Classification Results of Artificial Neural Networks for Alzheimer's Disease Detection

Alexandre Savio[1], Maite García-Sebastián[1],
Carmen Hernández[1], Manuel Graña[1], Jorge Villanúa[2]

[1]Computational Intelligence Group, University of the Basque Country

[2]Osatek, Hospital Donostia

10th International Conference on Intelligent Data Engineering and Automated Learning
(IDEAL 2009)

Outline

1. Introduction
   - Alzheimer's Disease
   - Motivation
   - Introduction to the Analysis Methods
   - Materials and Methods
2. Results
3. Conclusions and Further Work

Outline

1. Introduction
   - Alzheimer's Disease
   - Motivation
   - Introduction to the Analysis Methods
   - Materials and Methods
2. Results
3. Conclusions and Further Work

Alzheimer's Disease (AD)

- Neurodegenerative disorder and one of the most common cause of dementia in old people.
- Still incurable and terminal.
- Although noninvasive approaches for antemortem diagnosis of AD are under development, definitive diagnosis requires a postmortem study of the brain tissue.
- T1 weighted MRI scans (sMRI) promises to aid diagnosis and treatment monitoring of AD, offering the potential for easily obtainable surrogate markers of diagnostic status and disease progression.

Outline

1 Introduction
- Alzheimer's Disease
- Motivation
- Introduction to the Analysis Methods
- Materials and Methods

2 Results

3 Conclusions and Further Work

Objective

Objective

- Detection of patients with very mild to mild Alzheimer's disease.

Our approach

Using sMRI and standard classifiers:

- Feature extraction based on Voxel-based Morphometry (VBM) analysis
- Backpropagation (BP)
- Radial Basis Function Networks (RBFN)
- Learning Vector Quantization Networks (LVQ)
- Probabilistic Neural Network (PNN)

Differential features of our work

- This issue has been addressed in many other works.

The differences here are:

- Freely available database with good quality images and well-documented.
- The number of subjects selected for this study is relatively high.

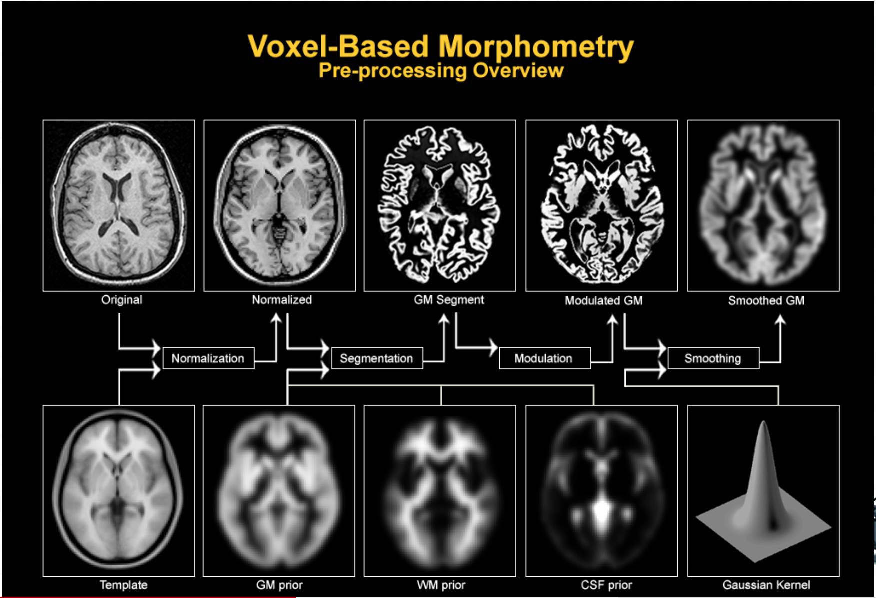
Outline

1. Introduction
   - Alzheimer's Disease
   - Motivation
   - Introduction to the Analysis Methods
   - Materials and Methods
2. Results
3. Conclusions and Further Work

Voxel-based Morphometry (VBM)

- Morphometry analyses allow a measurement of structural differences within or across groups throughout the entire brain.
- VBM measures differences in local concentrations of brain tissue, through a voxel-wise comparison of multiple brain images.
- The most popular brain morphometry analysis.

VBM Preprocessing Pipeline



VBM and the General Linear Model (GLM)

- After preprocessing we fit a linear statistical model to the data, each grey matter voxel independently.
- Use the estimated model parameter values to look for a specific effect we are interested in:
    - Identifying and characterizing structural differences in GM among populations.

