these sequences, using a rigid registration method. The skull stripping and the initial candidate selection are two other steps in the prepsocessing step. After preprocessing, the classification step is described in section 5.2. The feature extraction, feature selection and learning classifiers are discussed in this section. The postprocessing step for indicating the anatomical location of the CMBs is described in section 5.3. The applied tool for analysing and evaluating the system results is mentioned in section 5.4.
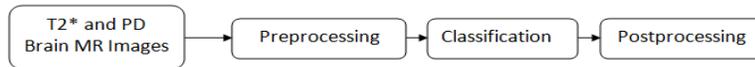
Figure 4: The generic system model

5.1   Preprocessing

Our preprocessing step can be divided in three stages: image registration, soft tissue segmentation and initial candidate selection. These stages are shown in Figure 5.
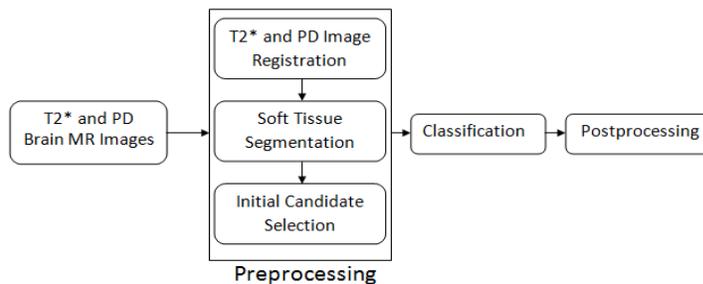
Figure 5: The details of preprocessing step in generic model

5.1.1   Image registration

In this section we describe primary nonrigid registration concepts and its application for our problem. The task of image registration is finding the transfor-

mation function $T(x)$ to perform spatial alignment of the target image I_t with the reference image I_r [17].

The MR imaging protocols can have different sequences such as: T1-weighted, proton density-weighted (PD) and T2*-weighted GRE (gradient-recalled echo). We have mentioned in section 2, the T2*-weighted MR images are used for distinguishing the CMBs. However in this thesis the features in PD and T2* sequences are used to distinguish CMBs. For this aim, we need to find the corresponding pixels in T2* and PD sequences by using different image registration techniques.

The similarity measure S is used for validation of alignment quality. The registration problem is an optimization problem where the cost function C is minimized w.r.t T [17]:

$$\widehat{T} = argmin_T C(T; I_r, I_t), \tag{1}$$

$$C(T; I_r, I_t) = -S(T; I_r, I_t) + \omega R(T), \tag{2}$$

where ω weights similarity against regularity and R is regularization term.

Similarity measures which are used in literatures can be divided into different categories [17]:

- **Sum of squared differences (SSD)**: only suited for the images which have the equal intensity distribution, i.e. for the images with the same modality.

- **Mutual information (MI)**: can be used less strictly as it assumes a linear relation between the intensity values of the images.

- **Kappa Statistic (KS)**: used for registration of binary images (segmentation). It measures the overlap of the segmentation.

- **Mutual information (MI)**: the relation between the probability distributions of the intensities of the target and reference images. The MI is suitable for the mono-modal and multi-modal applications.

5.1.2 Soft tissue segmentation

We implement skull-stripping to segment the soft tissue (brain) from skull. There is no presence of CMBs on the skull and with this hypothesis we can extract the skull without missing any microbleed in the subjects. Applied software and techniques for segmentation procedure are descirbed in section 6.

5.1.3 Initial candidate selection

The microbleeds are hyperintese elements of the image within specific size interval. This feature is used in the initial candidate selection. For image thresholding we started with a static fixed threshold value which is gained by evaluation of the distributions of both the intensity values and the sizes of the microbleeds. We extracted all the voxels in the image data which are labeled as CMBs to investigate the distribution of their sizes and intensity values. The procedure of candidate selection is discussed more in section 6.

5.2 Classification

In this section the details of classification steps: feature extraction, feature selection and learning classifiers, are discussed. The general model of the classification step is shown in Figure 6.
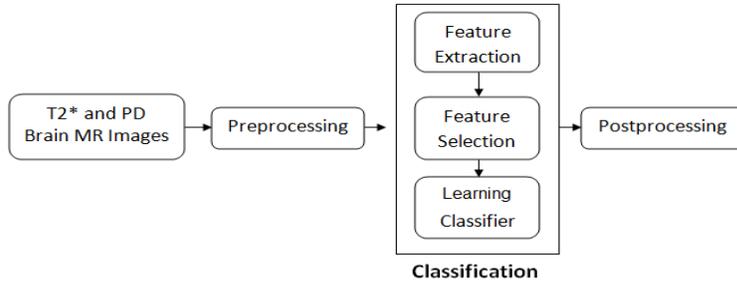
Figure 6: The details of classification step in generic model

5.2.1 Features extraction

In this section we describe all the features which can be used for distinguishing CMBs from false-positive candidates. A Feed-Forward and Feed-Backward features selection method is used to choose most significant features and reducing the computation time [5]. The area under the receiver operating characteristic (ROC) curve is used to evaluate the feature selection procedure [7]. The are different techniques to calculate this *area* which among them we can call uniform distribution, triangular distribution, beta distribution and gaussian distribution [3].

5.2.1.1 Geometric-based features

The following size and shape features which are gained from the binary images are [21]:

- The volume of the candidates: *vol*;

- The minimum, maximum and median length of bounding-box in x, y and z directions: $l_{min}, l_{max}, l_{med}$. The bounding-box in 3 dimensional images is the coordinates of the cubic border that encloses a region of interest;

- The ratios $\frac{l_{min}}{l_{max}}$ and $\frac{l_{max}}{l_{med}}$;

- Compactness1: vol/l_{max}^3;

- Compactness2: $vol/(l_x) * (l_y) * (l_z)$, where l_x, l_y, l_z are the length of the bounding box in x,y and z directions respectively.

5.2.1.2   Intensity-based features

The five features based on the intensity values of the candidates are extracted on T2*-MRI and Proton Density (PD) sequences.

- Mean grey value;

- Minimum gray value;

- Maximum gray value;

- Median of gray values;

- Standard Deviation of gray values.

These features are computed on the voxels of the bounding-box of initial candidates. These bounding-boxes are also dilated to cover their neighboring voxels. We estimated the intensity features on these dilated bounding boxes which makes the total intensity-based features up to ten.

5.2.1.3   Scale-Space representation for feature extraction

In the real-world, objects exist in different scales as meaningful substances. Scale-Space theory is a framework for representing the images at multiple scales to find the relevant scale. The scale-space method has been used for many image processing problems especially for feature detection in images.

The feature detection techniques based on the automatic scale detection method [19] is used in this work. As mentioned in the introduction section, the CMBs have spherical structure. This is used as the main feature for distinguishing CMB from non-CMB structures. The Hessian matrix which is a second order derivative matrix can be used for detecting different geometrical structures as linear-, spherical- or plate-like. In the following part we describe different Scale-Space features which are Determinant and Trace of Hessian matrix and multi-scale vessel enhancement.

Determinant and Trace of Hessian matrix

The scale-normalized trace and determinant of the Hessian matrix are used to detect Blob structures in the image [19]. The Laplacian of Gaussian (LoG) is one of bases of blob detectors which in two dimensional spaces can be described as:

$$G_\sigma(x, y) = \frac{1}{\sqrt{2\pi\sigma^2}} exp[-\frac{x^2 + y^2}{2\sigma^2}] \tag{3}$$

where σ is the scale which is the standard deviation. The LoG at scale σ is used as a kernel to measure the intensity difference of inside and outside of the kernel in the range $(-\sigma,\sigma)$ and in the derivative direction.

The convolution of an image matrix $I(x,y,z)$ with the Gaussian function $G(x,y,z;t)$ in scale or standard deviation t will result in a scale-space representation $L(x,y,z;t)$ as is shown here:

$$L(x, y, z; t) = G(x, y, z; t) * I(x, y, z) \tag{4}$$

where $(*)$ denotes the convolution operator. If $L(x_1, x_2, ..., x_n)$ be a given real-valued function, the Hessian matrix as a second-order partial derivative can be shown as:

$$H(L(x_{i,j})) = D_i D_j L(x_{i,j}) \tag{5}$$

where H is the hessian matrix of the function L and parameter x, and D_i and D_j are derivative operators over the function L with respect to the arguments i and j of this function. The determinant and trace of H in three dimensions can be shown as:

$$det H_{norm} L(x, y, z; t) =$$

$$t^{2\gamma}(L_{xx}L_{yy}L_{zz} + 2L_{xz}L_{xy}L_{yz} - L_{xx}L_{yz}^2 - L_{zz}L_{xy}^2 - L_{yy}L_{xz}^2), \tag{6}$$

$$trace H_{norm} L(x, y, z; t) = t^\gamma(L_{xx} + L_{yy} + L_{zz}) \tag{7}$$

where $\mathbf{det}HL$ and $\mathbf{trace}HL$(x,y,z;t) denotes the determinant and trace of Hessian matrix of function L, respectively. The constant t is the scale or variance in the Gaussian kernel g and γ is the normalization parameter which varies in different application. For the blob detection we set it as $\gamma = 1$.

The blob points can be detected by:

$$(x, y, z; t) = \mathbf{argmaxlocal}_t(det H_{norm} L(x, y, z; t)) \tag{8}$$

$$(x, y, z; t) = \mathbf{argmaxlocal}_t(trace H_{norm} L(x, y, z; t)) \tag{9}$$

where $argmaxlocal_t$ is the argument of the local maximum in scale t.

Multiscale vessel enhancement

We applied several second order image structure features computed at different scales. These features can be used to distinguish spherical structures from tubular or plate-like structures [10, 20, 27]. In addition, we used automatic scale selection [19] to detect the strongest blobs and vessel-like structures in the image.

The Taylor expansion is used for approximating the local behavior of the image with a higher order which for point x_o and its neighborhood [10],

$$I(x_o + \delta x_o, s) \approx I(x_o, s) + \delta x_o^T \nabla_{o,s} + \delta x_o^T H_{o,s} \delta x_o \tag{10}$$

where the $\nabla_{o,s}$ indicate the gradient vector. The third part of the formula $(\delta x_o^T H_{o,s} \delta x_o)$ is the directional second order derivative (Hessian matrix).

