USING LOCAL BINARY PATTERN TO CLASSIFY DEMENTIA IN MRI

K. Oppedal[⋆†]*, K. Engan*[⋆]*, D. Aarsland*[†]*, M. Beyer*[†]*, O.B. Tysnes*[‡]*, T. Eftestøl*[⋆]

[⋆]Dept. of Electrical Engineering and Computer Science, University of Stavanger, Stavanger, Norway
[†]Centre for age-related medicine, Stavanger University Hospital, Stavanger, Norway
[‡]Dept. of Neurology, Haukeland University Hospital, Bergen, Norway

ABSTRACT

White matter lesions (WML) are hyperintense signals in T2-weighted MRI of the brain. Volume and regional distribution of WML have been extensively studied in dementia, but not much attention has been given to texture analysis in these regions. We wanted to explore if it is possible to distinguish patients with dementia from healthy elderly in a classification framework testing different texture features in the WML regions, paying special attention to a feature called Local Binary Pattern (LBP). The results presented here, indicates that the LBP features used in our experiment are powerful features in a maximum likelihood classifier, when classifying demented from normal controls.

Index Terms— MRI, texture, LBP, classification, dementia

1. INTRODUCTION

White matter lesions (WML) are localized areas of abnormally high signal, in the brain white matter, visible on T2-weighted images. They are typically seen around the ventricles (periventricular WML), but also as focal lesions in the deep white matter. It is generally considered that WML is an indicator of small-vessel cerebrovascular disease (CVD), but other mechanisms may contribute as well and vary among individuals. WML occur with increased frequencies with increasing age in healthy subjects, but they are also found to be associated with vascular dementia, and Alzheimer disease (AD).

Successful automatic classification of patients with dementia from normal controls, have previously been done using pattern recognition techniques on magnetic resonance images (MRI) [1], where the brain MRI was segmented and the gray matter image intensities used as feature vectors in a support vector machine (SVM) classifier. Little is done in evaluating features based on image processing of the WML regions in MR images of the brain when diagnosing dementia, other than finding the *volume sizes* or regional distribution of the WML in the different regions of the brain. When WML are segmented in MR images, many features can be calculated from the WML regions and from the normal appearing white matter. An interesting question is if it is possible to distinguish different mechanisms, i.e. patients with dementia from healthy elderly, based on textural information in the WML regions. In this paper we apply an established texture feature called Local Binary Pattern (LBP) [2], an interesting feature that earlier has shown high discriminative power in medical imaging [3]. To evaluate the discriminative power of a feature or a combination of several features, we use a classification framework earlier published by Engan et al. [4]. We hypothesized that classification based on pattern recognition methods using regional WML volumes as features, would be able to discern demented from normal controls, and that a classifier using local binary pattern (LBP) texture features is able to discern demented from normal controls better than the calculated volume alone.

Helse Vest AS for funding

2. MATERIAL AND METHODS

2.1. Subjects

We recruited 245 patients with a first time diagnosis of mild dementia from three outpatient dementia clinics in Western Norway. A clinician performed a structured clinical interview of demographic, and clinical data with the patients and caregivers. 215 subjects fulfilled inclusion criteria and constituted the baseline population. Patients with acute delirium, terminal illness, recently diagnosed with a major somatic illness, previous bipolar disorder or psychotic disorder were excluded. Diagnoses were made after a detailed assessment, including the use of standardized clinical rating scales. Two of the investigators independently applied the diagnostic criteria after baseline and 2 years later. In cases of disagreement, and in patients fulfilling more than one set of operationalized diagnostic criteria, the final ascertainment was made based on consensus. Seven patients have come to autopsy, and the pathological diagnosis has been consistent with the clinical diagnosis in all cases. Further details of the recruitment process and diagnoses are provided elsewhere [5]. A control group consisting of 32 healthy elderly (NC) matched for sex, age, and education were selected from the Norwegian Park-

West study [6]. The NC were scanned with the same protocol and at the same scanners as the dementia patients. The study was approved by The Regional Committee for Medical Research Ethics, University of Bergen, Norway. All participating patients signed informed consent to participate in the study after the study procedures had been explained in detail to the patient and a caregiver, usually the spouse or offspring.

2.2. WML volume quantification

The subjects were scanned at three different sites. A 1.5 T scanner was used in all three centers, using the same scanner in each center during the entire study period, and a common study imaging protocol. In some cases, scans were excluded due to insufficient image quality, not having both FLAIR and T1 images for the patient, or movement artifacts and other artifacts. A total of 63 mild dementia subjects were included in this study, 52 with AD and 11 with Lewy Body Dementia.

