21st International Conference on Pattern Recognition (ICPR 2012)
November 11-15, 2012. Tsukuba, Japan

Using local texture maps of brain MR images to detect Mild Cognitive Impairment

Rita Simões*, Cornelis Slump* and Anne-Marie van Cappellen van Walsum†‡
*Signals and Systems Group, University of Twente, The Netherlands
Contact e-mail: a.r.lopessimoes@utwente.nl
†Department of Anatomy, Radboud University Nijmegen Medical Centre, The Netherlands
‡MIRA Institute for Biomedical Technology and Technical Medicine, University of Twente, The Netherlands

Abstract

Early detection of Alzheimer's disease is expected to aid in the development and monitoring of more effective treatments. Classification methods have been proposed to distinguish Alzheimer's patients from normal controls using Magnetic Resonance Images. However, their performance drops when classifying patients at a prodromal stage, such as in Mild Cognitive Impairment. Most often, the features used in these classification tasks are related to structural measures such as volume, shape and tissue density. However, microstructural changes have been shown to arise even earlier than these larger-scale alterations. Taking this into account, we propose the use of local statistical texture maps that make no assumptions regarding the location of the affected brain regions. Each voxel contains texture information from its local neighborhood and is used as a feature in the classification of normal controls and Mild Cognitive Impairment patients. The proposed approach obtained an accuracy of 87% (sensitivity 85%, specificity 95%) with Support Vector Machines, outperforming the 63% achieved by the local gray matter density feature.

1. Introduction

Alzheimer's Disease (AD) is the most common type of dementia and a major cause of disability worldwide [11]. Early detection of AD is essential to provide the patients with adequate and timely treatments and to help researchers monitor their effectiveness. Structural Magnetic Resonance Imaging (MRI) is a diagnostic tool that provides high-resolution images and a high brain tissue contrast. In addition, its non-invasiveness makes it a suitable imaging technique for follow-up studies.

A limitation of most state-of the-art MR image analysis methods in this field is that they often concern only group comparisons. Although these methods can provide a description of the location and magnitude of statistically significant differences between two groups, they have limited clinical value for individual patients.

This limitation has led to the development of classification methods to identify Alzheimer's patients from Normal Controls (NC) and, more recently, to distinguish NC from patients suffering from Mild Cognitive Impairment (MCI), which indicates high risk of developing Alzheimer's. As pointed out by a recent comparison study on various classification methods [2], the current major challenge is to discriminate patients who are at a very early stage of AD or even possibly before they start developing the disease. As shown by the comparison results, the performance of most classifiers dropped significantly when they attempted to classify between NC and MCI.

Typically, the features used by these classification methods concern the volume and/or the shape of specific brain structures, like the hippocampus [2]. Voxel-Based Morphometry (VBM) approaches have also been used, which analyze the local concentration of gray matter [2, 9].

However, such tools are limited by the segmentation quality of the structures of interest. Furthermore, it has been shown that the brain microstructure starts to deteriorate several years before the first symptoms arise and before structural alterations can be detected [6].

Texture Analysis (TA) is an image processing tool that has recently found applications in the study of various neurological diseases, including Alzheimer's. It extracts information that is otherwise not visible by a direct analysis of the image intensity and shape properties. In [5], the authors performed 2D texture analysis using the entire brain to classify between AD and NC. Because the whole brain was used, no discrimination between significant regions was performed. In [12], Zhang *et al*. also classified patients as AD or NC using 3D texture features computed at manually de-

978-4-9906441-1-6

fined spherical Regions of Interest (ROIs), in the hippocampus and the entorhinal cortex. However, and as the authors recognized, the results varied significantly with the location and the size of the chosen ROI. Furthermore, in neither of these two studies was an analysis with MCI patients performed. Other studies have carried out texture analysis in the corpus callosum and thalamus [4]. In all cases, the texture descriptors are computed at manually segmented ROIs, thereby requiring a priori knowledge about the disease and becoming dependent on the quality of the segmentations. Also, to the best of our knowledge, no comparisons between the two approaches (structural and textural) have been performed.