The second order structure of the image can be shown using the definition of the eigenvalues:

$$H_{o,s} u_{s,k} = \lambda_{s,k} u_{s,k} \tag{11}$$

which gives,

$$u_{s,k}^T H_{o,s} u_{s,k} = \lambda_{s,k} \tag{12}$$

where $\lambda_{s,k}$ is correspond to the k-th normalized eigenvector $u_{s,k}$ of the Hesian matrix H at scale s. The second order structure of the image can be shown as an ellipsoid whose axes directions are determined by the eigenvectors direction of the Hessian matrix and the length of its radii by the eigenvalues magnitude. The assumption for the order of the eigenvalues is that:

$$|\lambda_1| \geq |\lambda_2| \geq |\lambda_3| \tag{13}$$

In Table 1 the relation of geometrical structures with the eigenvalues λ_k in 3D is shown.

Table 1: Relation of geometrical structures with the eigenvalues λ_k in 3D. (H=high, L=low, N=noisy, $+/-$ indicate the sign of the eigenvalue) [10].

λ_1	λ_2	λ_3	orientation pattern
N	N	N	noisy, no preferred direction
L	L	H^-	plate-like structure (bright)
L	L	H^+	plate-like structure (dark)
L	H^-	H^-	tubular structure (bright)
L	H^+	H^+	tubular structure (dark)
H^-	H^-	H^-	blob-like structure (bright)
H^+	H^+	H^+	blob-like structure (dark)

The equations (12) and (13) show an ideal tubular and spherical structures in a 3D image, respectively:

$$|\lambda_1| \approx 0, |\lambda_1| \prec\prec |\lambda_2|, \lambda_2 \approx \lambda_3, \tag{14}$$

$$0 \succ\succ \lambda_1 \succ \lambda_2 \succ \lambda_3, \tag{15}$$

Based on these properties we define a ratio which is a measure of deviation from the blobness and can be shown as,

$$R_b = \frac{|\lambda_1|}{\sqrt{|\lambda_2 \lambda_3|}} \tag{16}$$

For the blob-like structures the R_b has its maximum value. The other ratio which can be used for distinguishing tubular-like from plate-like structures is,

$$R_a = \frac{|\lambda_2|}{|\lambda_3|}, \tag{17}$$

The changes in the image background are pretty small which means that the second derivative or the eigenvalues will be small as well. This leads to use the norm of the Hessian by using Frobenius matrix norm which gives us the measure of second order structures,

$$S = \|H\|_F = \sqrt{\sum_{j \leq D} \lambda_j^2}, \tag{18}$$

where D indicates the image dimension. The value of S is close to zero where there is no change in the intensity value in the image which means that the eigenvalues are small. Base on equation (14), (15) and (16) a vesselness measure can be defined as,

$$V_o = \begin{cases} 0, & \text{if } \lambda_2 > 0 \text{ or } \lambda_3 > 0 \\ (1 - exp(-\frac{R_a^2}{2\alpha^2}))(exp(-\frac{R_b^2}{2\beta^2}))(1 - exp\left(-\frac{S^2}{2c^2}\right)), & otherwise \end{cases} \tag{19}$$

α, β and c are control parameters for R_a, R_b and S respectively. In our work α and β are set to 0.5 and C is set as the half of the maximum value of Hessian norm [10]. This vesseleness measurement should be computed in different scales and the maximum of all scales s is taken as one of the features for our problem.

To summarize this section the list of the features used for the second classification step are listed in the following. All these features are calculated in three different standard deviation (scales).

- Eigenvalues for the Hessian matrix: λ_1, λ_2, λ_3 and $|\lambda_1| \geq |\lambda_2| \geq |\lambda_3|$;

- The maximum, minimum and median of the eigenvalues for the Hessian matrix;

- The ratio of maximum, minimum and the median of the eigenvalues for the Hessain matrix;

- The maximum of the determinant of Hessian matrix;

- The maximum of the trace of Hessian matrix;

- The maximum of the determinant of the multiplication of Hessian matrix [9].

- The maximum of the trace of the multiplication of Hessian matrix.

- The maximum of Vesselness;

The features are the local image descriptors of bounding-box's center-point for each candidate. The settings of parameters for these features is discussed in section 6.

5.2.2 Feature-space dimensionality reduction

In the following part we discuss about different dimensionality reduction techniques including feature selection and feature extraction [2, 5]. We analyse the results of these techniques in section 6 and the best performing method will be used in the final solution for our system.

The complexity of the classifier can increase by adding more features to the system, but it can also help to improve the performance. Based on the curse of dimensionality rule in machine learning, it has been observed in different machine learning problems that adding more features to the feature space may cause the performance get worse rather than better, beyond a certain point [15]. This is called overfitting or overtraining.

In machine learning problems it is accepted to keep the number of training data 10 times more than the number of features to avoid overfitting. Overfitting can be solved by combing or redesigning the features. This can be done by taking a subset of features using different feature selection or feature extraction methods which are discussed at the following sections.

Feature selection

Feature Selection is one of the methods used to reduce the dimension of the feature vector by gaining the highest sensitivity. This is one of the techniques for avoiding the curse of dimensionality which was discussed. The feed-forward and feed-backward feature selection techniques, FFFS and FBFS respectively, are another two different approaches [5]. The FBFS start with all features and reduces the dimension by removing them one by one. This continues while removing the features that increase the errors [2, 5]. The FFFS method starts without any features and adds the ones which decreases the errors significantly. This will continue while adding the other features that do not improve the performance or reduce the error significantly.

Feature extraction

The other dimensionality reduction technique is feature extraction. Feature extraction is mapping or transforming the feature vector to a lower dimension vector to extract the most relevant information from the data. There are different feature extraction methods of which some of them are: principal component analysis, semidefinite embedding, nonlinear dimensionality reduction, independent component analysis and partial least squares.

For our work, the Partial Least Square Mapping (PLSM) is a statistical method used as an feature extraction method [28]. In this technique the linear regression model can be found based on projecting both the predicted and observable variables to a new space. In contrast with principal component analysis which determine the main components of the covariance matrix of the observable data, in PLSM the covariance matrix of predicted variables are also considered.

5.2.3   Learning classifiers

In this part we describe the different classification methods which are tested in this work. The performance of these classifiers will be analyzed in section 6 and the best performing classifier will be used for the final model of the system.

The classifiers can be divided in two categories as discriminative and generative classifiers. In discriminative classifiers the parameters of class posterior $P(Y/X)$ is estimated directly, whereas generative classifiers estimate parameters for class priors $P(Y)$ and class conditional density $P(X/Y)$ [2]. Linear discriminant, quadratic discriminant and parzen classifiers are generative classifiers and KNN, a support vector and artificial neural network are the examples for discriminative classifiers. In the following part, the classifiers which are used in this thesis are discussed [5].

Support vector classifier

In Support Vector Machines (SVMs), the learning involves optimization of the convex function [2, 5]. The SVMs are one of the main solutions for different classifications and regression problems. In SVM, the dependency of the hypothesis on the data through the support vectors makes the model more understandable for the observer. The SVM exhibits a good generalization unlike the Artificial Neural Network (ANN) where the dependence of the data is not very clear.

Actual motivation for the subject of SVM is a convex optimization problem with one solution [2]. This unique solution to the problem is in contrast with the ANN, where the convex function may have several local (false) minima and the difficulty is finding the global minima [2]. Unlike the ANN, where we have many tuning parameters like the number of hidden nodes, layers, etc., in SVM

we are limited to a few parameters like the set margin and the kernel.

In binaries or two class problems we have:

$$\Theta = \{x_i, y_i | x_i \in \Re, y_i \in \{-1, 1\}\}, i = 1, ..., m, \tag{20}$$

where x_i is the input data and each data has a labeling value $y_i = \pm 1$ where y indicates the class to which x_i belongs and the index i indicates the number of data or labels.

The objective is to divide the data points with classes $y_i = 1$ and $y_i = -1$ with a margin hyper-plane. In Figure 7 an example of a support vector classifier for a binary classification problem is shown. The optimal hyper-plane, which is a unique solution and separates the training data with a maximal margin, can be formulated as a set of points,

$$w.x + b = 0, \tag{21}$$

where x is our data, b is the bias, w is the weight vector perpendicular to a hyper-plane and $(.)$ is dot product. we will consider a decision function of the form,

$$D(x) = sign(w.x + b) = \begin{cases} 1, & \text{if } w.x+b > 0 \\ -1, & \text{if } w.x+b < 0 \end{cases} \tag{22}$$

The distance between hyper-planes can be written as $\frac{2}{\|w\|}$ using geometrical properties where $\|w\|$ is the norm of w. Maximizing the distance between two parallel hyper-planes (margin), while they separate the classes of data by selecting w and b, is our solution to the classification problem. Therefore, the problem is minimizing the $\|w\|$ which is a difficult optimization problem since it is involved with the square root of $\|w\|$. We can change this difficult problem to a simpler one by replacing the square root form of norm of w with square form of w which changes the problem to a quadratic optimization problem [5]. Thus the optimization problem to find a saddle point will be,

$$min_{w,b} max_\alpha \left(\frac{1}{2} \|w\|^2 - \sum_{i=1}^{n} \alpha_i [y_i(w.x_i - b) - 1] \right), \tag{23}$$

where α_i indicates Lagrange multipliers. The solution to this optimization problem can be expressed as,

$$w = \sum_{i=1}^{n} \alpha_i y_i x_i, \tag{24}$$

where x_i are the support vectors as the solution to the equation. This solution results in the function which separates our classes to two different categories for our binary problem.
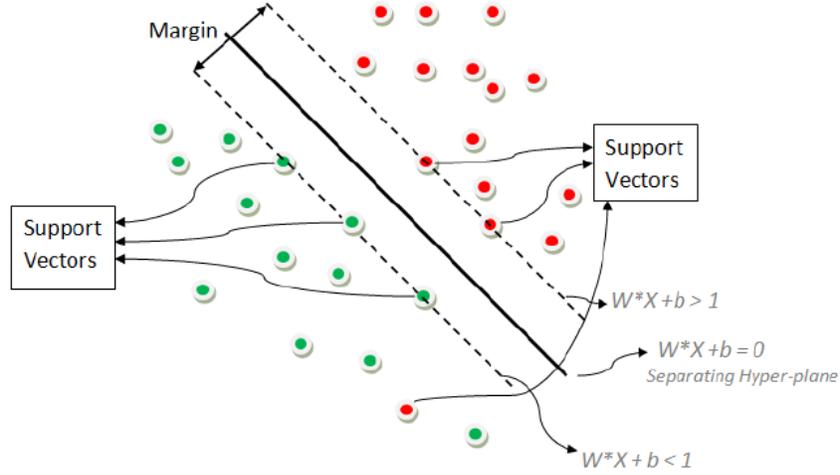
Figure 7: Support Vector Classifier; the support vectors which are shown in the figure indicate the margin of largest separation between the classes.