Image analysis was performed according to a method developed and previously published by MJ Firbank et al. [7]. Briefly, the non-brain regions were removed from the T1 image, using the segmentation routines in SPM5 (`http://www.fil.ion.ucl.ac.uk/spm`), and the WML were segmented on a slice-by-slice basis from the FLAIR image, with the images in native space, using a threshold determined from the histogram of pixel intensities for each image slice. To explore the regional distribution of WML throughout the brain, a region of interest (ROI) template in standard MNI space was used (Montreal Neurological Institute, http://www.bic.mni.mcgill.ca). This ROI template was transformed from MNI space to the image space (FLAIR) of each subject by use of the normalization routines in SPM5, and the volumes of WML in each ROI was calculated. The ROI map was based on the Brodmann template, which is provided as part of MRICro (`http://www.psychology.nottingham.ac.uk/staff/crl/lesion.html#brod`). Further details can be found in [7].

2.3. LBP texture feature

Ojala et al. [2] was the first to describe LBP as a texture feature. It is a simple yet very efficient texture operator which lables the pixels of an image by thresholding the neighborhood of each pixel and considers the result as a binary number, see Figure 2.3 for details. Since its discriminative power has shown to be high and at the same time computationally simple, LBP texture operators are a popular approach in various applications, and unifies traditionally divergent statistical and structural models of texture analysis. Adding an image contrast measure (C) as well as varying the texture neighborhood enhances the discriminative power of the LBP feature even further. In [8], Unay et al. demonstrated that the rotation invariant LBP is invariant to some common MRI artifacts.

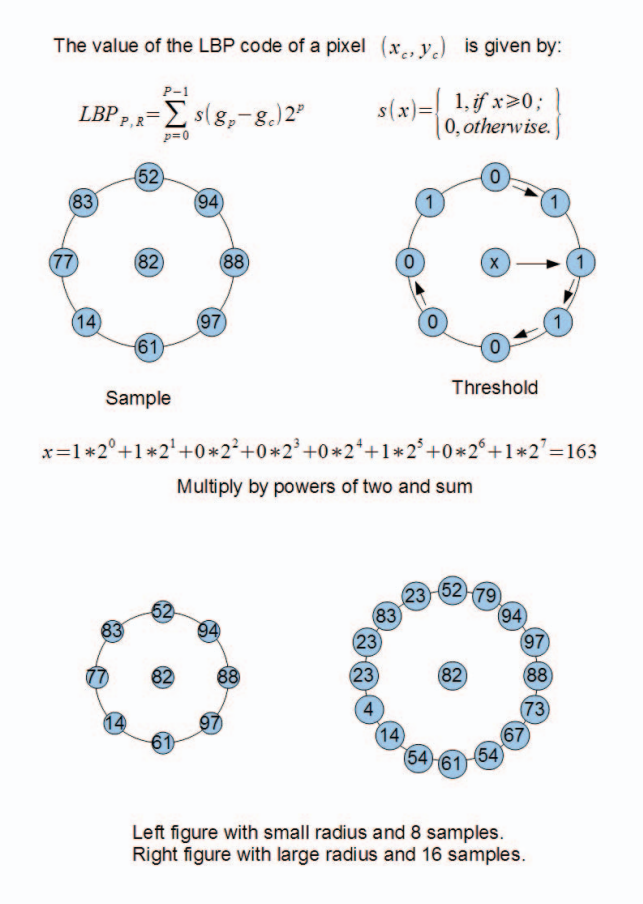


Fig. 1. Top figure demonstrates LBP thresholding. A neighbor to a center pixel is set to one if it has equal or higher pixel value and zero if it has lower. In an anti-clockwise manner every neighbor is multiplied by powers of two and summed as demonstrated in the equation below. The bottom figure demonstrates how the radius and number of samples can be varied in the choice of neighborhood.

2.4. ML classification and ROC analysis

On the basis of a training data set with observations whose sub-population is known, statistical classification makes it possible to identify the true sub-population of new observations. Because of its simplicity, computationally speed, and good convergence properties with increasing training samples, a maximum likelihood (ML) classifier was applied in this study. When applied to a data set and a given statistical model, ML estimation provides estimates for the model's parameters, assuming a normal distribution. For a fixed set of data and underlying statistical model, the method of ML selects values of the model parameters that produces a distribution that gives the observed data the greatest probability, i.e. parameters that maximize the likelihood function. Since we have a small dataset, a leave-one-out technique was used

as a strategy for cross-validation.

A receiver operating characteristic, or ROC curve, is a graphical plot of the sensitivity, or true positive rate, vs. false positive rate, for a binary classifier system as its discrimination threshold is varied. ROC analysis provides tools to select optimal models independently from the cost context or the class distribution. To handle the scarce data set a method earlier published by Eftestøl [9], was used to control the true positive rate in the ROC analysis. Area under curve (AUC), was used to quantify the performance of the classifier. The AUC is equal to the probability that a classifier will rank a randomly chosen positive instance higher than a randomly chosen negative one. AUC close to one is deemed as excellent performance of the classifier model.