In this work, we propose the use of *local* statistical (co-occurrence matrix based) texture maps as features to be used in the classification of NC and MCI. In these maps, each voxel contains texture information from its local neighborhood and is considered as a feature for classification. We perform a statistical significance analysis on these voxels as a feature selection step. Finally, we use Support Vector Machines (SVM) in a cross-validation scheme to classify the subjects. We compare our method with a structural approach that uses, as features, the voxels in the gray matter probability map [9].

Our contributions are the following: application of local statistical texture maps to the classification of NC and MCI, which make no assumption about the expected location of significant differences and that require no previous segmentation of brain structures; performance comparison of the proposed features and a widely used structural feature - the local gray matter density. To the best of our knowledge, no other texture studies have made such comparison.

2. Methods

2.1 Calculation of the feature maps

The Haralick features are based on the Gray Level Co-occurrence Matrix (GLCM), which gives information about the statistical distribution of voxel intensity pairs [7]. In this work, we refer to these texture descriptors by the following: F1 - angular second moment; F2 - contrast; F3 - correlation; F4 - sum of squares; F5 - inverse difference moment; F6 - sum average; F7 - sum variance; F8 - sum entropy; F9 - entropy; F10 - difference variance; F11 - difference entropy. For a complete description of the features, we refer the reader to [7].

As in previous texture studies [12, 4], we compute the first eleven Haralick features (according to [7]) at a $3 \times 3 \times 3$ sliding window centered on each brain voxel. This allows for texture analysis in the entire brain rather than at specific ROIs. The GLCM is determined for all 13 three-dimensional directions, considering voxel pairs at a distance of 1 voxel. In order to increase the computational speed of these calculations, and to avoid very sparse GLCMs, we quantize the original image intensities to 5 bits (range $[0, 31]$). After texture feature calculation, we obtain, for each subject, 11 feature maps.

3 Experiments and Results

3.1 Data and preprocessing

For this study, datasets from 15 Normal Controls (75.4 ± 4.5 years, 8 males and 7 females) and 15 MCI patients (73.3 ± 8.2 years, 10 males and 5 females) were retrieved from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database [8]. The data consist of three-dimensional T1 images acquired at 3T. These images have been previously corrected for acquisition artifacts such as bias field inhomogeneities, geometric distortions and scaling, as described in [8]. To eliminate global differences between brain shapes and volumes, we align all images to the same spatial reference using a non-linear diffeomorphic registration method, DARTEL [1].

3.2 Feature maps

We then register the features maps into the template space, by applying the same warp field that originated from the non-linear registration of the T1 images. An 8mm (FWHM) isotropic Gaussian kernel is finally applied to smoothen the aligned feature maps.

To obtain the gray matter density feature maps, we first segment the brain images using the probabilistic segmentation method offered by SPM8 (Wellcome Trust Centre for Neuroimaging, http://www.fil.ion.ucl.ac.uk/spm). Then, and similarly to what is done with the texture maps, we apply the respective warp field obtained in the non-linear registration step to the gray matter segmentations, followed by Gaussian smoothing. The resulting maps represent the local concentration of gray matter per voxel. Two-dimensional slices of all obtained feature maps are shown in Figure 1.

3.3 Classification

We use an SVM (implemented in the Python package *scikits-learn* [10]) to classify the datasets into one of the two classes: NC or MCI. To better evaluate the classifier's generalizability, we perform a random subsampling evaluation with 10 random permutations, in which the test set corresponds to 10% of the data samples. At each training fold, we carry out an analysis of

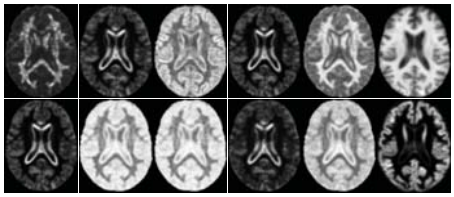

Figure 1: Feature maps of an MCI patient after non-linear registration to a common spatial reference. Bottom right: gray matter tissue probability map.

variance (ANOVA) test on the training samples and select the 5% most significant voxels, which are then used as features in the classification task. We perform a grid search (with 5-fold cross-validation on the training set) for the best SVM parameters: kernel type - linear or Radial Basis Function (RBF); the cost C and, for the RBF kernel, the scale γ. The best classifier is then evaluated on the test set. The final performance measures (accuracy, sensitivity and specificity) are computed as the average of the values obtained at each evaluation fold.

