Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 90 (2016) 9 – 14

Procedia Computer Science

International Conference On Medical Imaging Understanding and Analysis 2016, MIUA 2016, 6-8 July 2016, Loughborough, UK

Texture-based classification for the automatic rating of the perivascular spaces in brain MRI

Víctor González-Castro[a,*], María del C. Valdés Hernández[a], Paul A. Armitage[b], Joanna M. Wardlaw[a]

[a] *Department of Neuroimaging Sciences, Centre for Clinical Brain Sciences, University of Edinburgh, 49 Little France Crescent, Edinburgh EH16 4SB, United Kingdom*

[b] *Department of Infection, Immunity and Cardiovascular Disease, University of Sheffield, Royal Hallamshire Hospital, Sheffield S10 2JF, United Kingdom*

Abstract

Perivascular spaces (PVS) relate to poor cognition, depression in older age, Parkinson's disease, inflammation, hypertension and cerebral small vessel disease when they are enlarged and visible in magnetic resonance imaging (MRI). In this paper we explore how to classify the density of the enlarged PVS in the basal ganglia (BG) using texture description of structural brain MRI. The texture of the BG region is described by means of first order statistics and features derived from the co-occurrence matrix, both computed from the original image and the coefficients yielded by the discrete wavelet transform (WSF and WCF, respectively), and local binary patterns (LBP). Experimental results with a Support Vector Machine (SVM) classifier show that WCF achieves an accuracy of 80.03%.

Peer-review under responsibility of the Organizing Committee of MIUA 2016

Keywords: Brain MRI; Perivascular Spaces; Texture Descriptors; Discrete Wavelet Transform; Local Binary Pattern; Support Vector Machine.

1. Introduction

Perivascular spaces (PVS), also known as Virchrow-Robin spaces are interstitial fluid-containing cavities that surround the walls of the small arteries, veins and capillaries in the brain. They are normally microscopic but, when they are enlarged, they are visible on T2-weighted (T2W) brain magnetic resonance imaging (MRI) as round or linear structures, less than 3mm cross sectional diameter, with intensities close to those of the cerebrospinal fluid (CSF). An increase of enlarged PVS (EPVS) has been associated with worse cognition, depression at older ages, Parkinson's disease, inflammatory disorders, hypertension and cerebral small vessel disease (SVD) in the form of lacunar stroke and vascular dementia[1]. PVS are increasingly recognised as an important component of the brain's circulation and fluid drainage pathways. Therefore, quantifying the PVS is of significant interest.

* Corresponding author. Tel.: +44 (0)131 465 9528.
E-mail address: victor.gonzalez@ed.ac.uk

Peer-review under responsibility of the Organizing Committee of MIUA 2016
doi:10.1016/j.procs.2016.07.003

Manual volume quantification (i.e. segmentation) of PVS is very time consuming even if it is computer-aided, and the existent automatic methods present serious limitations, due to the overlap in PVS shape, intensity, location and size with those of lacunes[2]. Ramirez et al. assessed the quantitative volume densities of PVS, showing that PVS in the white matter are associated to AD-related vascular pathologies[3]. Recently Wang et al. presented a method based on thresholding T2W images acquired using a 1.5T MRI scanner to quantify PVS in the basal ganglia[4] which achieved good results, although it required manual intervention. Cai et al. developed a method to produce quantitative results of PVS volume densities. This method is automatic, produced good results and showed an increase in PVS density in patients with Alzheimer's disease (AD) compared to healthy controls, but it was used with high resolution 7T MRI scans[5] which still have limited applicability in clinical use.

As an alternative to quantitative measurements, several similar visual rating scales have been proposed in recent years to measure the EPVS burden. Potter et al. reviewed their ambiguities and advantages, and combined their strengths to develop a more comprehensive scale which proved to be robust in terms of its intra- and inter-observer reliability[1]. However, as with any visual recognition process, it is subject to observer bias and relatively insensitive to subtle changes. An automatic EPVS rating method (e.g. replicating the visual rating scale using image processing and pattern recognition) could overcome these and the drawbacks of the existent segmentation methods.