3. EXPERIMENTS AND RESULTS

The total- and regional WML volumes in various regions were used as features. Total brain-, total periventricular-, and total deep white matter WML volume, as well as the temporal- (Te), occipital- (Oc), frontal- (Fr), parietal- (Pa), cerebellar- (CB), basal ganglia- (BG), central white matter (CW), and corpus callosum (CC) in each of the three regions were used to form 27 features.

Mean, standard deviation, and entropy of the grayscale intensities in the concatenation of all the WML areas of a patient, were used to form three other features.

LBP and C values was calculated slice-wise for all pixels in the MR image for each subject using the matlab functions developed by Ojala et al. [10]. The WML regions was defined as region of interest (ROI), and only the LBP and C values therein was collected. LBP and C values for four different neighborhoods or combinations of radi from the centre pixel and number of samples in the neighborhood (r,p) were calculated, namely (1,8), (2,8), (4,16), and (8,16), collecting only the uniform and rotationally invariant LBP and C values. To incorporate the LBP with the C values a ratio of the two was also calculated. From these collections of LBP, C, and LBP/C values, the mean, median, standard deviation, entropy, skewness, and kurtosis were calculated to form 18 scalar features. The features were named using a prefix defining which combination of radius and number of samples in the neighborhood was used (f.eks. r1p8), an underscore, a letter (i.e. l for LBP, c for contrast, or lrc for LBP/C), following the name of the measure (i.e. mean, standard deviation, median, entropy, skewness, or kurtosis), to form f.eks. "r1p8_lmean".

In the first experiment, all possible 1D features were tested classifying the demented from the normal controls. The results using the WML volumes as features are presented in Table 1 with AUC values and their 95% confidence interval, AUC (CI_{95}). The confidence interval will seem wider than expected, since it was calculated for every point on the ROC curve. AUC equal to 0.73 (0.57, 0.83) were achieved for the total cerebellar volume, the cerebellar volume in deep white matter, and the frontal volume in the periventricular area. The AUC from the mean, standard deviation, and entropy of the grayscale values in the ROI were all close to or below 0.6. Because of the unpromising result, these features were not further used. The highest AUC values from the LBP features, were median of the LBP calculated with radius two using eight samples (r2p8_lmed) and similar with radius four using 16 samples with AUC equal to 0.86 (0.68, 0.93) and 0.82 (0.64, 0.91) respectively as shown in Table 2.

Feature	AUC (CI_{95})
[CB_tot]	0.73 (0.57, 0.83)
[CB_dwm]	0.73 (0.57, 0.83)
[Fr_pvh]	0.73 (0.54, 0.86)
[BG_dwm]	0.71 (0.52, 0.84)
[BG_pvh]	0.71 (0.51, 0.84)
[pvh]	0.70 (0.51, 0.84)

Table 1. The six 1D volume feature vectors giving the highest AUC classifying normal controls and demented.

Feature	AUC (CI_{95})
[r2p8_lmed]	0.86 (0.68, 0.93)
[r4p16_lmed]	0.82 (0.64, 0.91)
[r1p8_lmean]	0.79 (0.61, 0.90)
[r4p16_lent]	0.79 (0.61, 0.90)
[r2p8_lent]	0.79 (0.60, 0.90)
[r2p8_lmean]	0.77 (0.58, 0.89)
[r2p8_lskew]	0.75 (0.56, 0.88)
[r1p8_lskew]	0.75 (0.56, 0.88)
[r8p16_cmean]	0.75 (0.56, 0.87)
[r1p8_lent]	0.74 (0.55, 0.86)
[r8p16_cmed]	0.72 (0.53, 0.85)
[r8p16_lent]	0.70 (0.52, 0.83)

Table 2. The twelve 1D feature vectors giving the highest AUC classifying normal controls and demented.

The best performing 1D features were combined with all other possible 1D features to generate 2D feature vectors which were tested in the same manner as above, and provided a significant improvement over 1D features. The four 2D feature vector with the highest AUC were [r4p16_lent, r8p16_cmed] with AUC equal to 0.90 (0.74, 0.96), [r2p8_lmean, r8p16_cmed] with AUC equal to 0.89 (0.72, 0.96), [r2p8_lmed, r8p16_cmean] with AUC equal to 0.89 (0.72, 0.96), and [r1p8_lmean, r8p16_cmed] with AUC equal to 0.89 (0.72, 0.96), see Table 3 for further details.