The classification results are shown in Figures 2 and 3. The texture descriptor with the best performance (F3 - correlation) achieved a mean accuracy (percentage of correctly classified subjects) of 87%, at a sensitivity of 85% and a specificity of 95%. In contrast, the accuracy of the structural feature was 63%, with 75% sensitivity and 55% specificity. A Wilcoxon-Mann-Whitney statistical significance test on the evaluation folds' results showed that feature F3 significantly outperformed the gray matter density feature in terms of accuracy ($p = 0.007$) and specificity ($p = 0.01$). Feature F8 (sum entropy) showed also, at a high significance level ($p = 0.06$), a better accuracy than the structural feature.

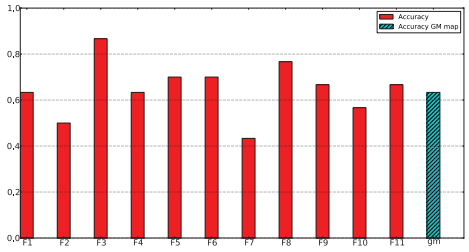

Figure 2: Mean accuracy obtained using the first 11 Haralick features and the Gray Matter (GM) tissue probability maps (rightmost blue hatched bar).

In addition, we show the brain voxels that were selected by the ANOVA test in one of the training folds (Figure 4a). We observe that using the correlation (F3) map voxels in the left hippocampus are detected as being statistically significant (and consequently used in the classification). Voxels in the brain ventricles, particularly near the edges, are also selected, as well as in the white matter and near the lateral sulcus. The higher accuracy of the classification using this feature map, when compared to the structural feature, indicates that these regions might play a role in MCI, even though their corresponding gray matter density is not significantly different between the groups. As a comparison, we show, in Figure 4b), the statistical differences between the same NC group and a group of 12 AD patients, where we clearly see, for both feature types, the two hippocampi being selected (the left being more significant). MRI signal changes which do not correlate with structural measurements have already been observed in ageing subjects [3]. The underlying cause for these alterations lies probably in the change of water, protein and mineral content of the tissues. A similar explanation can be given to why texture descriptors might be able to capture early signals of dementia.

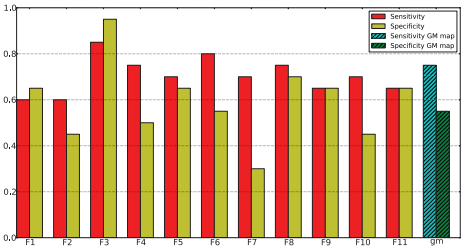

Figure 3: Mean sensitivitiy and specificity obtained using the first 11 Haralick features and the Gray Matter (GM) tissue density maps (rightmost green and blue hatched bars).

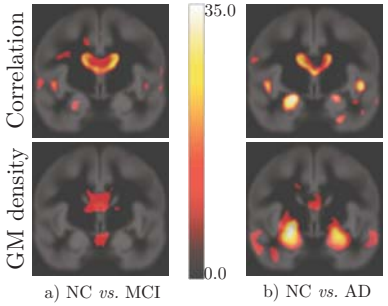

Figure 4: F-values of the statistical test (only the 5% most significant voxels are shown) for the correlation map and the GM local density map.

A final analysis was performed on the effect of varying the percentile of features selected for classification, although no significant changes in classification performance were obtained.

4. Conclusions and recommendations

In this work, we have analyzed the performance of statistical texture maps in classifying MCI patients and

normal elderly controls. We used a whole-brain voxel-wise approach, in which we made no assumptions about the expected location of differences between the two subject groups.

We obtained a mean accuracy of 87% (sensitivity of 85% and specificity of 95%) when using the correlation map voxels as features in an SVM classification task, outperforming the structural feature map - the local gray matter density. Remarkably, the voxels selected using the two feature maps were not the same, suggesting that texture- and structure-based features might be sensitive to distinct aspects of the disease. In particular, part of the left hippocampus was selected when using the texture map but not with the GM density map, possibly indicating an earlier sensitivity of the texture descriptor to changes in this region.

Further work will include a more thorough evaluation of other classifiers and feature selection/extraction methods. Also, the influence of the preprocessing steps on the classification performance should be assessed. This includes the non-linear registration to the common spatial reference and the smoothing applied to the registered feature maps.