Linear discriminant classifier (LDC)

Finding some linear combinations between the set of features means finding the best separation of different objects in different classes, which is known as the linear discriminant classification (LDC) method. In this method we model each class of data by a Gaussian model. The class conditional density can be modeled as multivariate Gaussian:

$$f_{k(x)} = \frac{1}{(2\pi)^{p/2} \left| \sum_k^{-1} \right|^{1/2}} e^{\frac{-1}{2}(x-\mu_k)' \sum_k^{-1}(x-\mu_k)}, \qquad (25)$$

where x is the data or feature vector, p is the dimension of the data or feature vector, k indicates the number of classes, $\sum$ is the covariance matrix, and μ is the mean vector. In the case where we have LDC, all classes have the same covariance or on the other hand $\sum_k$ in the equation (25) for all k's is the same. This means that the gaussian distribution are shifted version of each other. By this assumption the discriminant function for LDC can be written as:

$$F_i(x) = \frac{-1}{2}(x - \mu_i)^{-1} \sum^{-1}(x - \mu_i) + ln\pi_k, \qquad (26)$$

27

where the quadratic term is $x^{-1} \sum^{-1} x$ which is not dependent on classes and can be deleted,

$$-\frac{1}{2}(x - \mu_i)^{-1} \sum^{-1} (x - \mu_i) = -\frac{\|(x - \mu_i)\|^2}{2\sigma^2}, \qquad (27)$$

which simplify the equation (26) to,

$$F_i(x) = -\frac{\|(x - \mu_i)\|^2}{2\sigma^2} + ln\pi_k. \qquad (28)$$

The classifcation rule can be written as,

$$G(x) = argmax_i(F_i(x)). \qquad (29)$$

From equation (28) it can be seen that the problem of classification is the Euclidean distance from the center of the class. The simplification of this equation shows F as a linear function where we can write this discriminant function (F) as a function of the features x_i which are the components of the feature vector $\bar{X}$,

$$F(\bar{X}) = \theta_0 + \sum_{j=1}^{n} \theta_i x_i, \qquad (30)$$

where i indicates the number of classes and θ_i and θ_0 are the components of the weight vector and bias, respectively. Assuming a feature space, for a two-class problem, the decision boundary or surface is defined by $F(\bar{X}) = 0$, which for a linear case of $F(\bar{X})$ is a hyper-plane. A sample of data belongs to class θ_1 when the $F(\bar{X})$ is more than 0, and belongs to class θ_2 when it is less than 0.

The LDC classifier can be used when there is no equal within-class frequency. The assumption for using the linear classifier is the equivalence of covariance matrices for normal densities of all classes. If this assumption cannot be fulfilled then using linear classifier will not give a good result. An example of LDC classifier is shown in Figure 8.

Quadratic discriminant classifier (QDC)

In the case which the classes has their own correlation structure, the linear discriminant function should be replace by quadratic function which is a non-linear function. QDC estimates the covariance matrix $\sum_k$ for each class k, separately [2] [5]. QDC function can be written as,

$$F_i(x) = \frac{-1}{2}log\left|\sum_k\right| - \frac{-1}{2}(x - \mu_i)^{-1} \sum^{-1} (x - \mu_i) + ln\pi_k, \qquad (31)$$

where the classification rule is the same as equation (30). By adding additional terms to equation (28) involving the products of X components, we obtain the quadratic discriminant function F,
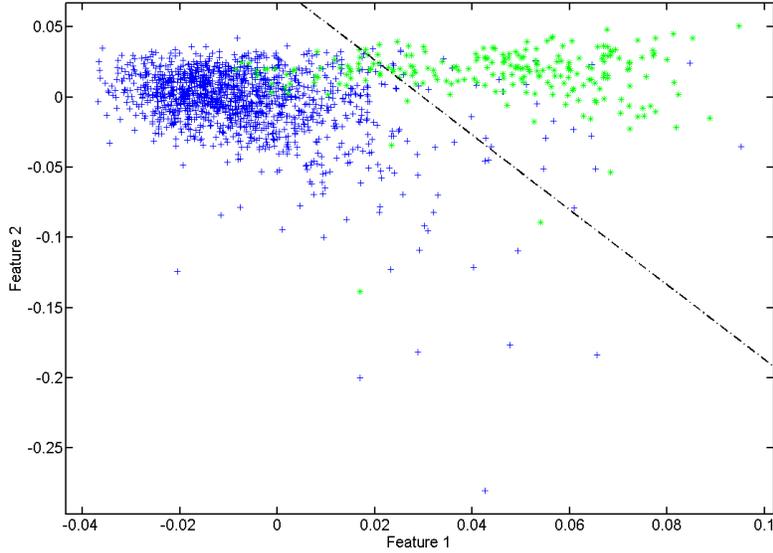
Figure 8: An example of a LDC classifier.

$$F(\bar{X}) = \theta_0 + \sum_{i=1}^{n}\sum_{j=1}^{n}\theta_{ij}x_ix_j, \tag{32}$$

The decision boundary or surface is defined by $F(\bar{X}) = 0$ which for a quadratic case of $F(\bar{X})$ is a second-degree or hyper-quadratic surface. Unlike LDC, there is no assumption for QDC that the covariance of classes be identical. Fitting the data using QDC is better than LDC but has more parameters to estimate. In Figure 9 the QDC classification is shown for a binary classification problem.

Parzen classifier

Parzen window method as a non-parametric classification or estimation technique uses a kernel as a weighting function for estimating the class conditional densities [2, 5]. This kernel has two properties as,

- $\int_{+\infty}^{-\infty} U(x)dx = 1$, which is a property of probability density function(PDF).

- $U(x)$ is a symmetric function so $U(x) = U(-x)\forall x$

The feature space can be divided to K d-dimensional hyper-cubic regions, R_k. The kernel U can be used for approximating of the number of samples

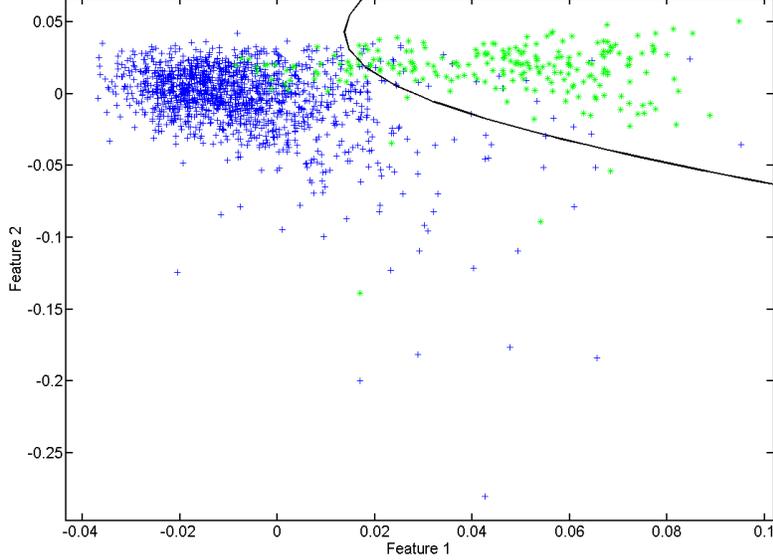Figure 9: An example of binary classification using QDC

which are located within R_k regions. The power density function (PDF) at point X can be formulated as,

$$f(X) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{K_n} U\left(\frac{X - X_i}{K_n}\right), \tag{33}$$

where n is the number of observations, U indicate the kernel, K_n indicate the kernel window width, X_i is independent and identically-distributed sample of random variables and the $\sum$ component indicates the number of samples located within the hyper-cube.

The Parzen classifier kernel parameters can be user defined or can be optimized by changing in the training process. There are different types of kernels in which some of them are uniform, Gaussian, cosine, and triangle. The Gaussian function usually is used as the kernel. In this case the U in the equation (31) can be written as,

$$U(x - x_i) = \frac{1}{(2\pi)^{d/2} s^d |\Sigma|^{1/2}} exp\left\{-\frac{1}{2s^2}(x - x_i)^T \Sigma^{-1}(x - x_i)\right\}, \tag{34}$$

where d, Σ and s indicate dimensionality, covariance matrix and smoothing parameter respectively.

30

The classification decision will be made by the calculated PDF function in equation (32). In a two class problem the test sample as a new input belongs to class **A** if $p_A f_A > p_B f_B$ otherwise it belongs to class **B**. Parameter p indicates the prior probability and f is the class estimated PDF in equation (32). The prior probability can be calculated based on the data frequencies in different classes. This can be reach by taking the ratio of the number of samples in each class and the total numbers of sample data.

5.3   Postprocessing

In the final stage a post-processing step was applied to determine the anatomical location of the microbleeds candidates. This helps the researchers to have further information about the CMBs location, to find the relation between the location of CMBs and its effect on brain function, using brain atlases.

The non-rigid registration of 20 atlas images consisting of 83 brain regions was used for this aim (Hammers atlas) [13]. The deformed atlas labels were then fused using a vote rule [14]. This segmentation was computed based on a T1-weighted image of the same subject using the Elastix software. Subsequently the segmented T1-weighted image was co-registered to the T2*-weighted image to propagate the labels.

