Small White Matter Lesion Detection in Cerebral Small Vessel Disease

Mohsen Ghafoorian[a,b], Nico Karssemijer[b], Inge van Uden[c], Frank Erik de Leeuw[c], Tom Heskes[a], Elena Marchiori[a] and Bram Platel[b]

[a] Institute for Computing and Information Sciences, Radboud University, Nijmegen, The Netherlands;

[b] Diagnostic Image Analysis Group, Department of Radiology, Radboud University Medical Center, Nijmegen, The Netherlands

[c] Donders Institute for Brain, Cognition and Behaviour, Department of Neurology, Radboud University Medical Center, Nijmegen, The Netherlands;

ABSTRACT

Cerebral small vessel disease (SVD) is a common finding on magnetic resonance images of elderly people. White matter lesions (WML) are important markers for not only the small vessel disease, but also neuro-degenerative diseases including multiple sclerosis, Alzheimer's disease and vascular dementia. Volumetric measurements such as the "total lesion load", have been studied and related to these diseases. With respect to SVD we conjecture that small lesions are important, as they have been observed to grow over time and they form the majority of lesions in number. To study these small lesions they need to be annotated, which is a complex and time-consuming task. Existing (semi)automatic methods have been aimed at volumetric measurements and large lesions, and are not suitable for the detection of small lesions.

In this research we established a supervised voxel classification CAD system, optimized and trained to exclusively detect small WMLs. To achieve this, several preprocessing steps were taken, which included a robust standardization of subject intensities to reduce inter-subject intensity variability as much as possible. A number of features that were found to be well identifying small lesions were calculated including multimodal intensities, tissue probabilities, several features for accurate location description, a number of second order derivative features as well as multi-scale annular filter for blobness detection. Only small lesions were used to learn the target concept via Adaboost using random forests as its basic classifiers. Finally the results were evaluated using Free-response receiver operating characteristic.

Keywords: small white matter lesions, white matter hyperintensities, computer aided detection, small vessel disease

1. INTRODUCTION

Cerebral small vessel disease (SVD) encompasses all the pathological processes that affect the small vessels of the brain. SVD is common in elderly people; some studies have reported its prevalence to reach up to 95 percent. On magnetic resonance imaging (MRI), SVD is characterized by lacunes, white matter lesions, brain microbleeds and brain subcortical atrophy. In a small proportion of cases SVD will eventually lead to cognitive, motor and mood impairment, dementia and Parkinsonism[1,2,3]. Studies are being conducted to investigate the common characteristics of those cases[4,5]. Due to lack of better measures, these studies have been focusing on the overall volume and segmentations of WMLs. It is assumed, however, that accurate detection of lesions, large or small, is a valuable asset for the study of SVD and its personalized prognosis.

There are several studies in the literature focusing on (semi)automatic WML quantification, most of which are performing white matter segmentation for different disorders including multiple sclerosis[6,7], vascular dementia[8], leukoaraiosis[9] and neuropsychiatric systemic lupus erythematosus[10]. Some methods follow a non-classification approach. As two instances of this category, Ong et al. use a histogram based method that adaptively detects intensity outliers[11], and Wu et al. utilize a fuzzy connected algorithm to iteratively grow the seed points obtained from FLAIR image intensity histogram[12]. Other methods employ voxel classification: Anbeek et al. use KNN classifier on spatial and intensity features[13]. Kloppel et al. combine intensity, Gabor and spatial information using SVM and KNN classifiers[14]. Steenwijk et al. also make use of KNN with intensities, MNI normalized spatial information and tissue type priors[15]. Finally, Ithapu et al. train SVM and random forest classifiers on intensity and texture features in order to segment WMLs for the patients

diagnosed with Alzheimer's diseases[16].

Since these researches focus on the segmentation of WMLs, and are tuned and validated with volume or Dice measurements, smaller sized lesions have been ignored in the development of these algorithms because these smaller lesions do not contribute to the total volume significantly. As a consequence, the small lesions go mostly undetected by these (semi)automatic methods. From manual annotations of over 500 SVD patients, we have learned that over 60 percent of the lesions' effective diameter is smaller or equal to 3 mm. Therefore many WML descriptors like the number of WMLs, locational distribution, progress stage and progress speed of the WMLs will be better assessed, if a better detection of smaller lesions is available. The only method that already worked on WML detection is by Riad et al. that train a general WML voxel classifier which uses a special WML sampling to equally affect the classifier by smaller and larger lesions[17]. Considering the fact that smaller and larger lesions are completely heterogeneous in shape and intensity, one cannot expect a single classifier to perfectly detect both small and large lesion concept.