The influence of the size of the local window chosen to compute the features should be evaluated. In this work, we focused on very fine-scale statistical textures. A multi-scale analysis will provide further insight on also larger-scale texture properties. Other feature types, such as higher-order statistical features and Gabor wavelets, as well as combinations of various features, need also be considered.

Additionally, a comparison between the results obtained with images acquired at 3T and at the most commonly available 1.5T is desirable. In particular, it is worth investigating how both structure- and texture-based features perform at the two field strengths. Similarly, other MRI modalities (such as T2 images) should be considered.

Finally, the classification must be performed with a larger number of samples to allow for stronger conclusions. However, these preliminary results seem to indicate that microstructural information, such as that provided by local texture descriptors, can play a useful role towards better and earlier detection of Alzheimer's disease.

5 Acknowledgements

This work is part of the VIP-BrainNetworks project, which is funded by the department of Economic Affairs of the Netherlands and the provinces of Gelderland and Overijssel. The authors would also like to acknowledge the Alzheimer's Disease Neuroimaging Initiative (ADNI) for providing the data.

References

[1] J. Ashburner. A fast diffeomorphic image registration algorithm. *Neuroimage*, 38(1):95–113, Oct 2007.

[2] R. Cuingnet, E. Gerardin, J. Tessieras, G. Auzias, S. Lehricy, M.-O. Habert, M. Chupin, H. Benali, O. Colliot, and A. D. N. Initiative. Automatic classification of patients with Alzheimer's disease from structural MRI: a comparison of ten methods using the ADNI database. *Neuroimage*, 56(2):766–781, May 2011.

[3] C. Davatzikos and S. M. Resnick. Degenerative age changes in white matter connectivity visualized in vivo using magnetic resonance imaging. *Cereb Cortex*, 12(7):767–771, Jul 2002.

[4] M. S. de Oliveira, M. L. F. Balthazar, A. D'Abreu, C. L. Yasuda, B. P. Damasceno, F. Cendes, and G. Castellano. MR imaging texture analysis of the corpus callosum and thalamus in amnestic Mild Cognitive Impairment and mild Alzheimer disease. *AJNR Am J Neuroradiol*, 32(1):60–66, Jan 2011.

[5] P. A. Freeborough and N. C. Fox. MR image texture analysis applied to the diagnosis and tracking of Alzheimer's disease. *IEEE Trans Med Imaging*, 17(3):475–478, 1998.

[6] G. B. Frisoni, N. C. Fox, C. R. Jack, P. Scheltens, and P. M. Thompson. The clinical use of structural MRI in Alzheimer disease. *Nat Rev Neurol*, 6(2):67–77, Feb 2010.

[7] R. Haralick, K. Shanmugam, and I. Dinstein. Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics*, 3:610–621, 1973.

[8] C. R. Jack, M. A. Bernstein, N. C. Fox, P. Thompson, G. Alexander, D. Harvey, B. Borowski, P. J. Britson, J. L. Whitwell, C. Ward, A. M. Dale, J. P. Felmlee, J. L. Gunter, D. L. G. Hill, R. Killiany, N. Schuff, S. Fox-Bosetti, C. Lin, C. Studholme, C. S. DeCarli, G. Krueger, H. A. Ward, G. J. Metzger, K. T. Scott, R. Mallozzi, D. Blezek, J. Levy, J. P. Debbins, A. S. Fleisher, M. Albert, R. Green, G. Bartzokis, G. Glover, J. Mugler, and M. W. Weiner. The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. *J Magn Reson Imaging*, 27(4):685–691, Apr 2008.

[9] S. Klöppel, C. M. Stonnington, C. Chu, B. Draganski, R. I. Scahill, J. D. Rohrer, N. C. Fox, C. R. Jack, J. Ashburner, and R. S. J. Frackowiak. Automatic classification of MR scans in Alzheimer's disease. *Brain*, 131(Pt 3):681–689, Mar 2008.

[10] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and D. E. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[11] A. Wimo and M. Prince. World Alzheimer Report 2010 - The Global Economic Impact of Dementia. Technical report, Alzheimer's Disease International, 2010.

[12] J. Zhang, C. Yu, G. Jiang, W. Liu, and L. Tong. 3D texture analysis on MRI images of Alzheimer's disease. *Brain Imaging Behav*, 6(1):61–69, Mar 2012.