Alzheimer Disease Classification on Diffusion Weighted Imaging features

M. Termenón[1], A. Besga[2], J. Echeveste[3], A. Gonzalez-Pinto[2], M. Graña[1]

[1]Computational Intelligence Group
Basque Country University, UPV-EHU

[2]Unidad de Investigación en Psiquitaría
Hospital de Santiago Apostol, Vitoria-Gasteiz

[3]Departamento de Resonancia Magnética
Osatek, Vitoria-Gasteiz

Outline

Introduction

- The present paper will focus on the application of Machine Learning (ML) algorithms for the computer aided diagnosis (CAD) of Alzheimer Disease (AD).
- The aim of this paper is to obtain discriminant features from scalar measures of Diffusion Tensor Imaging (DTI) data and to train and test classifiers able to discriminate AD patients from controls on the basis of features selected from DTI volumes.

Alzheimer Disease

- AD is a neurodegenerative disorder, which is one of the most common cause of dementia in old people.
- This degenerative disorder presents a cognitive and behavioral impairment that interferes with the daily life of the individual and its social network, with a high economical and psychological cost.
- The diagnosis of AD can be done after the exclusion of other forms of dementia but a definitive diagnosis can only be made after a post-mortem study of brain tissue.
- This is one of the reasons why early diagnosis based on Magnetic Resonance Imaging (MRI) is a current research hot topic in the neurosciences.

Diffusion Weighted Imaging

- Diffusion Weighted Imaging (DWI) provides a measure of the integrity of the White Matter (WM) fibers measuring the movements of the water molecules inside the brain.
- This information can be used to provide structural information in vivo through the computation of diffusion tensors, the so called Diffusion Tensor Imaging (DTI).
- Scalar measures of diffusion computed from DTI are fractional anisotropy (FA) and mean diffusivity (MD), which give information about the magnitude of the diffusion process at each voxel, though they do not give direction.

Outline

Feature Database

- Thirty five men and women (aged 60-89), twenty controls and fifteen patients, were the subjects of this study.
- Patients also include two cases of very mild to mild AD.
- Structural MRI and DTI data were used for this experiment.

Image Processing

Algorithm 1: T1 and DWI data processing pipeline to obtain corrected FA and MD.

1. Convert DICOM to nifti
2. Skull stripping T1 volumes
3. Affine registration of T1 skull stripped volumes to template MNI152.
4. Correct DWI scans.
5. To obtain skull stripped brain masks for each DWI corrected scans.
6. To apply the diffusion tensor analysis.
7. Rigid registration 6DoF of FA and MD volumes to T1 affine registered volumes, resulting of Step3.

Feature Extraction process

- Procedure:
  - Considering each voxel site independently, we compose a vector with the FA or MD intensities at the voxel site across all the subjects.
  - We compute Pearson's correlation coefficient between this vector and the control variable, (Control=0; Patients=1) obtaining two independent volumes, one for FA and other for MD, of correlation values at each voxel.
  - We select a threshold corresponding to a percentile of the absolute correlation distribution, retaining the voxel sites with absolute value of correlation above this threshold.
  - For each percentile selected, we compose two feature vector for each subject, one extracted from the FA data and other from MD data.

Feature Extraction result

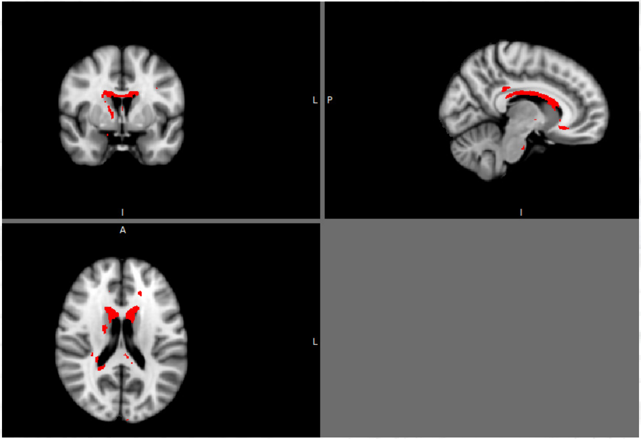

Figure: Voxel sites for FA features selected with a 99,5% percentile on the correlation distribution.

Feature Extraction result

- Voxel sites selected to build the feature vectors, were localized in many different regions of the brain.
  - FA: most significant differences were found in the thalamus, temporal lobe and corpus callosum. In white matter, we found discriminant voxel values in the cingulum gyrus, anterior thalamic radiation, corticoespinal tract and uncinate.
  - MD: there were also findings in the inferior fronto-occipital fasciculus.