In this paper, we present a CAD system that has been trained and optimized for detection of small WMLs. This optimization includes selection of appropriate features and the classifier. To assess the detection performance of our automated system, we use a free-response receiver operating characteristic (FROC) analysis.

2. MATERIALS AND METHODS

2.1 Data

The data used in this research comes from RUN DMC[4], which includes three MR images of 503 patients diagnosed with SVD who showed mild cognitive impairment evidences. A single 1.5 Tesla scanner (Magnetom Sonata, Siemens Medical Solutions, Erlangen, Germany) was used to obtain the MRI scans. The protocol included a 3D T1 magentization-prepared rapid gradient-echo sequence (TR/TE/TI 2250/3.68/850 ms; flip angle 15◦; voxel size 1.0×1.0×1.0 mm); FLAIR pulse sequences (TR/TE/TI 9000/84/2200 ms; voxel size 1.0×1.2×5.0 mm, interslice gap 1 mm) and transversal T2* weighted gradient echo sequence (TR/TE 800/26 ms; voxel size 1.3×1.0×6.0 mm, interslice gap 1 mm). Later two experienced neurologists manually annotated WMLs in a slice by slice manner, as white matter signal hyperintensities in both supra and infratentorial regions on FLAIR scans except for gliosis surrounding infarcts. 50 out of these 503 cases were manually annotated by the both experts.

2.2 Preprocessing

There are four main steps that were taken to prepare the raw T1, T2* and FLAIR images ready for the feature calculation. To setup a voxel classification framework, first of all, the three modalities were aligned so that all the corresponding voxels of the three images were at the same position. For this, T1 and T2* images were linearly registered to the FLAIR image using mutual information measure with trilinear interpolation resampling as implemented in FSL-FLIRT[18]. All subject images were also registered to the ICBM152[19] atlas so that we obtained a mapping from each subject space to atlas space for a later use.

Then we performed brain extraction to remove the skull, eyes and all other non-brain tissue. We utilized FSL-BET[20] operating on T1 since it had the highest resolution. Furthermore due to presence of the magnetic field inhomogeneity, which is prone to affect lesion classification, bias field correction was performed using FSL-FAST[21].

Even after bias field correction, there might be quite considerable intensity variability between the different subjects. As FLAIR Intensity is the most important feature for WML identification task, normalization of the image intensities across patients is imperative. One of the subjects was considered as the reference and other subjects were transformed in a way that their intensity characteristics became similar after the transformation. To achieve this, we used a bi-variate Gaussian mixture modeling (GMM) method on T1 and FLAIR intensities to extract gray matter (GM), white matter (WM) and cerebrospinal fluid (CSF). Then we projected each Gaussian on the FLAIR axis resulting in a 1D Gaussians for the FLAIR image. Given a voxel, considering its signal intensity and the tissue type it belongs to, the intensity was transformed to a new value with a similar position in the corresponding tissue type Gaussian distribution in the reference image. Based on the membership probabilities of a given intensity that the GMM process provided, a weighted average of transformed intensities were calculated. This way the Gaussians representing different tissues of all subjects' FLAIR images, were aligned with those of the reference image. The same procedure was applied to standardize the T1 images.

2.3 Features

As mentioned before, a voxel classification framework is established and twenty four different voxel features are calculated for the classifier to learn the WML concept. The group of intensity features are the most intuitive features one can consider

for this task. Intensities of the co-registered, bias field corrected and intensity standardized T1, FLAIR and T2* images are used. Since WMLs are related to WM and not to the other tissue types, three features representing the probabilities that a voxel belongs to each tissue, were added to the features set. These probabilities are obtained from the GMM process.

WMLs are not uniformly distributed over the whole brain space and thus voxel location information might be a very discriminating feature. For instance, gray matter voxels might look like WMLs in terms of intensity, but a neurology expert can easily distinguish between them, noting their position in the brain. Therefore, several features describing the voxels location were calculated: X, Y and Z of ICBM152 mapping of voxels, 2D Euclidean distances of voxels from left and right ventricles, 2D distance of voxels from the brain cortex as well as the 2D distance of voxels from the mid-sagittal brain surface were calculated. Furthermore, considering the relatively high number of available subjects, a pretty accurate prior probability of WMLs in each location was obtained by calculating the proportion of patients with an annotated WMLs in the corresponding atlas position. This proportion was used as another location-based feature.