VBM and GLM

- The GLM equation expresses the observed response variable in terms of a linear combination of regressors.

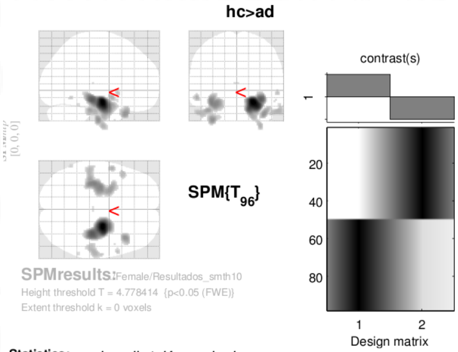
$$Y = X\beta + \varepsilon$$

- Y: observation vector (Mx1)
- X: design matrix (MxL). Each column corresponds to an effect that the user has built into the experiment or that may confound the results.
- β: regressor or covariate vector (Lx1). Unknown parameters
- ε: vector of error terms (Mx1)

VBM (Statistical Inference)

- On the results of GLM a t-test is computed at each voxel.
- The t-test values constitute a Statistical Parametric Map (SPM).
- The decision threshold for the test is set using Random Field Theory to account for spatial dependencies.

SPM Result



Statistics: ***p*–values adjusted for search volume**

set-level		cluster-level			voxel-level							
p	c	$p_{corrected}$	k_E	$p_{uncorrected}$	$p_{FWE-corr}$	$p_{FDR-corr}$	T	$(Z_{\equiv})$	$p_{uncorrected}$	mm	mm	mm
0.000	12	0.000	1764	0.000	0.000	0.000	7.59	6.70	0.000	24	−8	−16
					0.000	0.000	6.43	5.85	0.000	34	−24	−12
					0.007	0.000	5.33	4.98	0.000	34	−8	−48
		0.000	1355	0.000	0.001	0.000	5.91	5.44	0.000	−34	−20	−16
					0.001	0.000	5.87	5.41	0.000	−30	−12	−18
					0.002	0.000	5.71	5.29	0.000	−34	−14	−40
		0.000	161	0.004	0.002	0.000	5.68	5.26	0.000	40	24	−30
		0.000	195	0.002	0.006	0.000	5.36	5.00	0.000	58	−20	−28
					0.018	0.000	5.07	4.76	0.000	62	−10	−22
		0.005	42	0.106	0.007	0.000	5.33	4.98	0.000	−56	4	−8
		0.003	60	0.058	0.011	0.000	5.21	4.88	0.000	−48	20	−32
		0.018	13	0.358	0.015	0.000	5.12	4.80	0.000	58	12	−2
		0.024	8	0.475	0.030	0.000	4.93	4.64	0.000	−58	−54	−10
		0.034	3	0.679	0.038	0.000	4.86	4.58	0.000	56	10	−12
		0.027	6	0.541	0.041	0.000	4.84	4.57	0.000	62	−44	−18
		0.034	3	0.679	0.043	0.000	4.82	4.55	0.000	0	−22	10
		0.042	1	0.830	0.049	0.000	4.79	4.52	0.000	48	8	−40

Outline

1. Introduction
   - Alzheimer's Disease
   - Motivation
   - Introduction to the Analysis Methods
   - Materials and Methods
2. Results
3. Conclusions and Further Work

Subjects

- The set of subjects used consists in 98 women selected from the Open Access Series of Imaging Studies (OASIS) database

	Very mild to mild AD	Normal
No. of subjects	49	49
Age	78.08 (66-96)	77.77 (65-94)
Education	2.63 (1-5)	2.87 (1-5)
Socioeconomic status	2.94 (1-5)	2.88 (1-5)
CDR (0.5 / 1 / 2)	31 / 17 / 1	0
MMSE	24 (15-30)	28.96 (26-30)

- We find many subjects with high MMSE and low CDR.

Feature Extraction

- The clusters (regions) detected as result of VBM were used as a mask on the grey matter (GM) segmentation images to select the potentially most discriminant voxels.

Standard Classifiers

- Four supervised classification models were used:
  - Backward propagation of errors or Backpropagation (BP)
  - Radial Basis Function Networks (RBFN)
  - Probabilistic Neural Networks (PNN)
  - Learning Vector Quantization (LVQ)

Backward propagation of errors or Backpropagation (BP)

- A non-linear generalization of the squared error gradient descent learning rule for updating the weights of the artificial neurons in a single-layer perceptron.
- We have used the resilient backpropagation, which uses only the derivative sign to perform the weight updating.

