Automated White Matter Lesion Segmentation in MRI using Box-Whisker Plot Outlier Detection

Ong Kok Haur[1]
khong.com08@student.usm.my
Dhanesh Ramachandram[1]
dhaneshr@cs.usm.my
Mandava Rajeswari[1]
mandava@cs.usm.my
Ibrahim Lutfi Shuaib[2]
ibrahim@amdi.usm.edu.my

[1] School of Computer Sciences
Universiti Sains Malaysia
11800, Pulau Pinang, Malaysia
Email: dhaneshr@cs.usm.my
Website : http://cvrg.cs.usm.my/

[2] Advanced Medical and Dental Institute
Universiti Sains Malaysia,
11800 Pulau Pinang, Malaysia

Abstract

We report a new automated method for White Matter Lesions (WMLs) Segmentation in cranial MR Imaging. WMLs are diffuse white matter abnormalities which are often presented as hyperintense regions. In our approach, the presence of these abnormalities are detected as outliers in the intensity distribution of the FLAIR sequence using the histogram tails analysis and Box-Whisker plot technique. In addition, our method includes post-processing to reduce False Positives attributed to MRI artefacts commonly observed in FLAIR sequences. We validated our approach using 19 cases of cranial MRI. A high correlation is seen between our automated approach and the results of a manual visual scoring approach performed by an expert radiologist.

1 Introduction

White Matter Lesions (WMLs), also known as White Matter Hyperintensities, have been shown to be predictors of several neurological disorders such as Multiple Sclerosis, Vascular dementia, Stroke and Alzheimer's Disease. In recent years, there have been a number of computer-aided WML segmentation approaches reported in the literature. In this paper, our focus is on threshold-based techniques. Threshold based techniques aim to find an optimal threshold value from the intensity histogram as a cut-off point to segment WMLs. In an early study, Hirono *et al.* [3] defined a threshold value of 3.5 SD of the White Matter(WM) voxel intensity distribution to segment WMLs. Jack *et al.* [4] implemented a more complex regression model to define a cut-off threshold for the FLAIR sequence. In yet another study [8], a white matter probability map (MNI 152 brains[1]), was used as a weighting function to favour the areas that are most likely populated with white matter. Various brain tissue types are modelled statistically. Voxels with intensities beyond $\mu_{WM} + 6\sigma_{WM}$ are classified as severe

[1] MNI Brains are various brain atlases modeled by the Montreal Neurological Institute and are popularly used as standard brain in neuro-radiological studies

WMLs while voxels having intensities in the range of $\mu_{WM} + 3\sigma_{WM}$ are classified as mild & moderate WMLs. Boer et al. [2], performed WML segmentation using an intensity threshold defined by: T = $\mu + \alpha\sigma$, where α is an optimized threshold parameter. To find the optimum value for α, segmentation results using various values of α were compared against two expert delineations. We present a fully automated WMLs detection and segmentation method, which uses Box-Whisker plot introduced by Tukey [6] to detect WMLs pixels, which are regarded in our case, as outliers. Apart from being a statistically sound approach for outlier detection, Box-Whisker plots also emphasize the histogram tails, which is of particular interest to us since WMLs too are distributed primarily in the right-tail end of the voxel intensity distribution in FLAIR images. A key advantage of our proposed technique is that it does not require prior training or modeling. Moreover, since our algorithm is computationally inexpensive, the segmentation can be performed in real-time. In this paper, we report the correlation between our automated WML segmentation approach and the visual score as determined by an expert radiologist, as a means to validate the proposed approach.

2 Materials and Methods

The dataset used in this study are MRI sequences obtained from the Advanced Medicine and Dentistry Institute(AMDI), Universiti Sains Malaysia. Cranial MR images of 19 subjects comprising of T1-weighted (T1-W) and Fast Fluid Attenuated Inversion Recovery (FLAIR) sequences. The subjects were scanned using 1.5T magnetic field strength with acquisition matrices of 512 x 512 for axial FLAIR (mean TR 8002±0 ms, mean TE 127.13±4.26 ms) and axial T1-weighted (mean TR 489.47±29.34 ms, mean TE 14±0 ms) sequences. Both sequences have a slice thickness of 5.0 mm. Subjects were between 39 and 75 years of age (mean age 58.31±9.53 yrs), whose WMLs visual scores[7] ranged between 2 and 18 (mean WMLs score 5.84±3.88). The volume of WMLs segmented by the proposed automated method is then used to compare with the gold standard assessments based on manual expert visual scoring.

Our WML segmentation approach uses multispectral information from T1-weighted, and FLAIR sequences. We adopted the model-based level set approach proposed by Zhuang *et al.* [9] to perform skull stripping. The skull stripped T1-Weighted sequence is subsequently used as a mask to remove the skull in the corresponding FLAIR sequence. Since 95% of WMLs occur within the WM, the WM region must be reliably identified first so that hyperintense voxels which are not part of the WM could be later discarded. Using the T1-Weighted sequence, voxels belonging to the WM, as well as the GM, CSF and the background (BG) regions are classified using the Fuzzy-C-Means clustering. To improve the clustering results, we apply N3-inhomogeneity correction [5].