A number of second order derivative features describing blob-like structures, were considered useful for small WML detection and were calculated and added to the features set: These features include Multi-scale Laplacian of Gaussian, determinant of Hessian matrix, vesselness and gauge derivative in the direction of the normal vector[22]. Furthermore because small WMLs mostly appear as isolated hyperintense blob-like structures, multi-scale gray-scale annular filter, as another blobness detector was used[23]. Table 1 summarizes the mentioned features.

Table 1. A summary of the features used to train the classifier for detection of small WMLs.

Group	**Feature**
Intensities	FLAIR intensity
	T1 intensity
	T2* intensity
Tissue Probabilities	WM probability
	GM probability
	CSF probability
Location information	X, Y and Z in MNI space
	2D Euclidean distance from left and right ventricle
	2D Euclidean distance from brain cortex
	2D Euclidean distance from mid-sagittal brain surface
	WMLs prior probability
Second order derivatives	Multi-scale Laplacian of Gaussian (t=1,2,4 mm)
	Multi-scale determinant of Hessian (t=1,2,4 mm)
	Vesselness (σ=1 mm)
	Gauge derivative in the direction of the normal vector
Annular filter	Multi-scale gray-scale annular filter (t=1,2,4)

2.4 Sampling and Training

A random subset of 100 images was selected for training. Considering the fact that there are much more negative samples compared to the positives, all voxels in WMLs having effective diameter smaller or equal to 3 mm were selected as positive samples. On the other hand, all trivial negative samples which had FLAIR intensity less than a specified threshold, were omitted, and 2% of the remaining voxels were taken randomly as the negative samples.

A random forest with 20 sub-trees was selected as the base classifier. Then three iterations of Adaboost were run over the random forest. In the Adaboost learning algorithm samples that were misclassified in previous iterations get higher chance to be selected for the next iterations. This causes the whole ensemble to be better capable to handle harder samples. The resulting classifier that concentrates more on harder samples, is a better fit for small WML classification that is a tough

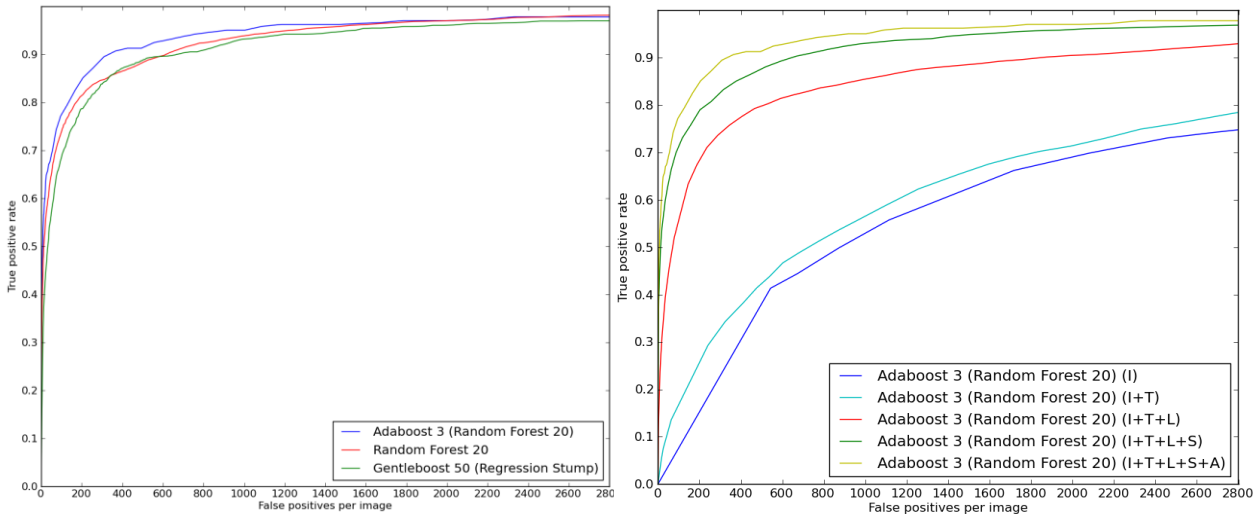
task due to the partial volume effect, presence of dirty white matter as well as small noises that resemble small WMLs.