Outline

Support Vector Machines

- Support Vector Machines (SVM) approach is a pattern recognition technique based on statistical learning theory.
- Its training principle consists of finding the optimal linear hyperplane that minimize the expected classification error.

$$y(x;w) = \sum_{i=1}^{N} w_i K(x, x_i) + w_0$$

Support Vector Machines

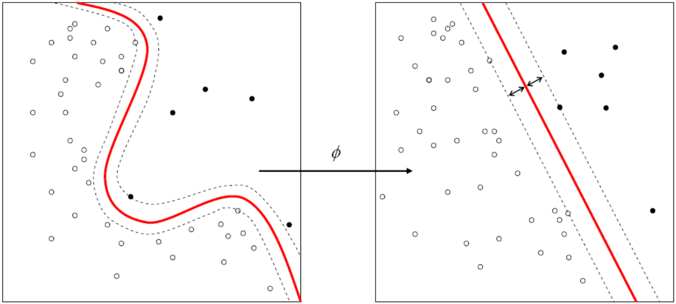

Figure: SVM linear separation.

Relevance Vector Machines

- Relevance Vector Machine (RVM) is a Bayesian sparse kernel technique for classification and regression.
- It is a sparse Bayesian model that provides probabilistic predictions through Bayesian inference.
- The benefit of a sparser classifier is that its results are more generalizable.

$$y(x;w) = \sum_{i=1}^{M} w_i \psi_i(x) = w^T \phi(x)$$

Relevance Vector Machines

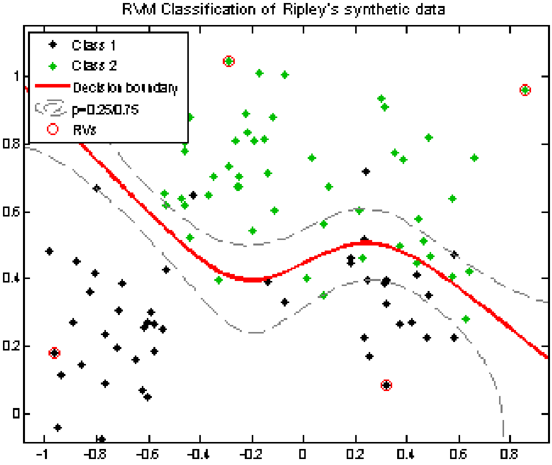

Figure: RVM classification.

Outline

Methodology

- To evaluate the performance of the classifier, we use 10-fold cross-validation, repeated 50 times.
- To quantify the results, we measured the Accuracy, Sensitivity and Specificity.
- We have labelled controls as class 0 and patients as class 1.

Approaches

- Tested approaches are:
  - 1NN.
  - RVM with linear kernels.
  - SVM with linear kernels.

Classification Results

	Pcr 99.50%		Pcr 99.90%		Pcr 99.95%		Pcr 99.99%	
1NN	**FA**	**MD**	**FA**	**MD**	**FA**	**MD**	**FA**	**MD**
Acc.	0.67	0.74	0.72	0.80	0.85	0.81	0.94	0.75
Sens.	0.68	0.78	0.69	0.81	0.86	0.81	0.87	0.75
Spec.	0.67	0.70	0.75	0.79	0.84	0.80	0.99	0.75
SVM	**FA**	**MD**	**FA**	**MD**	**FA**	**MD**	**FA**	**MD**
Acc.	0.96	0.99	0.96	0.99	0.97	0.98	0.99	0.94
Sens.	0.98	0.99	0.97	0.99	0.99	0.95	0.98	0.90
Spec.	0.95	0.98	0.95	0.99	0.96	0.99	0.99	0.97
RVM	**FA**	**MD**	**FA**	**MD**	**FA**	**MD**	**FA**	**MD**
Acc.	0.83	0.71	0.89	0.63	0.89	0.63	0.91	0.57
Sens.	0.78	0.67	0.88	0.65	0.88	0.67	0.89	0.67
Spec.	0.87	0.73	0.87	0.63	0.89	0.60	0.91	0.53

Table: Classification results for the FA and MD volumes.

Outline

Conclusions

- The aim of this paper was to test the hypothesis that features extracted from DTI images of Alzheimer patients and control subjects could be differentiated using classification techniques based on Machine Learning.
- Our main conclusion is that the proposed feature extraction is very effective providing a good discrimination between AD patients that can easily be exploited by the classifier construction algorithms.
- The selected voxels correspond to findings reported in the medical literature.
- The sensitivity and specificity results are well balanced, contrary to other classifiers that show some bias towards one of them.

Further work

- The main limitation of this study is that the results come from a small database. Therefore, more extensive testing will be needed to confirm our conclusions.

Thanks

- Contact:
  - Professor Manuel Graña.
  - Grupo de Inteligencia Computacional.
  - Universidad del País Vasco, San Sebsatián, Spain.
  - E-mail: **manuel.grana@ehu.es**
  - **http://www.ehu.es/computationalintelligence**