The input to the segmentation algorithm is the skull-stripped FLAIR sequence. Generally, WMLs can be regarded extreme outliers in the voxel intensity distribution. In order to detect these outliers, we use the Box-Whisker plot. In the Box-Whisker plot method, outliers, f_3, are defined as Eq. 1:

$$f_3 = Q_3 + 1.5 * IQR \tag{1}$$

where IQR is the Inter Quartile Range which denotes points falling within the 25 percentile and 75 percentile of the voxel distribution(see Eq. 2):

$$IQR = Q_3 - Q_1 \tag{2}$$

In addition, extreme outliers F_3 are defined as:

$$F_3 = Q_3 + 3 * IQR \tag{3}$$

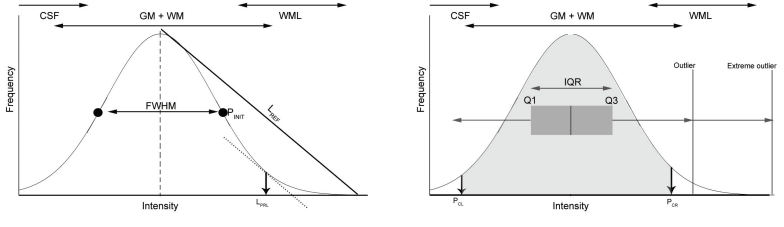
As a prerequisite to detect outliers, it is important to first determine the range of intensities which represent normal brain tissue. In our case, this would be the range of intensities that are occupied by GM and WM. Due to the partial volume effect (PVE), the intensity ranges for CSF, GM+WM and WMLs typically overlap. Therefore, we need to have a good estimate of the range of voxel distribution for the normal brain tissue (i.e. GM + WM). In our proposed approach, we heuristically estimate the range of intensities directly from the histogram of a given image. The procedure follows the following steps:

1 The histogram of the skull-stripped FLAIR image is first constructed. Histogram smoothing is then applied using a 1-Dimensional Gaussian kernel.

2 Considering only the right half of the histogram, an initial point, P_{INIT}, is set at the full width at half maximum (FWHM) point on the smoothed histogram. P_{INIT} is then used as the starting point to iteratively search for the rightmost point, P_{CR} (see Fig. 1(b)), that best bounds the GM + WM voxels. We define P_{CR} to be the tangent point between the curve of the histogram and the line, L_{PRL} , (dashed line) that is parallel to the "reference slope", L_{REF}, for the right tail of the histogram. L_{REF} is the line that connects the peak of the smoothed histogram to rightmost tip as shown in Fig. 1(a).

3 Step 2 is repeated for the left tail of the histogram, which would result in point P_{CL} being determined (see Fig. 1(b)).

4 The points P_{CL} and P_{CR} define the *potential* range of intensities for the WM + GM voxels, in other words, an estimate of the normal brain tissue distribution.



(a) (b)

Figure 1: 1(a) A smoothed histogram depicting overlapping intensity distributions for CSF, GM, WM and WMLs. 1(b) The outliers and extreme outliers are determined using the Box and Whisker plot using the intensity distribution of the GM and WM voxels.

As mentioned earlier, the points P_{CL} and P_{CR} only represent an initial estimate of the range of intensities for the GM and WM. A more statistically sound measure called the (Inter-Quartile Range) IQR, is then used to compute a more accurate estimate of the bounds of the normal tissue distribution. Technically, the IQR represents 50% of the middle data for a given normal distribution. The notion of using the IQR to represent normal tissue distribution is appealing because its value is not affected by extreme potential outliers in the data, which can often distort the computing of a measure of spread, and thereby lessening the

sensitivity to outliers. Hence, we compute the IQR for the range of intensities between the points P_{CL} and P_{CR} to obtain a more accurate estimate of the normal brain tissue distribution. Subsequently, we detect outliers and extreme outliers for the FLAIR image using Eq. 1 and Eq. 3. Results obtained after testing 2100 MRI images, indicate that the existence of *extreme outliers* is a highly probable indicator of WMLs. Voxels detected as being *outliers*, on the other hand, also indicate the presence of WMLs but with lesser probability. Even though outliers have lesser probability of signalling the presence of WML, a considerable number of voxels do actually have intensity values that fall between the outlier and the extreme outlier points. Therefore, in our approach, extreme outlier points are first used to initially detect the presence of WMLs in the skull-stripped FLAIR sequence. If an extreme outlier point is found, the WMLs are then segmented using the range of intensities that fall between the outlier point and the extreme outlier point in the given histogram. In Box-Whisker plot analysis, this range of intensities is known as the *outer fence* [6].