2.5 Validation

To assess the quality of the proposed CAD system, we make use of FROC analysis in order to emphasize more in detection of smaller lesions. The detailed procedure is as follows: Firstly local maxima of the lesions likelihood map resulted from the voxel classification are calculated in 3×3 neighborhoods. Then at different decision boundary thresholds on local maxima likelihoods, each local maxima that is found outside of the manual WML annotations is considered as a false positive, and every annotation segment hit by at least one local maxima, a true positive.

3. RESULTS

The described validation process was performed on 32 test subjects. Figure 1 (a) compares the FROC curves of a single random forest, Gentleboost using regression stumps and Adaboost using random forest. As the figure suggests, Adaboost using random forest as its base classifier, dominates the other two. Figure 1 (b) shows the performance of Adaboost using random forest with incrementally added sets of features.



(a) (b)

Figure 1. FROC evaluations of the CAD system: (a) FROC curve for three different classifiers: 3 iterations of Adaboost using random forest with 20 sub-trees, a single random forest with 20 sub-trees and Gentleboost with 50 regression stump classifiers (b) FROC curves for models created with different feature sets: (I=Intensities, T=Tissue Probabilities, L=Locations, S=Second order derivatives, A=multi-scale Annular filter)

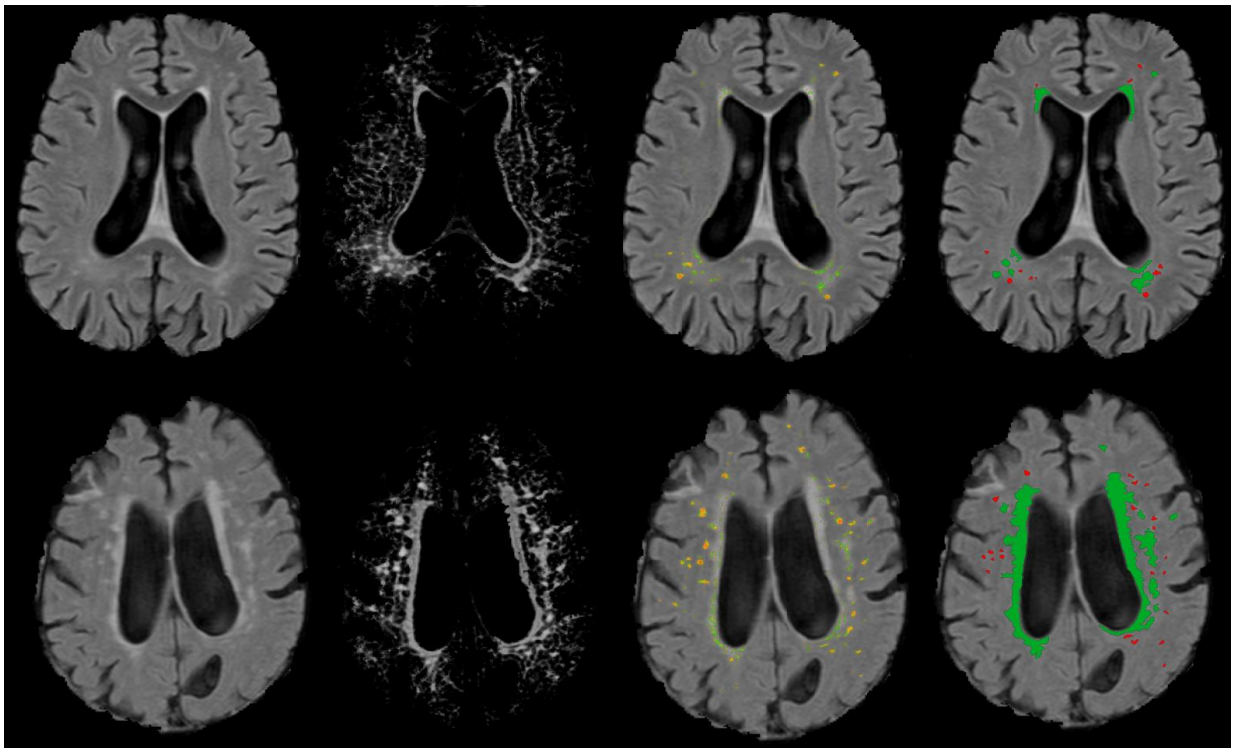
For a visual assessment of the quality of the presented CAD system, figure 2 demonstrates the detection results and compares it to the human expert annotations for small WMLs.