5.4   Applied tools for analysing the system results

In this thesis, the area under the Free Receiver Operating Characteristic (FROC) graph is used as the criterion for selecting the best performing feature selection method and learning classifier [7]. A FROC graph is a tool for visualizing and characterizing the performance of a system at all decesion thresholds [1]. The analysis of the FROC curve provides us with the highest accuracy of the system since it estimates all the sensitivity and specificity combinations that the system is able to provide. Figure 10 shows two FROC curves for classifiers **A** and **B**. The horizontal and vertical axis are named as *average number of false positives* and *normalised sensitivity*, respectively. By assuming the total number of N false positives for M subjects in the dataset, *the average number of false positives* will be $\frac{M}{N}$.

The *normalized sensitivity* of the system can be written as:

$$Sensitivity = \frac{TP}{P}, \tag{35}$$

where TP and P indicate the number of detected true positives and the total number of positives, respectively. The value of normalized sensitivity is between 0 and 1. For comparing different classifiers we can reduce FROC performance to a single scalar value which presents the expected performance. A method for this aim is the area under the ROC curve (AUC). According to these curves,
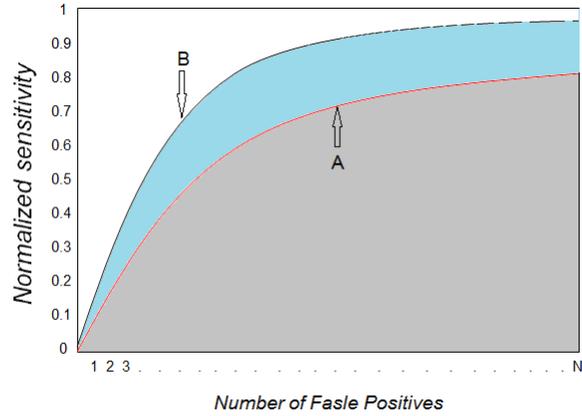
Figure 10: The FROC curves of **A** and **B** classifiers. The area under curve **A** (AUC_A) is shown with gray color. The area under curve B (AUC_B) is the sum of AUC_A added to the area of the region shown by blue color.

classifier **B** has greater area under the curve and therefore better performance [7]. There are different techniques to calculate the *area* which among them we can call uniform distribution, triangular distribution, beta distribution and Gaussian distribution [3].

6 Experiments and results

In this part we analyse different methods, described in section 5, and compare their performances to find the most suitable solution to our system. The applied hardware and software for this work is mentioned in section 6.1. The dataset which is used for our experiments is described in section 6.2. In section 6.3, image registration, skull-stripping and initial candidate selection, being a part of preprocessing, are described. The analysis of the results for the first and second classification steps are shown in section 6.4 and 6.5, respectively.

6.1 Applied hardware and software

Different software and toolboxes are used in this thesis. The image processing and statistics toolboxes of Matlab R2007b have been used as the main platform. The pattern recognition and machine learning toolbox, PRtools [6], is another software used for classification part. Mevislab 2.0 and Visual C++ are used for visualizing the data. In the preprocessing stage the FSL [24] software is used for skull stripping and Elastix [17] software for registration. An Intel Pentium 4, with 3.60 GHz CPU and 3.50 GB RAM, is used for running the programs.

6.2 Dataset for the experiments

In this work, the Rotterdam Study ErgoPlus database [26], which contains scans of a 1.5T (GE healthcare) scanner with eight channel head coil, is used. The scans which are MRI protocols for all participants include a 3D T1-weighted, a 2D proton density-weighted (PD), a 2D Flair (fluid-attenuated inversion recovery) and a 3D T2*-weighted GRE (gradient-recalled echo). The slice thickness of the scans is 1.6 mm for T2*-weighted GRE sequences which after zeropadding is reduced to 0.8 mm without any contrast material.

From this database, we use 237 labeled T2*-weighted GRE and proton density-weighed MR images all of which contain CMBs of the patients with the mean age of 79.2 years old, from a range of ages of 69.7 to 96.7 years old. The labeling was scored by one radiologist and one neurologist using PACS software (Picture Archiving and Communication System). The data are divided into two major subsets such as a *test* and a *train* set. The train and test sets contain 156 subjects ($\approx$ 448 CMBs) and 81 subjects($\approx$ 183 CMBs), respectively. The train set is further divided into two subsets of train and test set. The *train subset of the training set* is used for training and the *test set of train set* is used for regularizing the features and classifiers.

In our dataset Cohen's kappa, κ, coefficient which is a statistical measure of inter-rater agreement is used for scoring the manual labeling by labelers. All microbleeds in the 237 images in our dataset are labeled by expert raters. The inter-rater reliability, which is a score of agreement between the raters, is used for measuring the reliability of the rating by a human observer. In our data,
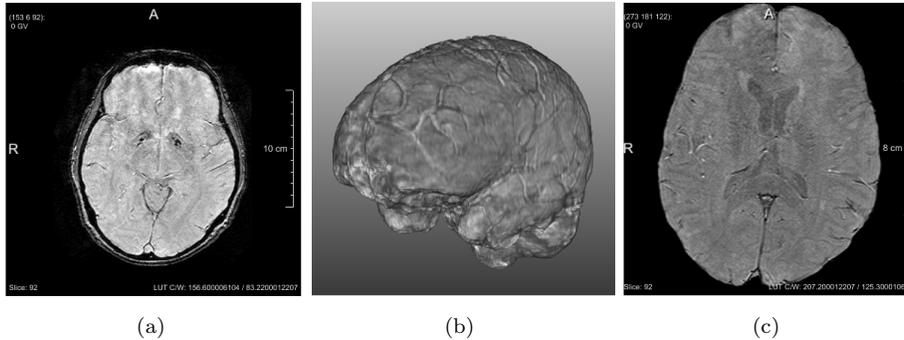
Figure 11: (a) An original T2* brain MR image (b) brain segmentation result for a subject using FSL (3D view) (c) segmented and cropped brain using FSL (2D view)

inter-observer reliabilities for visual detection of microbleeds were $\kappa = 0.85$ at the individual level and $\kappa = 0.82$ at the microbleed level respectively, therefore corresponding to very good agreement [26].

6.3 Evaluation of preprocessing methods

In this chapter we discuss the preprocessing step including image registration, skull-stripping and initial candidate selection which were described in section 5.1.

Different image registration techniques are mentioned in section 5.1.1. For our application the MI similarity measure is chosen since the registrations of multi-modal images which are T2- and PD-MRI images, are required.

For brain extraction, which was is described in section 5.1.2, among different software which are available, we decided to use a robust and accurate surface model approach which is implemented in Brain Extraction Tool (BET) in FSL package [29] [25]. To reduce the computation time, the images were cropped to eliminate the black backgrounds which contain no data of the brain tissues. The 2D and 3D views of the result of skull-stripping for a T2*-weighted MR image is shown in Figure 11. It can be seen in this figure that only soft brain tissues are left after implementing the FSL software for skull-stripping.

For initial candidate selection, as discussed in section 5.2, microbleeds are visible on T2*-weighted images as low intensity foci with a specific size range. A binary image was produced by thresholding the intensity values of the T2*-weighted image. To find the region of interests (ROI), the bounding-box of connected components were filtered by size to remove unrealistically large or small candidates. All three thresholds were optimized using the training set.

An 2D example of the thresholding procedure is shown in Figure 12. The original image (Figure 12(a)) will be converted to a binary image (Figure 12(b)) by thresholding the pixels with intensity values between 0 to 55, since the CMBs intensity values are within this intensity interval. The pixels with intensity values higher than 55 are set to 0 (black pixels in Figure 12(b)) and pixels with lower than this intensity are set to 1 (white pixels in Figure 12(b)). The white pixels in Figure 12(b) are the region of interests (candidates) in the image. These regions then are segmented by finding the connected components. The result of segmentation is shown in Figure 12(c). Different colors show different region of interest (candidates).

As it was described in the introduction section, the CMBs are in a certain size interval. This interval is between 2 mm to 10 mm, which is equal to 4 voxels and 64 voxels respectively, in our imaging data. By counting the number of pixels in every candidate, we can estimate the size of them. So we can filter the canidates which are out of this size interval. The result of thresholding the candidates by size is shown in Figure 12(d). After the prerocesssing step, on average, we have 652 candidates per subject for further classification on the training set. In the next section we will describe how to remove among these candidates the ones which are false positives while keeping the accurcay of the system high by keeping the true positives.
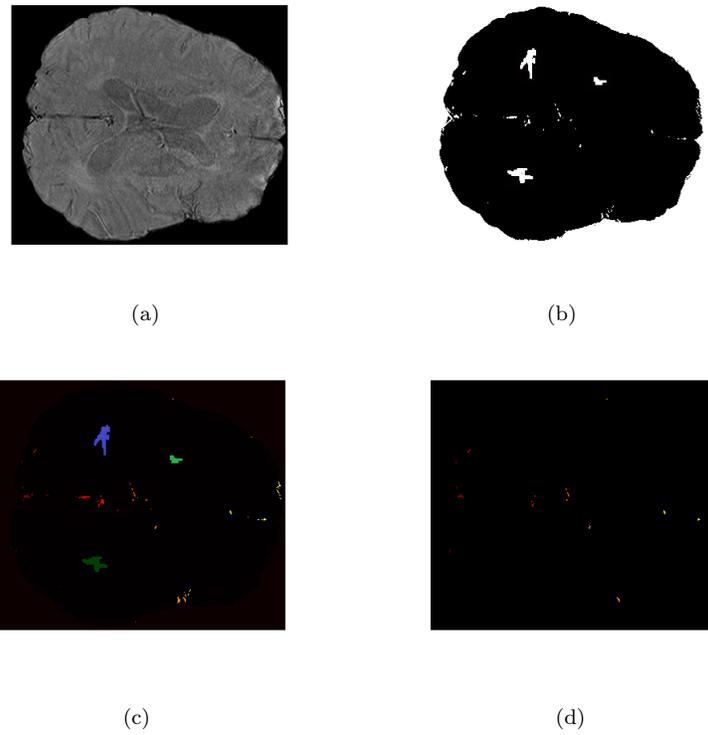
(a)

(b)

(c)

(d)

Figure 12: (a) original 2D MR brain image (b) binary image created by threshloding the original image (c) connected components (candidates) which their difference is shown by different colors (d) remained candidates after filtering the large and small regions.