Pattern recognition and machine learning have been widely applied in medical imaging, for segmentation[6,7] or for computer-aided diagnosis[8]. Chen et al. proposed a framework based on multiple instance learning to distinguish between absent/mild vs. moderate to severe SVD in computer tomography (CT) scans which achieved approximately 75% accuracy[9]. However, to the best of our knowledge, there are no similar methods that deal with the EPVS burden (i.e. replicating the EPVS visual rating scale).

In this paper we propose an automatic method to rate EPVS in the basal ganglia (BG) from structural T2W MRI. To do this we describe the texture of the BG region by means of Local Binary Patterns (LBP)[10] and first order statistical features and co-occurrence features proposed by Haralick[11] combined with the Discrete Wavelet Transform (DWT)[12]. These methods have already been applied in medical image classification works with good results[8,13].

The paper is organised as follows: In Section 2 the dataset and the proposed method are explained. Section 3 introduces the experimental setup, results and discussion. Finally, the conclusions are presented in Section 4.

2. Materials and methods

2.1. Imaging data

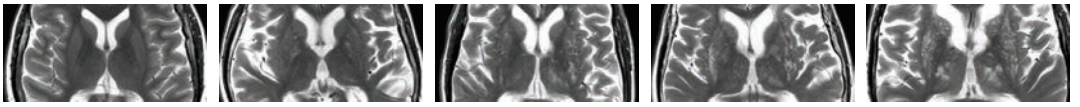
We used data from 264 patients who gave written consent to participate on a study of stroke mechanisms[14]. Brain MRI was conducted on a 1.5 tesla GE Signa LX clinical scanner (General Electric, Milwaukee, WI), equipped with a self-shielding gradient set and 8-channel-phased array heal coil following the protocol described in[14]. For our classification we used the T2W images (TE/TR = 147/9002 milliseconds, field of view 240×240 mm, acquisition matrix 256×256, slice thickness 5 mm, 1 mm inter-slice gap and voxel size $0.94 \times 0.94 \times 6.5$ mm).

2.2. EPVS visual rating scale

EPVS were assessed using the visual rating scale developed by Potter et al.[1]. It rates the EPVS separately in three major anatomical brain regions, i.e. midbrain, basal ganglia (BG) and centrum semiovale (CS) using T2W MRI, assigning 0 (no EPVS), 1 (mild; 1-10 EPVS), 2 (moderate; 11-20 EPVS), 3 (frequent; 21-40 EPVS) or 4 (severe; >40 EPVS). More detais about this visual rating scale can be found at http://www.sbirc.ed.ac.uk/documents/epvs-rating-scale-user-guide.pdf. Examples of each rating for the BG region are shown in Figure 1. In this paper, we will focus only on the EPVS in the BG, since moderate to severe EPVS in this region is considered a marker of cerebral SVD, associated with cognitive decline[15], vascular dementia and stroke[16].

2.3. Pre-processing

The intracranial volume and cerebrospinal fluid (CSF) were segmented automatically using optiBET[17] and FSL-FAST[18]. All subcortical structures were segmented automatically using other tools from the same FMRIB Software

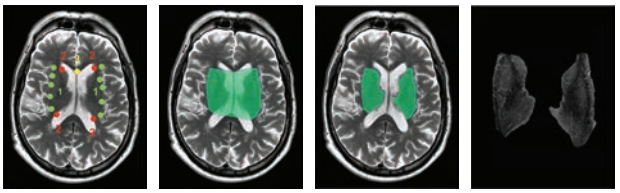


(a) (b) (c) (d) (e)

Fig. 1. Example for the EPVS ratings in the BG 0 to 4 ((a) to (e), respectively).

Library (FSL) on a pipeline described in [14]. For our classifier, we choose the axial slice with the highest number of PVS containing at least one characteristic BG structure, following the indications from Wang et al. [4]. As first approximation, from the slices that contained BG structures, we selected those in which the total area of these structures was more than 5 % of the intracranial area defined on the 2D slice.