4. CONCLUSIONS AND DISCUSSIONS

In this paper a method for the automatic detection of small WMLs was presented. Current research has focused solely on volumetric measurements or lesion segmentations, and have largely ignored these small lesions. Our method makes uses of features relevant to small WMLs, some of which are novel to be used in this domain. Contribution of each feature groups has been evaluated. Based on the results, it is notable that there is significant amount of information in standardized image intensities and our accurate location information feature group. Also interestingly, the not very well-known grayscale annular filter significantly contributes in detection performance, even though there were already Laplacian of Gaussian and determinant of Hessian as two other multi-scale blobness detectors in the features set. Also the Adaboost classifier has shown increased performance over a single random forest classifier, due to its ability to concentrate on harder samples.

Finally we have to note once more that small WML detection is a tough task with a large inter- and intra-reader variability, mostly due to the presence of dirty WM, which is the WM in transition to WML, image noise due to imaging deficiencies, patient movement and thick image slices that cause partial volume effect. The classifier reached 0.8 true

positive rate with less than 140 false positives per scan on average. Further investigation on false positives by aligning them to follow up scans from 5 years later, which enjoys thinner slices and higher contrast, showed that a significant proportion of false positives were either an obvious small WML present in the original scans that were missed by the expert annotators or high potential dirty white matters that were already transited to an obvious lesion in the follow-up scans.



(a) (b) (c) (d)

Figure 2. A visual demonstration of the performance of the proposed CAD system. (a) sample FLAIR slices (b) small WML likelihood maps generated by the CAD system (c) overlayed heatmaps for the thresholded WML likelihoods ($\tau = 0.7$) (d) expert annotations for the small lesions in red and larger lesions in green.

REFERENCES

[1] Vermeer, Sarah E., Niels D. Prins, Tom den Heijer, Albert Hofman, Peter J. Koudstaal, and Monique MB Breteler. "Silent brain infarcts and the risk of dementia and cognitive decline." *New England Journal of Medicine* 348, no. 13 (2003): 1215-1222.

[2] van Uden, Ingeborg WM, Anil M. Tuladhar, Karlijn F. de Laat, Anouk GW van Norden, David G. Norris, Ewoud J. van Dijk, Indira Tendolkar, and Frank-Erik de Leeuw. "White Matter Integrity and Depressive Symptoms in Cerebral Small Vessel Disease: The RUN DMC Study." *The American Journal of Geriatric Psychiatry* (2014).

[3] de Laat, Karlijn F., Anouk GW van Norden, Rob AR Gons, Inge WM van Uden, Marcel P. Zwiers, Bastiaan R. Bloem, Ewoud J. van Dijk, and Frank-Erik de Leeuw. "Cerebral white matter lesions and lacunar infarcts contribute to the presence of mild parkinsonian signs." *Stroke* 43, no. 10 (2012): 2574-2579.

[4]

[5] van Norden, Anouk GW, Karlijn F. de Laat, Rob AR Gons, Inge WM van Uden, Ewoud J. van Dijk, Lucas JB van Oudheusden, Rianne AJ Esselink et al. "Causes and consequences of cerebral small vessel disease. The RUN DMC study: a prospective cohort study. Study rationale and protocol." *BMC neurology* 11, no. 1 (2011): 29.

[6] Pantoni, Leonardo, Anna Maria Basile, Giovanni Pracucci, Kjell Asplund, Julien Bogousslavsky, Hugues Chabriat, Timo Erkinjuntti et al. "Impact of age-related cerebral white matter changes on the transition to disability–the LADIS study: rationale, design and methodology." *Neuroepidemiology* 24, no. 1-2 (2004): 51-62.