6.4 Evaluation of classification methods

In this section the classification procedures described in section 5.2 are tested and analyzed. As it was mentioned in section 5.2.1 different features can be used for classifying the CMB from Non-CMB candidates. However based on a curse of dimensionality rule in machine learning, we need to reduce the feature space dimension. For these aim, two techniques which are FFFS and PLSM, are described. We test and analyze these techniques to find the better performing one, to apply it in our final system design.

A dimensionally reduced feature set, is fed to different learning classifiers which were described in section 5.2.3. By testing and analyzing the performance of these classifiers, the best performing one is selected to use it for the final classification step.

It was mentioned in section 4 that we plan to use a two stage, rather than a single stage, a classification step for decreasing the computation time of the system. The general diagram of the system including two classification steps is shown in Figure 13. The details of the classification blocks are described in the following sections.
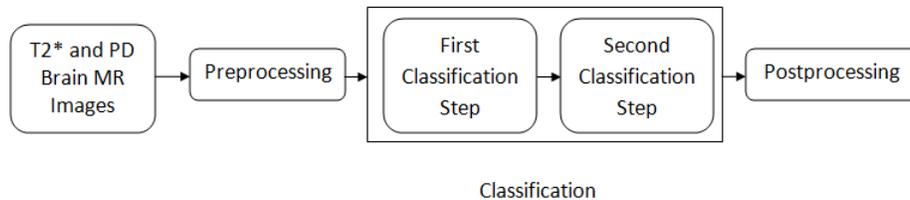

Classification

Figure 13: General diagram of the system wit two classification steps.

6.4.1 First classification

The aim of the first classification step is to remove the false-candidates, which passed in the preprocessing step, based on simple features. In this step different classifiers are tested which are LDC, QDC and Parzen. After analyzing their performances over train set, the best performing classifier is selected for this classification step. The block diagram of the first classification step is shown in Figure 14.

As discussed in the section 5.2.1, 18 geometry- and intensity-based features are estimated on the pre-selected candidates. The FFFS and PLSM techniques are tested on the train set to reduce the dimension of the feature set, and creating the subsets of train set which contains the most significant features. These
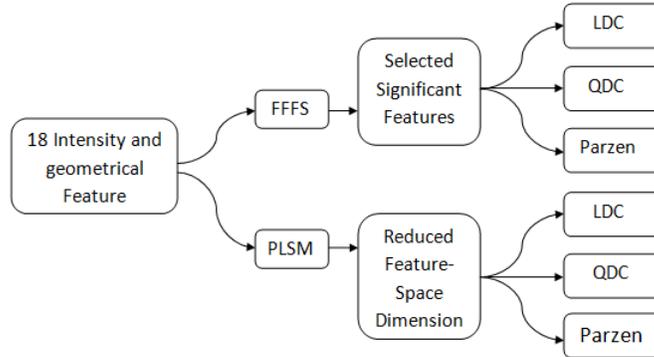
Figure 14: Diagram for first classification step evaluation

subsets are used for testing different learning classifiers, described in section 5, to find the best performing feature selection method and classifier.

The FROC curves for different classifiers are shown in Figure 15. In Figure 15(a) the FROC curves of classifiers by using all features are shown. In Figure 15(b) the FROC curves of classifiers after reducing feature dimension by PLSM method can be seen. In Figure 15(c) the FROC curves of classifiers by using FFFS feature selection method is shown.

According to section 4, the areas under the FROC curve (AUC) for LDC, QDC and Parzen classifiers, are calculated for comparing the performance of FFFS and PLSM. The results of AUC in Table 2 shows that FFFS method, while using the Parzen classifier, with the performance of 95% is the best one. Based on this analysis, we decide to use FFFS technique to find the most significant feature set and Parzen classifier for the first classification step, in the final system.

In the first classfication steps the threshold vlaue of the posterior probability for the learning classifier was selected to have the the best performance of the classifier while removing the maximum number of false candidates in this classification step. The best performance is achieved by Parzen classifier by removing 98% of the FPs while having 94% sensitivity. This sensitivity was set as the threshold for the first classification step in the test set.

Among 18 intensity- and geometric-based features, described in section 5.2.1.1 and 5.2.1.2, 6 features which are selected after FFFS method, in the order of their importance, can be listed as:

- Mean gray value;

- Compactness2;

- Standard Deviation of gray values on dilated bounding-box;

- Minimum gray value;

- Minimum length of bounding-box;

- Standard Deviation of gray values.

Table 2: Area under the FROC curves for LDC, QDC and Parzen classifiers using FFFS and SPLSM feature selection methods.
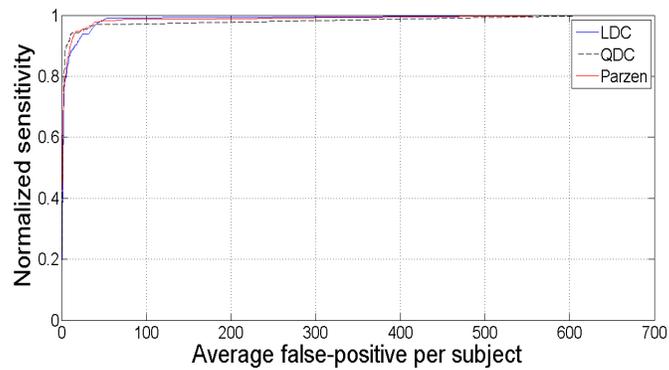
Feature Selection Technique	LDC	QDC	Parzen
No Feature Selection	0.88	0.91	0.89
SPLSM	0.89	0.92	0.90
FFFS	0.91	0.92	0.95
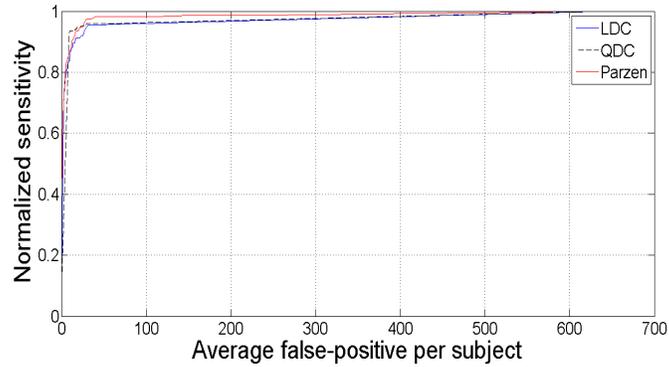
(a)
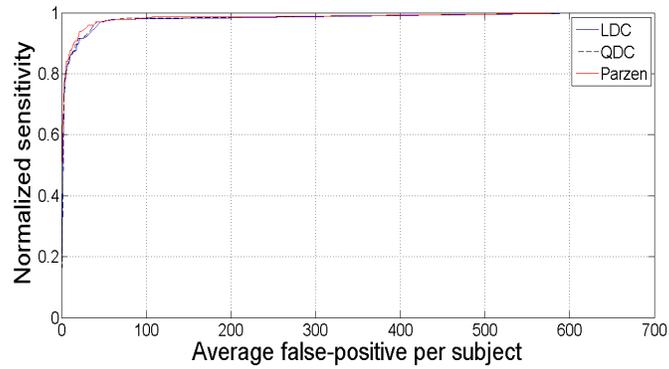
(b)

(c)

Figure 15: FROC curves of the training set for:(a) 18 geometrical and intensity features described in section 5.3.1 and 5.3.2 (b) using PLSM feature extraction method (c) using Feed-Forward feature selection method. The horizontal axes indicates the average number of false positives per subject while vertical axes indicate the normalized sensitivity.

40

6.4.2   Second classification

In this step different classifiers are tested which are LDC, QDC, Parzen and support vector classifiers. After analyzing the performance of the classifiers over train set, the best performing classifier is selected for this classification step evaluation. The block diagram of the second classification step, is shown in Figure 16 with more details.
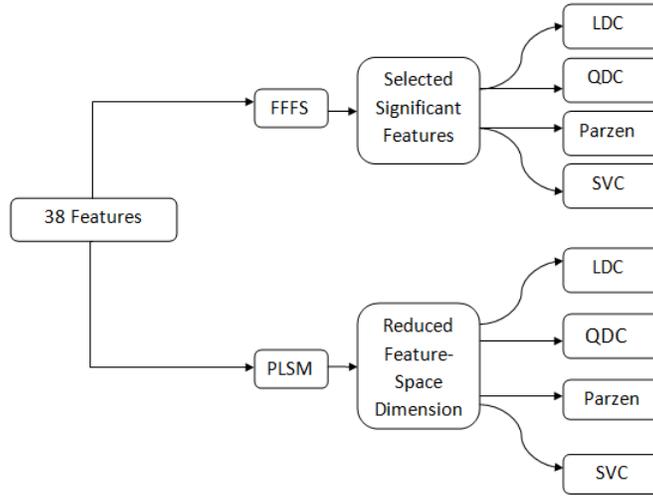
Figure 16: Diagram for second classification step evaluation.

The second classification step uses the features, which need more computation time compare to the features used in the first classification step, for removing the false-positives from the candidates which remain from first classification step.

As it was described in section 5.2.1.3, the LoG at scale σ can be used for multiscale features extraction. The images in our data have lower resolution in Z direction compare to X and Y directions. To avoid removing too much detail of the image in this direction, compare to the other directions, we need to use lower scale (σ). Three different scales with standard deviations of 0.5, 1, 2 in X *and* Y directions and 0.3, 0.5, 1 in Z direction, are selected.