False Positives in FLAIR can be attributed to numerous factors including incomplete skull stripping and flow artefacts [1]. In our approach, we use the voxels classified as CSF and WM in the T1-weighted sequences together with morphological processing to reduce False Positives. Firstly, we apply a dilation operation on the CSF voxels with a 3x3 structuring element, before using the dilated region as a mask to remove flow artefacts present predominantly at the peri-ventricular region. Next, we remove voxels detected as WMLs but which do not overlap with the WM region. It should be noted that it is impossible to remove all false positives for a given image as there are not clear-cut distinction between WMLs and artefacts, both appear with similar brightness characteristics. Our morphological post processing can however minimize the effects of FLAIR-related artefacts, thereby potentially producing accurate segmentation. In fact, there are still numerous approaches [2, 3, 4, 8] which seem to ignore the effect of flow artefacts altogether and do not report any form of post-processing.

3 Results and Discussion

We performed regression analysis to measure the correlation between the proposed automated WMLs segmentation approach and the visual scoring approach. Our results indicate that there exists a significant correlation between our approach and the manual visual scoring approach ($R = 0.8506$, $P = 3.94 \times 10^{-6}$)(Table 1). It is evident that our approach is consistent with the visual score provided by experts and can be reliably used to automate the analysis and quantification of the WMLs on large scale of data. The sample results using our approach are shown in Fig. 2.

Score	2	4	3	3	2	2	3	4	4	6
WMLs load(mm^3)	3574	1442	1248	1900	765	3386	702	8683	4944	2531
Score	10	18	8	6	8	6	6	10	6	
WMLs load(mm^3)	15443	127235	16365	4484	23262	2526	4364	12712	10769	

Table 1: Visual score and calculated total lesion load for 19 subjects

This paper presents a new approach to WMLs detection and segmentation in MR Images. It includes preliminary evaluation of our approach on a 19 subjects of dataset with encouraging results. Moreover, the presented approach has been tested on FLAIR datasets

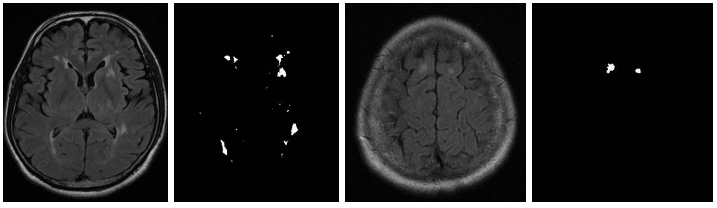


Figure 2: Automated segmentation results for 2 sample MR images with varying lesion loads.

that were acquired using different MRI scanner parameter settings. We are currently conducting a more thorough evaluation of our approach on white matter lesions using MRI data obtained from our university's hospital(HUSM) and the Advanced Medicine and Dentistry Institute(AMDI), USM.

4 Acknowledgements

This work is funded through Universiti Sains Malaysia's Research University Grant "Delineation and 3D Visualization of Tumor and Risk Structures", Grant No: 1001/PKOMP/817001.

References

[1] R. Beare, V. Srikanth, J. Chen, T. G. Phan, J. Stapleton, R. Lipshut, and D. Reutens. Development and validation of morphological segmentation of age-related cerebral white matter hyperintensities. *NeuroImage*, 47(1):199–203, 2009.

[2] R. de Boer, H. A. Vrooman, F. van der Lijn, M. W. Vernooij, M. A. Ikram, A. van der Lugt, M. M. B. Breteler, and W. J. Niessen. White matter lesion extension to automatic brain tissue segmentation on mri. *NeuroImage*, 45(4):1151–1161, 2009.

[3] N. Hirono, H. Kitagaki, H. Kazui, M. Hashimoto, and E. Mori. Impact of white matter changes on clinical manifestation of alzheimer's disease: A quantitative study. *Stroke*, 31(9):2182–2188, 2000.

[4] C. R. Jack, P. C. O'Brien, D. W. Rettman, M. M. Shiung, Y. Xu, R. Muthupillai, A. Manduca, R. Avula, and B. J. Erickson. Flair histogram segmentation for measurement of leukoaraiosis volume. *Journal of Magnetic Resonance Imaging*, 14(6):668–676, 2001.

[5] J. G. Sied, A. P. Zijdenbos, and A. C. Evans. A nonparametric method for automatic correction of intensity nonuniformity in mri data. *IEEE Transactions on Medical Imaging*, 17(1):87–97, 1998.

[6] J. W. Tukey. *Exploratory data analysis*. Reading, Mass: Addison-Wesley Publishing Company, 1977.

[7] L. O. Wahlund, F. Barkhof, F. Fazekas, L. Bronge, M. Augustin, M. Sjogren, A. Wallin, H. Ader, D. Leys, L. Pantoni, F. Pasquier, T. Erkinjuntti, and P. Scheltens. A new rating scale for age-related white matter changes applicable to mri and ct. *Stroke*, 32(6):1318–1322, 2001.

[8] W. Wen and P. Sachdev. The topography of white matter hyperintensities on brain mri in healthy 60- to 64-year-old individuals. *NeuroImage*, 22(1):144–154, 2004.

[9] A. H. Zhuang, D. J. Valentino, and A. W. Toga. Skull-stripping magnetic resonance brain images using a model-based level set. *NeuroImage*, 32(1):79–92, 2006.