On each of the slices initially selected, a polygon enclosing the BG, internal and external capsules and thalami was automatically drawn by joining anatomical points in the insular cortex, the closest points to them in the lateral ventricles (frontal and occipital horns) and the intercept of the genu of the corpus callosum with the septum; and subtracting from it the region occupied by the CSF. These steps are illustrated in Figure 2.



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

Fig. 2. Steps of the BG segmentation: (a) Detection of the vertices in the insular cortex (1), lateral ventricles (2) and genu (3); (b) creation of the polygon; (c) subtraction of the CSF from the BG polygonal region and (d) segmented BG region.

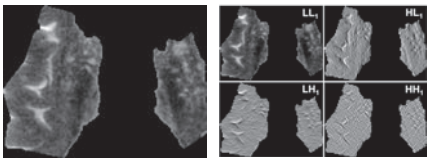
From this subset of slices, the single slice where our classifier operated was selected after applying contrast-limited adaptive histogram equalisation (CLAHE) [19] to the polygonal region and thresholding it at 0.43 times the maximum intensity level [2,4] (Fig. 2(d)). Later, we counted the number of hyperintense blobs with area between 3 and 15 times the in-plane voxel dimensions [4]. Although this procedure overestimates the number of PVS in the presence of other SVD markers (e.g. small lesions and lacunes) [2], it provides a good estimate of the axial slice with the most PVS.

2.4. Texture descriptors extracted from the wavelet transform

Information represented by spatial frequencies has often been used for texture description with successful results. Therefore, the discrete wavelet transform (DWT) has been applied to the images to characterise their textures because of its frequency domain localization capability. We have used the Haar family of wavelets, which have already been used in other medical image classification applications [8]. The DWT extracts the high-frequency components of a signal so they can be analysed separately. When the transform is applied to an image, four matrices of coefficients are obtained: approximations and horizontal, vertical and diagonal details, henceforth called LL_i, LH_i, HL_i and HH_i respectively, where i stands for the level of decomposition (Figure 3). The first one holds the low frequency components and is used for further decompositions (if any), while the other three hold the high frequency details.

The first descriptor based on the DWT is formed by the mean and standard deviation of the histograms of the original image and of each one of the matrices of coefficients yielded after three wavelet decompositions (i.e. LL_1, LH_1, HL_1, HH_1, LL_2, LH_2, HL_2, HH_2, LL_3, LH_3, HL_3 and HH_3). Hence, each region is represented by a vector of 26 features. This descriptor has been called Wavelet Statistical Features (WSF) [12,8].

We have also extracted the Haralick features *Contrast*, *Correlation*, *Energy* and *Homogeneity* from the Grey Level Co-occurrence Matrices (GLCMs) [11] of (a) the original image and (b) the coefficients obtained after a wavelet decomposition of the image (i.e. LL_1, LH_1, HL_1 and HH_1). Hence, each BG region is represented by a vector of 20 features, and the descriptor is called Wavelet Co-occurrence Features (WCF) [12,8]. All the Haralick features have been averaged



(a) (b)

Fig. 3. Example of a wavelet decomposition in one scale (b) of a BG region (a).

over the GLCM orientations 0°, 45°, 90° and 135° to provide them with a certain invariance to rotation. The impact of the GLCM distance parameter d was assessed by computing the GLCMs with distances 1 to 5.

2.5. Local binary patterns

Local binary patterns (LBPs) were introduced by Ojala et al. In their original version, they worked in a 3×3 pixel block, but they were later generalised to remove any limitations on the size of the neighbourhood or the number of sampling points [10]. Given a pixel c with coordinates (x_c, y_c), a pattern code is computed by comparing it with the value of its P neighbours as stated in Equation (1).

$$LBP_{P,R}(x_c, y_c) = \sum_{p=0}^{P-1} s(g_p, g_c) 2^p, \tag{1}$$

where g_c and g_p are the intensities of pixel c and its p-th neighbour (whose coordinates are (x_p, y_p), being $x_p = x_c + R\cos(2\pi p/P)$ and $y_p = y_c - R\sin(2\pi p/P)$), respectively. P is the total number of neighbours and R is the radius of the neighbourhood (i.e. distance from c to its neighbours). The function $s(z)$ returns 1 if $z \geq 0$ or 0 otherwise.