In total 38 features discussed in the section 5.2.1.3 were measured on the segmented candidates. Among all of them, four features are intensity-based features on PD sequences and 34 features are based on second derivative of the structures in three scales. As an example, the features of the determinant and the trace of the Hessian matrix, for three standard deviations (scales), are shown in Figure 17 and 18 for CMBs and Non-CMBs candidates, respectively.
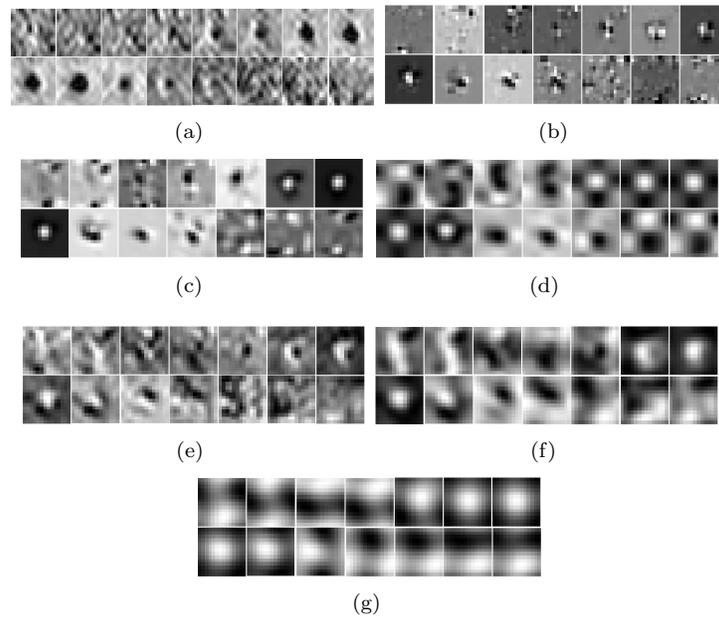
Figure 17: An example of CMB candidate (a) Original Image. Determinant of the hessian in (b) first (c) second and (d) third scales. Trace of the hessian in (e) first (f) second and (g) third scales.
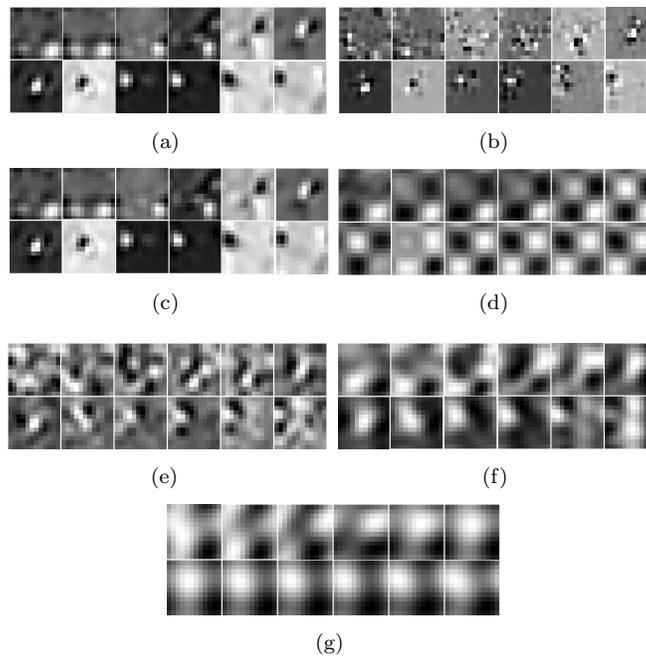
Figure 18: An example of Non-CMB candidate: (a) Original Image. Determinant of the hessian in (b) first (c) second and (d) third scales. Trace of the hessian in (e) first (f) second and (g) third scales.

As it was described in section 6.4.1, for comparing the performance of FFFS and PLSM feature selection techniques, their reduced feature sets are tested for training LDC, QDC, Parzen and SVC classifiers. The FROC curves for different classifiers are shown in Figure 19.
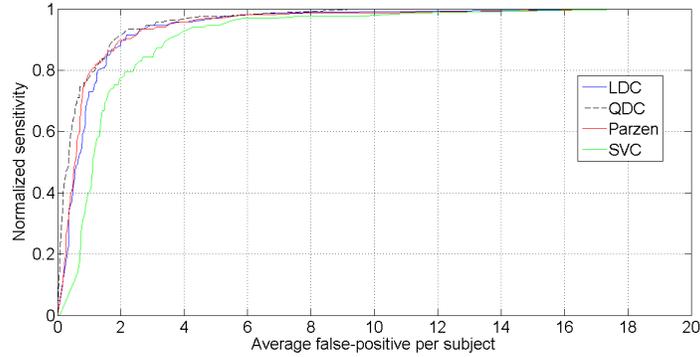

Figure 19: Area under the FROC for LDC,QDC and Parzen classifiers on the train set for the second classification step.

The results of AUC for these classifiers are shown in Table 3. This table shows the highest performance for FFFS method, while using the QDC classifier, with the performance of 98.6%. Based on these results, we decide to use FFFS technique to find the most significant feature set and QDC classifier for second classification step, in the final system. The results of these techniques on the validation (test) set are presented in section 7.

The features which are selected in the second classification step by applying FFFS can be listed as (in the order of importance):

- Maximum of determinant of the Hessian matrix in second scale;

- Maximum of gray values in PD;

- Maximum of trace of Hessian matrix in second scale;

- Maximum of the trace of Hessian matrix in the second scale;

- Maximum of the trace of the multiplication of Hessians matrix $(H * H)$ in the first scale;

- Minimum of the gray values in PD;

- Maximum of the trace of the multiplication of Hessian matrix $(H * H)$ in the second scale;

- Maximum of the eigenvalues in the second scale;

44

- (Minimum of the eigenvalues in the first scale) / (Median of the eigenvalues in the first scale);

- (Maximum of the eigenvalues in the second scale) / (Median of the eigenvalues in the second scale);

- Vesselness in the third scale;

- Maximum of the eigenvalues in the first scale;

- standard deviation of gray values in PD.

Table 3: Area under the FROC curve for LDC, QDC, Parzen and SVC classifiers using FFFS and SPLSM feature selection methods in second classification step.

Feature Selection Technique	LDC	QDC	Parzen	SVC
No Feature Selection	0.912	0.901	0.894	0.809
SPLSM	0.932	0.925	0.942	0.902
FFFS	0.951	0.953	0.986	0.921

7 System configuration and verification

In section 6, different methods were applied to find the best solution to our problem. In this section we present system consisting of the most suitable methods and validate the proposed system with the test set of our data. The best performing system is illustrated in Figure 20.

7.1 Preprocessing verification

In the preprocessing step, the soft tissue segmentation is performed on the validation set. The initial candidate selection, is done by using the thresholding values, applied in section 6.3, for size and intensity in T2* MR images. Afte this step, on average, 652 candidates per subject were detected for further classification on the test set. In the preprocessing step 1.6% of the CMBs are removed, which in this step, reduce the sensitivity of the system to 98.4%.

The large numbers of detected candidates are mainly vessels, noise and artifacts in the image, which most of them are removed in the next classification steps.

7.2 First classification step verification

In section 6.4.1, it was mentioned that the best performing feature selection method, FFFS, was used to reduce the dimension of the feature space. In the test set, after feature selection by FFFS method, 6 most significant features were used for further classification. It was also mentioned that, the Parzen classifier, shows the best performance on the train set. The classification results of the first classifier on the test set, are shown in Figure 21. By setting the sensitivity of 94% for the first classifier, which was selected in the train set, resulted in 19 false-positives, on average, per each subject, in our test set.

7.3 Second classification step verification

In section 6.4.2, it was described that in the second classification step, the FFFS technique can be used for feature selection and the best performing classifier was QDC for the train set. For the test set, after feature selection by FFFS method, 13 most significant features were used as the input of the QDC classifier in the second classification step. Figure 22 shows the FROC curve for the second classification step on the test set. Different threshold for the posterior probability of the classifiers can be set for having the best performance of the classifier, while removing the specific number of the false candidates. However for this step, the thershold of 98.6%, which was mentioned in section 6.4.2, is applied on the train set.
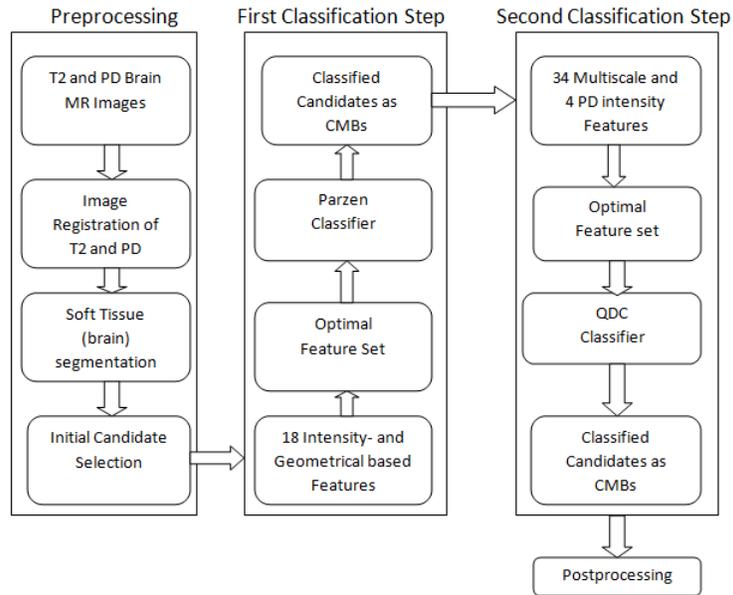
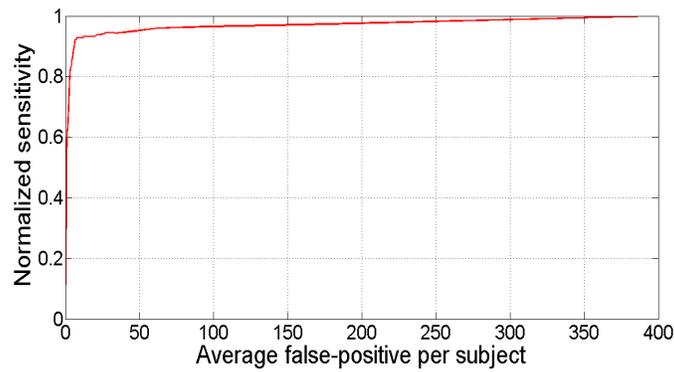Figure 20: Final solution to the system.