In the same manner 3D feature vectors were made combining the best performing 2D feature vector with all other 1D feature vectors only performing marginally better than the best 2D features. The two 3D feature vectors with the highest AUC were [r4p16_lent, r8p16_cmed, r8p16_lmed] with

Feature	AUC (CI_{95})
[r4p16_lent, r8p16_cmed]	0.90 (0.74, 0.96)
[r2p8_lmean, r8p16_cmed]	0.89 (0.72, 0.96)
[r2p8_lmed, r8p16_cmean]	0.89 (0.72, 0.96)
[r1p8_lmean, r8p16_cmed]	0.89 (0.72, 0.96)
[r1p8_lmean, r4p16_cmean]	0.88 (0.71, 0.95)
[r4p16_lent, r8p16_cmean]	0.88 (0.71, 0.96)
[r8p16_lent, r8p16_cmed]	0.88 (0.71, 0.96)
[r2p8_lmed, r4p16_cmean]	0.88 (0.71, 0.95)

Table 3. The eight 2D feature vectors giving the highest AUC classifying normal controls and demented.

AUC equal to 0.91 (0.75, 0.97) and [r4p16_lent, r8p16_cmed, r8p16_lkurt] with AUC equal to 0.90 (0.74, 0.96) which, see Table 4.

Feature	AUC (CI_{95})
[r4p16_lent, r8p16_cmed, r8p16_lmed]	0.91 (0.75, 0.97)
[r4p16_lent, r8p16_cmed, r8p16_lkurt]	0.90 (0.74, 0.96)

Table 4. The two 3D feature vectors giving the highest AUC classifying normal controls and demented.

4. DISCUSSION

Based on image processing of WML in MR images of demented and normal controls, the LBP features used in our experiment are proposed as powerful features in a maximum likelihood classifier, when classifying demented from normal controls.

Earlier Kloppel et al. [1] successfully used whole brain images as features in a SVM classifier to distinguish subjects suffering from Alzheimer disease from healthy elderly. To our knowledge though, regional WML volumes, LBP, and C values calculated from WML regions in MR images have not earlier been used as features to classify demented from normal controls. Since the distribution of the features in the dataset may not be normal, this approach can underestimate the true AUC values.

The small dataset of 95 subjects is a limitation of the study, but a leave-one-out cross-validation technique is adopted to exploit the dataset as good as possible. It is not certain that the features we used are normally distributed, and thus it is possible that the results would improve if a non-parametric classification technique was adopted.

5. REFERENCES

[1] Kloppel S, Stonnington CM, Chu C, Draganski B, Scahill RI, Rohrer JD, Fox NC, Jack CR Jr, Ashburner J, and Frackowiak RS, "Automatic classification of MR scans in Alzheimer's disease," *Brain*, vol. 131, pp. 681–689, March 2008.

[2] Ojala T, Pietikáinen M, and Harwood D, "A comparative study of texture measures with classification based on feature distributions," *Pattern Recognition*, vol. 29, no. 1, pp. 51–59, 1996.

[3] Nanni L, Lumini A, and Brahnam S, "Local binary pattern variants as texture descriptors for medical image analysis," *Artificial Intelligence in Medicine*, vol. 49, pp. 117–125, 2010.

[4] Engan K, Eftestøl T, Ø rn S, Kvaløy JT, and Woie L, "Exploratory data analysis of image texture and statistical features on myocardium and infarction areas in cardiac magnetic resonance images," in *32nd Annual International Conference of the IEEE EMBS*, Buenos Aires, Argentina, September 2010.

[5] Aarsland D, Rongve A, Nore SP, Skogseth R, Skulstad S, and Ehrt U et al., "Frequency and case identification of dementia with lewy bodies using the revised consensus criteria," *Dement Geriatr Cogn Disord*, vol. 26, no. 5, pp. 445–452, 2008.

[6] Alves G, Muller B, Herlofson K, HogenEsch I, Telstad W, and Aarsland D et al., "Incidence of Parkinson's disease in norway: the Norwegian ParkWest stydy," *J Neurol Neurosurg Psychiatry*, vol. 80, no. 8, pp. 851–857, 2009.

[7] Firbank MJ, Minett T, and O'Brien JT, "Changes in DWI and MRS associated with white matter hyperintensities in elderly subjects," *Neurology*, vol. 61, no. 7, pp. 950–954, 2003.

[8] Unay D, Ekin A, Cetin M, Jasinschi R, and Ercil A, "Robustness of local binary patterns in brain MR image analysis," in *Proceedings of the 29th Annual International Conference of the IEEE EMBS*, August 2007.

[9] Trygve Eftestøl, "Controlling true positive rate in ROC analysis," in *IEEE Computer in Cardiology Conference Proceedings*, September 2009.

[10] Ojala T, Pietikáinen M, and Máenpáá T, "A generalized Local Binary Pattern operator for multiresolution gray scale and rotation invariant texture classification," in *Advances in Pattern Recognition, ICAPR 2001 Proceedings, Lecture Notes in Computer Science*, 2001.