Radial Basis Function Networks (RBFN)

- Are ANN that use radial basis functions as activation functions.
- RBFNs consist of a two layer neural network, where each hidden unit implements a radial activated function.
- Training consists of the unsupervised training of the hidden units followed by the supervised training of the output units' weights.

Probabilistic Neural Networks (PNN)

- A PNN is a special type of ANN that uses a kernel-based approximation to form an estimate of the probability density function of categories in a classification problem.
- The distance is computed from the point being evaluated to each of the other points and a RBF is applied to the distance to compute the weight for each point.
- The most common RBF function used is the Gaussian function, where a spread value must be set.
- We performed a search for the best sigma value in the range (0, 1).

Learning Vector Quantization (LVQ)

- LVQ provides a method for training competitive layers in a supervised manner.
- The system is composed of an unsupervisedly trained competitive layer which performs a partitioning of the input space.
- The supervisedly trained output layer provides the labeling of the input data according to its belonging to an input region (crisp clustering) or to its degree of membership (soft clustering).
- The basic versions proposed by Kohonen are known as the LVQ1 and LVQ2.

Features extracted

1. Mean and standard deviation of grey matter probability voxels within each cluster (MSD)
2. All grey matter voxels within clusters in a vector (VV)

Results

- All the results were extracted from the VBM detected clusters.
- We performed 10 times a 10-fold cross-validation for each experiment.
- For each experiment we show:
  - Size of feature vector
  - Classification accuracy
  - Sensitivity (AD patients correctly classified)
  - Specificity (controls correctly classified)

Backprop Results

Feature extracted	#Features	#Hidden units	%Accuracy	Sensitivity	Specificity
MSD	24	10	78.0 (0.12)	0.69 (0.14)	0.88 (0.13)
VV	3611	10	78.0 (0.11)	0.72 (0.17)	0.84 (0.18)

Table: Classification results with a BP network with resilient backpropagation. Mean (Standard deviation) of 10 cross-validations.

RBF Network Results

Feature extracted	#Features	Spread	%Accuracy	Sensitivity	Specificity
MSD	24	0.02	66.00 (0.13)	0.65 (0.24)	0.68 (0.14)
VV	3611	0.852	72.5 (0.10)	0.65 (0.21)	0.80 (0.17)

Table: Classification results with a RBF network. Mean (Standard deviation) of 10 cross-validations.

PNN Results

Feature extracted	#Features	Spread	%Accuracy	Sensitivity	Specificity
MSD	24	0.02	77.8 (0.09)	0.62 (0.14)	0.94 (0.1)
VV	3611	0.852	74.2 (0.14)	0.68 (0.20)	0.81 (0.17)

Table: Classification results with a PNN network. Mean (Standard deviation) of 10 cross-validations.

LVQ1 Results

Feature extracted	#Features	#Hidden units	%Accuracy	Sensitivity	Specificity
MSD	24	10	81.0 (0.18)	0.72 (0.27)	0.90 (0.14)
VV	3611	10	79.3 (0.13)	0.76 (0.23)	0.82 (0.19)

Table: Classification results with a LVQ1 network . Network training parameters:MSD: 200 epochs, goal: 0.01 and learning rate: 0.01; *VV*: 150 epochs, goal: 0.10 and learning rate: 0.010.Mean (Standard deviation) of 10 cross-validations.

LVQ2 Results

Feature extracted	#Features	#Hidden units	% Accuracy	Sensitivity	Specificity
MSD	24	10	83.0 (0.12)	0.74 (0.23)	0.92 (0.1)
VV	3611	10	77.0 (0.15)	0.76 (0.23)	0.78 (0.17)

Table: Classification results with a LVQ2 network . Network training parameters: *MSD*: 200 epochs, goal: 0.01 and learning rate: 0.01; *VV*: 50 epochs, goal: 0.01 and learning rate: 0.005. Mean (Standard deviation) of 10 crossvalidations.

Conclusions

- We performed feature extraction processes based on VBM analysis to classify MRI volumes of AD patients and normal subjects.
- We used the basic GLM design without any covariate to detect subtle changes between AD patients and controls. The best accuracy result is 83% with the LVQ2, but this result is not far from the results of LVQ1 and PNN.
- As we don't have post-mortem confirmation of AD subjects, the very mild demented subjects could be false positives. Post-mortem confirmation data of AD diagnosed subjects could improve the results.

Further work

- Using other morphometry methods such as Deformation-based andTensor-based morphometry.
- Try these methods with real clinical subjects and different types of dementia like MD1 and FTD.

Questions?

Thank you for your attention.