Figure 21: FROC curve for Parzen classifier in the first classification step, over test set, using 6 selected features by FFFS feature selection method.
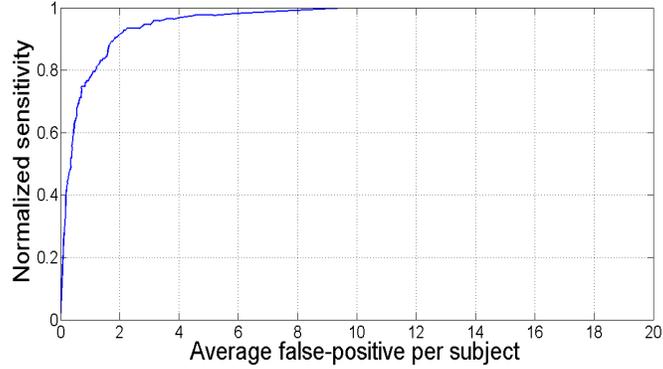
Figure 22: FROC curve for QDC classifier in the second classification step over test set, using 13 selected features by FFFS feature selection method.
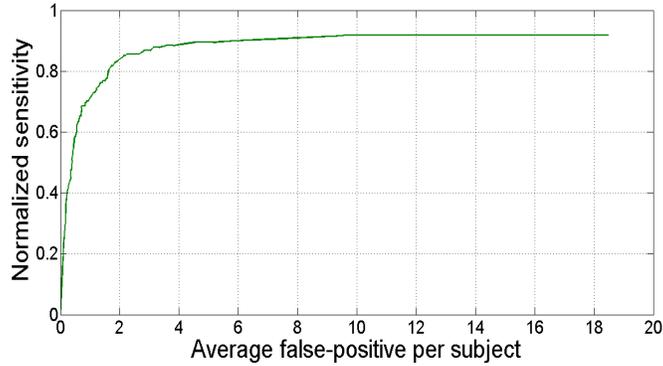

Figure 23: Cumulative FROC of the system.

7.4   Cumulative results

In each step of the proposed system, some of false candidates were removed. However, some of the true positives were also removed by the system, which this results in decreasing the sensitivity of the system in each step. The cumulative results can be reached by including this missing sensitivity in each step. Our system miss 1.6% sensitivity in the preporcessing step, 6% sensitivity in the first classification step and 1.4% sensitivity in the second classification step. Concluding the system miss 9.1% of total number of CMBs on the test set which gives the system total sensitivity at level of 90.9%. The cumulative FROC curve can be seen in Figure 23. As it can be seen from the FROC curve, at the sensitivity of 90.9% , on average, 4 false positves per subject were detected. The summary of the results for different classification steps can be seen in Table 4.

   In Figure 24, 25 and 26 some of detected true-positives, false-positives and

Table 4: Classification results on the test set. In each step we miss some sensitivity. By reducing this missing sensitivity from the missing sensitivity in previous steps, we can reach the cummulative sensitivity.

	Initial Candidate Selection	After First Classification Step Using Parzen	After Second Classification Step Using QDC
Sensitivity	98.4%	94%	98.6%
Average FP	705	19	4.1
Cumulative Sensitivity	98.4%	92.4%	90.9%

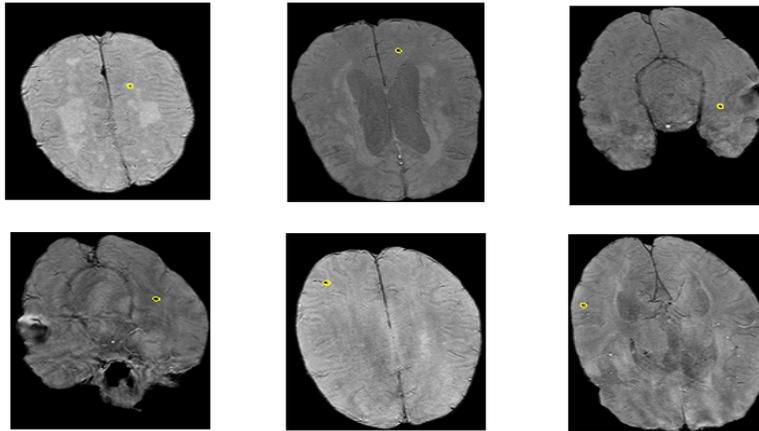Figure 24: Some results of true-positives in T2* MRI images. The yellow circles indicate the location of the CMBs.

false-negatives are shown. Most of the false-positives were caused by blood vessels but there are also other mimics as calcium and iron decompositions, partial volume artifact from bone, cavernous malformations and metastatic melanoma [12] which can be among the false-positives.
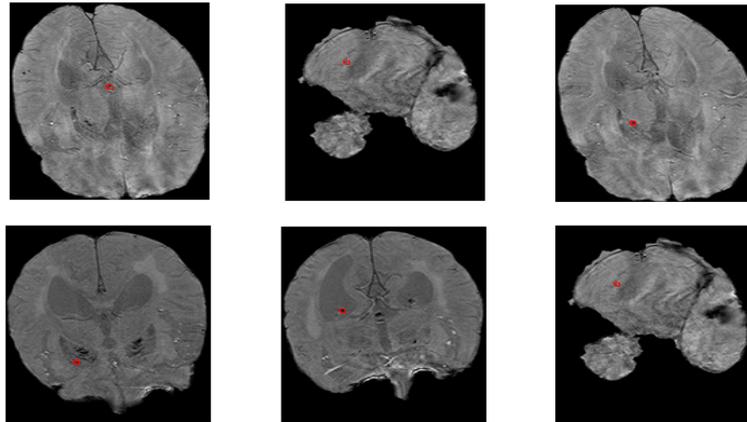
49

Figure 25: Some results of false-positives in T2* MRI images. The red circles indicate the location of the FPs.
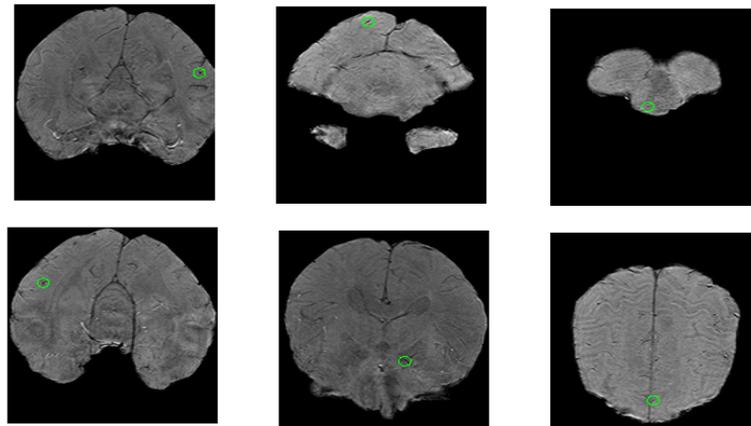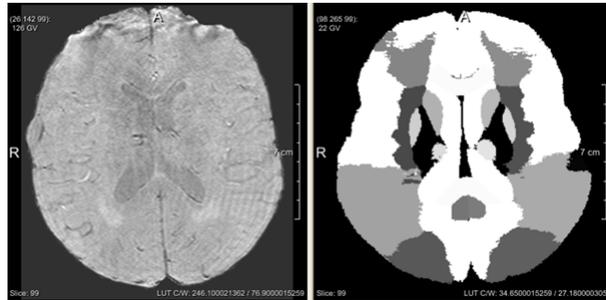

Figure 26: Some results of false-negatives in T2* MRI images. The green circles indicate the location of the FNs.

(a)

Figure 27: Left: Slice of the brain, Right: Hammers atlas shows different anatomical regions of the brain

7.5   Postprocessing verification

As discussed in section 5.3, in the final stage a postprocessing step was applied to determine the anatomical location of the microbleeds candidates. An example of Hammers atlas for a slice in our data can be seen in Figure 27.

The anatomical location of the true positive and false positive candidates, according to Hammers atlas, are shown for 76 subjects of the test set in Figure 28 and 29, respectively. The horizontal axis of histograms plot indicates different anatomical locations in the atlas.

As there are a low number of the samples in our study and a high number of the regions in Hammers atlas, 83, having a statistical validation about the distribution of CMBs in the brain is not possible. This can be reached by having a higher number of samples in future studies.
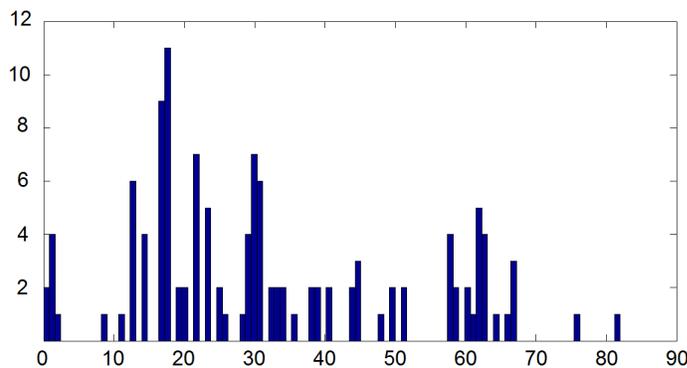

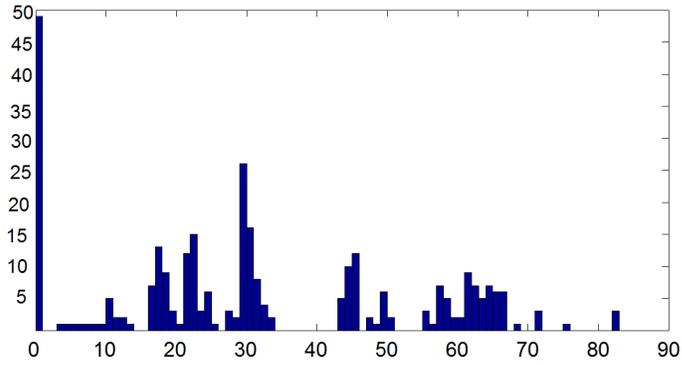Figure 28: The anatomical regions of true positives based on Hammers atlas.