[7] Lao, Zhiqiang, Dinggang Shen, Dengfeng Liu, Abbas F. Jawad, Elias R. Melhem, Lenore J. Launer, R. Nick Bryan, and Christos Davatzikos. "Computer-assisted segmentation of white matter lesions in 3D MR images using support vector machine." *Academic radiology* 15, no. 3 (2008): 300-313.
[8] Khayati, Rasoul, Mansur Vafadust, Farzad Towhidkhah, and Massood Nabavi. "Fully automatic segmentation of multiple sclerosis lesions in brain MR FLAIR images using adaptive mixtures method and Markov random field model."*Computers in biology and medicine* 38, no. 3 (2008): 379-390.
[9] Yamashita, Yasuo, Hidetaka Arimura, and Kazuhiro Tsuchiya. "Computer-aided detection of ischemic lesions related to subcortical vascular dementia on magnetic resonance images." *Academic radiology* 15, no. 8 (2008): 978-985.
[10] Schwarz, Christopher, Evan Fletcher, Charles DeCarli, and Owen Carmichael. "Fully-automated white matter hyperintensity detection with anatomical prior knowledge and without FLAIR." In *Information processing in medical imaging*, pp. 239-251. Springer Berlin Heidelberg, 2009.
[11] Scully, Mark, Blake Anderson, Terran Lane, Charles Gasparovic, Vince Magnotta, Wilmer Sibbitt, Carlos Roldan, Ron Kikinis, and Henry J. Bockholt. "An automated method for segmenting white matter lesions through multi-level morphometric feature classification with application to lupus." *Frontiers in human neuroscience* 4 (2010).
[12] Ong, Kok Haur, Dhanesh Ramachandram, Rajeswari Mandava, and Ibrahim Lutfi Shuaib. "Automatic white matter lesion segmentation using an adaptive outlier detection method." *Magnetic resonance imaging* 30, no. 6 (2012): 807-823.
[13] Wu, Minjie, Caterina Rosano, Meryl Butters, Ellen Whyte, Megan Nable, Ryan Crooks, Carolyn C. Meltzer, Charles F. Reynolds III, and Howard J. Aizenstein. "A fully automated method for quantifying and localizing white matter hyperintensities on MR images." *Psychiatry Research: Neuroimaging* 148, no. 2 (2006): 133-142.
[14] Anbeek, Petronella, Koen L. Vincken, Floris Groenendaal, Annemieke Koeman, Matthias JP Van Osch, and Jeroen Van der Grond. "Probabilistic brain tissue segmentation in neonatal magnetic resonance imaging." *Pediatric research* 63, no. 2 (2008): 158-163.
[15] Klöppel, Stefan, Ahmed Abdulkadir, Stathis Hadjidemetriou, Sabine Issleib, Lars Frings, Thao Nguyen Thanh, Irina Mader, Stefan J. Teipel, Michael Hüll, and Olaf Ronneberger. "A comparison of different automated methods for the detection of white matter lesions in MRI data." *NeuroImage* 57, no. 2 (2011): 416-422.
[16] Steenwijk, Martijn D., Petra JW Pouwels, Marita Daams, Jan Willem van Dalen, Matthan WA Caan, Edo Richard, Frederik Barkhof, and Hugo Vrenken. "Accurate white matter lesion segmentation by k nearest neighbor classification with tissue type priors (kNN-TTPs)." *NeuroImage: Clinical* 3 (2013): 462-469.
[17] Ithapu, Vamsi, Vikas Singh, Christopher Lindner, Benjamin P. Austin, Chris Hinrichs, Cynthia M. Carlsson, Barbara B. Bendlin, and Sterling C. Johnson. "Extracting and summarizing white matter hyperintensities using supervised segmentation methods in Alzheimer's disease risk and aging studies." *Human brain mapping* (2014).
[18] Riad, Medhat M., Bram Platel, Frank-Erik de Leeuw, and Nico Karssemeijer. "Detection of white matter lesions in cerebral small vessel disease." In *SPIE Medical Imaging*, pp. 867014-867014. International Society for Optics and Photonics, 2013.
[19] Jenkinson, Mark, and Stephen Smith. "A global optimisation method for robust affine registration of brain images." *Medical image analysis* 5, no. 2 (2001): 143-156.
[20] Mazziotta, John, Arthur Toga, Alan Evans, Peter Fox, Jack Lancaster, Karl Zilles, Roger Woods et al. "A four-dimensional probabilistic atlas of the human brain." *Journal of the American Medical Informatics Association* 8, no. 5 (2001): 401-430.
[21] Smith, Stephen M. "Fast robust automated brain extraction." *Human brain mapping* 17, no. 3 (2002): 143-155.
[22] Zhang, Yongyue, Michael Brady, and Stephen Smith. "Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm." *Medical Imaging, IEEE Transactions on* 20, no. 1 (2001): 45-57.
[23] Kuijper, Arjan. "Geometrical PDEs based on second-order derivatives of gauge coordinates in image processing." *Image and Vision Computing* 27, no. 8 (2009): 1023-1034.
[24] Moshavegh, Ramin, B. E. Bejnordi, Andrew Mehnert, K. Sujathan, Patrik Malm, and Ewert Bengtsson. "Automated segmentation of free-lying cell nuclei in Pap smears for malignancy-associated change analysis." In *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, pp. 5372-5375. IEEE, 2012.