Finally, the whole image is described by means of a histogram of the LBP values of all pixels given by Eq. (1). We have also assessed the rotation invariant LBP (called $LBP^{ri}_{P,R}$) and rotation invariant LBP with "uniform" paterns, $LBP^{riu2}_{P,R}$ introduced in [10]. As the BG regions and EPVS are rather small, we have set the parameters $P = 1$ and $R = 8$. With this configuration, each image was described by means of 256, 36 and 10 features with the $LBP_{1,8}$, $LBP^{ri}_{1,8}$ and $LBP^{riu2}_{1,8}$ descriptors, respectively.

3. Experiments and results

3.1. Experiments

Our sample consists of 264 datasets of MRI structural scans and their corresponding visual rating scores. The latter were made by an experienced neuroradiologist (Table 1). Since a "moderate to severe" rating of EPVS in the basal ganglia is a marker of SVD [15], we dichotomised the EPVS scores into 0 (EPVS scores 0-1) and 1 (EPVS scores 2-4), as per [16]. This resulted in 133 and 131 datasets with classes 0 (negative) and 1 (positive), respectively.

Table 1. Distribution of the visual ratings in the dataset.

EPVS rating	0	1	2	3	4	TOTAL
Num. images (%)	5 (1.89 %)	128 (48.48 %)	68 (25.76 %)	44 (16.67 %)	19 (7.20 %)	264

The BG region slices selected from each scan as explained in section 2.3, characterised by means of the texture descriptors based on the Wavelet transform (i.e. WSF and WCF) and LBPs (i.e. $LBP_{1,8}$, $LBP^{ri}_{1,8}$ and $LBP^{riu2}_{1,8}$) (sections 2.4 and 2.5) and classified by means of a Support Vector Machine (SVM) with Radial Basis Function (RBF) kernel (i.e. $K(\mathbf{x}, \mathbf{x}') = \exp(-\gamma\|\mathbf{x} - \mathbf{x}'\|^2)$) [20]. A linear kernel was also assessed, but results were worse. Several combinations of the regularization parameter of the SVM, C (i.e., 1, 5, 10, 50, 100, 250 and 500) and γ of the RBF kernel (i.e., 10^{-5}, 10^{-4}, 10^{-3}, 0.01, 0.1 and 1) have been assessed with all descriptors, in order to find the optimal configuration. Classification has been carried out using stratified 10-fold cross validation. In order to avoid any possible random

effects in the selection of the folds, this procedure has been repeated 10 times, and the results presented in this paper are an average of these 10 runs. Data were normalised so that they had mean 0 and standard deviation 1.

3.2. Results and discussion

The results of the experiments have been measured in terms of accuracy, as the dataset is quite balanced, using the dichotomised visual ratings as ground truth. Table 2 shows the best results using the descriptors based on the wavelet transform (i.e. WSF, WCF) and the descriptors based on local binary patterns with $R = 1$ and $P = 8$ (i.e. $LBP_{1,8}$, $LBP^{ri}_{1,8}$ and $LBP^{riu2}_{1,8}$).

Table 2. Overall accuracy, TNR and TPR, global standard deviation (std) and the standard deviation for the negative (std_N) and positive class (std_P) of the SVM classification along the 10 iterations. Also, the parameters C and γ these results were obtained with are provided.

	C	γ	**Accuracy (%)**	**TNR (%)**	**TPR (%)**	**std.**	**std_N**	**std_P**
WCF	**500**	**0.01**	**80.03**	**79.36**	**80.67**	**1.36**	**1.89**	**1.66**
WSC	250	10^{-4}	75.07	80.90	69.25	1.27	1.29	1.61
$LBP_{1,8}$	1	10^{-3}	70.22	77.39	62.89	2.01	1.36	2.93
$LBP^{ri}_{1,8}$	10	10^{-3}	69.78	73.28	66.24	0.96	0.79	1.37
$LBP^{riu2}_{1,8}$	1	0.1	71.15	75.29	66.97	1.55	1.44	1.78