Figure 29: The anatomical regions of false positives based on Hammers atlas

8   Conclusion

Detection of all cerebral microbleeds using a multi-step classification technique with a low number of false positives is desired by medical staff. We have presented a system for semi-automatic detection of the cerebral microbleeds in MR images, based on local structures and prior-knowledge of the CMBs location. Our proposed system is able to determine the anatomical location of the CMBs in the brain regions [11]. We trained and tested the system on a large database. The sensitivity and specificity of this method has been verified by comparing with manual method since there is not any fully or semi-automated system to compare with.

In the preprocessing step, by filtering the voxels which were not in the intensity range of CMBs, the system could keep the regions of interest. The system could remove connected components which were not in the CMBs' size interval. In the preprocessing step, on average, 705 false positives per subject were detected while missing 1.6% of the total number of true positives (CMBs) in the test set.

Most of candidates, which were left after preprocessing step, there were vessels and noisy structures. The intensity- and geometrical-based features are applied to remove a large fraction of false candidates, left from the preprocessing step. The most important features were selected by FFFS feature selection method. In the first classification step, by applying Parzen classifier, among 705 false positives which were left from preprocessing step, on average, 19 false positives per subject were passed for further processing in second classification step while missing 6% of the system sensitivity.

In the second classification step the QDC classifier, as the best performing classifier, was able to distinguish the CMBs from vessels and noisy structures. The most significant features for the QDC were selected by FFFS features selection method. The QDC classifier in the second classification step detected, on average, 4.1 false candidates per subject while missing 1.4% of the system sensitivity.

The radiologists at the Erasmus Rotterdam, have expected a system performance higher than 90%, with on average, less than 5 false candidate per subject. The proposed method fulfills the requirements with the average number of false positives and the sensitivity, 4.1 and 90%, respectively.

Although our proposed system contains some FPs, the semi-automatic detection of CMBs helps the radiologists to facilitate the manual segmentation and reduce needed time. The radiologist can use the system results just to examine detected candidates, and distinguish the CMBs from false postives, instead of searching all the MR image sequences of the subject.

In the future work, the system can be redesign by integrating the classification steps in a single one.

In the preprocessing step, for initial candidate selection, the size and intensity threshold values were selected by observing the train set and applying these threshold values over the test set. However, it is preferred to use a technique which can find the threshold values adaptively. This means there is a need to find a technique which is able to set the thresholding parameters with different imaging protocols.

In the classification steps different features such as the local image descriptors for PD sequences can be used. The performance of the system with the other classification methods such as artificial neural networks and kNN should be investigated.

References

[1] A.I. Bandos, H.E. Rockette, T. Song, and D. Gur. Area under the free-response roc curve (FROC) and a related summary index. *Biometrics*, 65(1):247–256, 2009.

[2] C.M. Bishop. *Neural Networks for Pattern Recognition*. Oxford University Press, 1995.

[3] M. Coffin and S. Sukhatme. Receiver operating characteristic studies and measurement errors. *Biometrics*, pages 823–837, 1997.

[4] L. Derex, N. Nighoghossian, M. Hermier, P. Adeleine, F. Philippeau, J. Honnorat, H. Yilmaz, P. Dardel, J.C. Froment, and P. Trouillas. Thrombolysis for ischemic stroke in patients with old microbleeds on pretreatment MRI. *Cerebrovasc Dis*, 17(2-3):238–241, 2004.

[5] R.O. Duda, P.E. Hart, and D.G. Stork. *Pattern classification*, volume 2. Wiley New York, 2001.

[6] R.P.W Duin. PRTools Version 3.0, A Matlab Toolbox for Pattern Recognition. In *Proceedings of SPIE in Medical Imaging*, page 1331. Citeseer, 2000.

[7] T. Fawcett. An introduction to ROC analysis. *Pattern recognition letters*, 27(8):861–874, 2006.

[8] F. Fazekas, R. Kleinert, G. Roob, G. Kleinert, P. Kapeller, R. Schmidt, and H.P. Hartung. 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. *American Journal of Neuroradiology*, 20(4):637, 1999.

[9] L.M.J. Florack. *The syntactical structure of scalar images*. Universiteit Utrecht, Faculteit Geneeskunde, 1993.

[10] A.F. Frangi, W.J. Niessen, K.L. Vincken, and M.A. Viergever. Multiscale vessel enhancement filtering. *Lecture Notes in Computer Science*, pages 130–137, 1998.

[11] B. Ghafary Asl, F. van der Lijn, M. Poels, H. Vrooman, M.A. Ikram, W.J. Niessen, A. van der Lugt, M. Vernooij, and M. de Bruijne. A computer aided detection system for cerebral microbleeds in brain MRI. In *IEEE International Symposium on Biomedical Imaging*, 2012 [Accepted].

[12] S. M. Greenberg, M.W. Vernooij, C. Cordonnier, A. Viswanathan, R. Al-Shahi Salman, S. Warach, L. J. Launer, M.A. Van Buchem, and M.Mb. Breteler. Cerebral microbleeds: a guide to detection and interpretation. *Lancet Neurol*, 8(2):165–74, 2009.

[13] A. Hammers, R. Allom, M.J. Koepp, S.L. Free, R. Myers, L. Lemieux, T.N. Mitchell, D.J. Brooks, and J.S. Duncan. Three-dimensional maximum probability atlas of the human brain, with particular reference to the temporal lobe. *Human Brain Mapping*, 19(4):224–247, 2003.

[14] R.A. Heckemann, J.V. Hajnal, P. Aljabar, D. Rueckert, and A. Hammers. Automatic anatomical brain MRI segmentation combining label propagation and decision fusion. *NeuroImage*, 33(1):115–126, 2006.

[15] A.K. Jain, R.P.W. Duin, and J. Mao. Statistical pattern recognition: A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(1):4–37, 2000.

[16] W. Kakuda, VN Thijs, MG Lansberg, R. Bammer, L. Wechsler, S. Kemp, ME Moseley, MP Marks, and GW Albers. Clinical importance of microbleeds in patients receiving IV thrombolysis. *Neurology*, 65(8):1175, 2005.

[17] S. Klein, M. Staring, K. Murphy, M.A. Viergever, and J.P.W. Pluim. ELASTIX: a toolbox for intensity-based medical image registration. *IEEE Transactions on Medical Imaging*, 29(1):196–205, 2010.

[18] Y. Lee, T. Hara, H. Fujita, S. Itoh, and T. Ishigaki. Automated detection of pulmonary nodules in helical ct images based on an improved template-matching technique. *IEEE Transactions on Medical Imaging*, 20(7):595–604, 2001.

[19] T. Lindenberg. Feature detection with automatic scale selection. *International Journal of Computer Vision*, 30(2):77–116, 1998.

[20] P. Lo and M. de Bruijne. Voxel classification based airway tree segmentation. *Progress in Biomedical Optics and Imaging-Proceedings of SPIE*, 6914, 2008.

[21] K. Murphy, B. van Ginneken, A.M.R. Schilham, B.J. de Hoop, H. Gietema, and M. Prokop. A Large Scale Evaluation of Automatic Pulmonary Nodule Detection in Chest CT using Local Image Features and k-Nearest-Neighbour Classification. *Medical Image Analysis*, 13:757–770, 2009.

[22] N. Nighoghossian, M. Hermier, P. Adeleine, K. Blanc-Lasserre, L. Derex, J. Honnorat, F. Philippeau, JF Dugor, JC Froment, and P. Trouillas. Old microbleeds are a potential risk factor for cerebral bleeding after ischemic stroke: a gradient-echo t2*-weighted brain mri study. *Stroke*, 33(3):735, 2002.

[23] G. Roob, A. Lechner, R. Schmidt, E. Flooh, H.P. Hartung, and F. Fazekas. Frequency and location of microbleeds in patients with primary intracerebral hemorrhage. *Stroke*, 31(11):2665, 2000.

[24] S.M. Smith, M. Jenkinson, M.W. Woolrich, C.F. Beckmann, T.E.J. Behrens, H. Johansen-Berg, P.R. Bannister, M. De Luca, I. Drobnjak, D.E. Flitney, et al. Advances in functional and structural mr image analysis and implementation as fsl. *Neuroimage*, 23:S208–S219, 2004.

[25] S.M. Smith, M. Jenkinson, M.W. Woolrich, C.F. Beckmann, T.E.J. Behrens, H. Johansen-Berg, P.R. Bannister, M. De Luca, I. Drobnjak, D.E. Flitney, et al. Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage*, 23:208–219, 2004.

[26] M.W. Vernooij, A. van der Lugt, and M. Breteler. Risk of Thrombolysis-Related Hemorrhage Associated With Microbleed Presence. *Stroke*, 39(7):e115, 2008.

[27] R. Wiemker, T. Bulow, and T. Blaffert. Unsupervised extraction of the pulmonary interlobar fissures from high resolution thoracic CT data. *International Congress Series of Computer Assisted Radiology and Surgery*, 1281:1121–1126, 2005.

[28] H. Wold. Partial least squares. *International Journal of Cardiology*, 147(2):581–591, 1985.

[29] M.W. Woolrich, S. Jbabdi, B. Patenaude, M. Chappell, S. Makni, T. Behrens, C. Beckmann, M. Jenkinson, and S.M. Smith. Bayesian analysis of neuroimaging data in FSL. *NeuroImage*, 45(1S1):173–186, 2009.

[30] J.M. Zabramski, T.M. Wascher, R.F. Spetzler, B. Johnson, J. Golfinos, B.P. Drayer, B. Brown, D. Rigamonti, and G. Brown. The natural history of familial cavernous malformations: results of an ongoing study. *Journal of Neurosurgery*, 80(3):422–432, 1994.

[31] X. Zhang, G. McLennan, E.A. Hoffman, and M. Sonka. Automated detection of small-size pulmonary nodules based on helical CT images. In *Proceedings of Information Processing in Medical Imaging: 19th International Conference, IPMI*. Springer, 2005.

[32] B. Zhao. Automatic detection of small lung nodules on ct utilizing a local density maximum algorithm. *Journal of Applied Clinical Medical Physics*, 4(3):248–260, 2003.