The best accuracy was obtained by WCF (80.03%) – achieved with a GLCM distance $d = 3$, $C = 500$ and $\gamma = 0.01$. The other wavelet-based descriptor (WSF) obtained 75.07% – $C = 250$ and $\gamma = 10^{-4}$. Moreover, WCF achieves balanced true positive rate (TPR) (i.e. images with moderate to severe EPVS correctly classified) and true negative rate (TNR) (i.e. images with non or mild EPVS correctly classified) whereas in the others they are very imbalanced (Figure 4). The TPR achieved by WCF is slightly higher than TNR, whereas it is quite lower in the other cases.

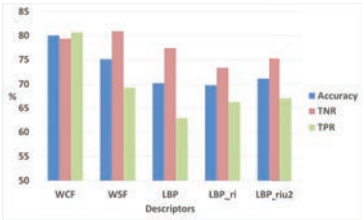


Fig. 4. Global accuracies, TNR and TPR of the different descriptors.

We are aware that this accuracy may not seem high enough, but to the best of our knowledge there are no other methods that apply pattern recognition to automatically assign a rating for the density of perivascular spaces. This being the first attempt to automatically classify EPVS burden in brain MRI, we consider our approach a promising starting point. It may seem that we are missing information by extracting 2D descriptors from a single slice of a 3D volume. However, we based this decision on a clinically proven visual rating scale[1] and validated guidelines[4]. In addition, the relatively low resolution in the slice direction (5 mm) compared to the size of the EPVS (1 to 5 mm diameter) makes it unrealistic to use a 3D approach to characterise these structures.

4. Conclusion

In this paper we present an automatic method to assess the density of EPVS in the BG (one of the markers of small vessel disease) on brain MRI into (a) none or mild and (b) moderate to severe.

The BG region is described by means of (a) texture descriptors based on the first and second order statistical descriptors extracted from the results of the DWT (WSF and WCF respectively), and (b) local binary patterns (LBP), "rotation invariant" (LBP^{ri}) and "uniform rotation invariant" (LBP^{riu2}) LPBs. These descriptors have been classified by a support vector machine with radial basis function kernel.

WCF achieves the highest accuracy (80.03%), improved balance between the accuracies in the detections of both classes and a slightly higher accuracy in the detection of the brains with moderate to severe EPVS than the other

descriptors. This is the first time that this type of automatic rating has been carried out, so 80.03% is a promising starting point. A next step will be to evaluate this approach against the output from other methods that quantify EPVS[5,4,2] and to see if it works in other brain regions such as centrum semiovale.

A limitation of this work may be the fact that the ground truth we are using is a visual (and, therefore, subjective) rating. However, the visual rating scale has shown a high inter-rater reliability in previous works[1] so this should not affect the validity of the results shown in this paper too much. Future work should compare the inter-rater agreement with the performance of our method using the ratings from a second rater on this sample, not available at the moment. In addition, we could try to avoid the selection of the relevant BG slice by fusing descriptors computed from each slice and selecting the most relevant ones by means of an automatic feature selection method.

Acknowledgements

We would like to thank Dr. Stephen Makin (patient recuitment), study participants, radiographers and staff at the Brain Research Imaging Centre Edinburgh, a SINAPSE (Scottish Imaging Network A Platform for Scientific Excellence) collaboration centre, the Wellcome Trust for funding the primary study that provided the data (Ref. No. 088134/Z/09) and the Row Fogo Charitable Trust (Grants Nos. R35865 and R43412).

References

1. Potter, G.M., Chappell, F.M., Morris, Z., Wardlaw, J.M.. Cerebral perivascular spaces visible on magnetic resonance imaging: development of a qualitative rating scale and its observer reliability. *Cerebrovasc Dis* 2015;**39**(3-4):224–231.
2. Valdés Hernández, M.d.C., Piper, R.J., Wang, X., Deary, I.J., Wardlaw, J.M.. Towards the automatic computational assessment of enlarged perivascular spaces on brain magnetic resonance images: a systematic review. *J Magn Reson Imaging* 2013;**38**(4):774–785.
3. Ramirez, J., Berezuk, C., McNeely, A.A., Scott, C.J.M., Gao, F., Black, S.E.. Visible Virchow-Robin spaces on magnetic resonance imaging of Alzheimer's disease patients and normal elderly from the Sunnybrook Dementia Study. *J Alzheimers Dis* 2015;**43**(2):415–424.
4. Wang, X., Valdés Hernández, M.D.C., Doubal, F., Chappell, F.M., Piper, R.J., Deary, I.J., et al. Development and initial evaluation of a semi-automatic approach to assess perivascular spaces on conventional magnetic resonance images. *J Neurosci Methods* 2016;**257**:34–44.
5. Cai, K., Tain, R., Das, S., Damen, F.C., Sui, Y., Valyi-Nagy, T., et al. The feasibility of quantitative MRI of perivascular spaces at 7T. *J Neurosci Methods* 2015;**256**:151–156.
6. Ithapu, V., Singh, V., Lindner, C., Austin, B.P., Hinrichs, C., Carlsson, C.M., et al. Extracting and summarizing white matter hyperintensities using supervised segmentation methods in Alzheimer's disease risk and aging studies. *Human brain mapping* 2014;**35**(8):4219–4235.
7. de Boer, R., Vrooman, H.A., van der Lijn, F., Vernooij, M.W., Ikram, M.A., van der Lugt, A., et al. White matter lesion extension to automatic brain tissue segmentation on MRI. *Neuroimage* 2009;**45**(4):1151–1161.
8. Alegre, E., González-Castro, V., Alaiz-Rodríguez, R., García-Ordás, M.T.. Texture and moments-based classification of the acrosome integrity of boar spermatozoa images. *Comput Methods Programs Biomed* 2012;**108**(2):873–881.
9. Chen, L., Tong, T., Ho, C.P., Patel, R., Cohen, D., Dawson, A.C., et al. Identification of cerebral small vessel disease using multiple instance learning. In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Springer; 2015, p. 523–530.
10. Ojala, T., Pietikainen, M., Maenpaa, T.. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2002;**24**(7):971–987.
11. Haralick, R.M., Shanmugam, K., Dinstein, I.. Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics* 1973;**SMC-3**(6):610–621.
12. Arivazhagan, S., Ganesan, L.. Texture classification using wavelet transform. *Pattern recognition letters* 2003;**24**(9):1513–1521.
13. Sørensen, L., Shaker, S.B., De Bruijne, M.. Texture classification in lung CT using local binary patterns. In: *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2008*. Springer; 2008, p. 934–941.
14. Valdés Hernández, M.d.C., Armitage, P.A., Thrippleton, M.J., Chappell, F., Sandeman, E., Muñoz Maniega, S., et al. Rationale, design and methodology of the image analysis protocol for studies of patients with cerebral small vessel disease and mild stroke. *Brain Behav* 2015; **5**(12):e00415.
15. Staals, J., Makin, S.D.J., Doubal, F.N., Dennis, M.S., Wardlaw, J.M.. Stroke subtype, vascular risk factors, and total MRI brain small-vessel disease burden. *Neurology* 2014;**83**(14):1228–1234.
16. Potter, G.M., Doubal, F.N., Jackson, C.A., Chappell, F.M., Sudlow, C.L., Dennis, M.S., et al. Enlarged perivascular spaces and cerebral small vessel disease. *Int J Stroke* 2015;**10**(3):376–381.
17. Lutkenhoff, E.S., Rosenberg, M., Chiang, J., Zhang, K., Pickard, J.D., Owen, A.M., et al. Optimized brain extraction for pathological brains (optiBET). *PLoS One* 2014;**9**(12):e115551.
18. Zhang, Y., Brady, M., Smith, S.. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Transactions on Medical Imaging* 2001;**20**(1):45–57.
19. Zuiderveld, K.. Contrast limited adaptive histogram equalization. In: *Graphics gems IV*. Academic Press Professional, Inc.; 1994, p. 474–485.
20. Vapnik, V.. *The Nature of Statistical Learning Theory*. Springer; 2nd ed.; 1995.