

---

## AUTOMATIC CLASSIFICATION OF STRUCTURAL MRI FOR DIAGNOSIS OF NEURODEGENERATIVE DISEASES

### Clasificación automática de IRM estructural para el diagnóstico de enfermedades neurodegenerativas

GLORIA DÍAZ<sup>1</sup>, Ph. D. (c); EDUARDO ROMERO<sup>1</sup>, Ph. D.;  
JUAN ANTONIO HERNÁNDEZ-TAMAMES<sup>2</sup>, Ph. D.; VICENTE  
MOLINA<sup>3</sup>, M.D.; NORBERTO MALPICA<sup>2</sup>, Ph. D.

<sup>1</sup>Grupo de Investigación Bioingenium, Universidad Nacional  
de Colombia. Bogotá D.C., Colombia.

<sup>2</sup>Laboratorio de Análisis de Imagen Médica y Biometría,  
Universidad Rey Juan Carlos. Madrid, España.

<sup>3</sup>Hospital Clínico de Salamanca. Salamanca, España.

Correspondencia: Eduardo Romero Castro. Universidad Nacional  
de Colombia. Facultad de Medicina. Carrera 30 # 45-03, edificio 471 -  
Medicina, piso 1. Bogotá D.C., Colombia. edromero@unal.edu.co

Presentado 15 de marzo de 2010, aceptado 4 de junio de 2010, correcciones 23 de junio de 2010.

#### ABSTRACT

This paper presents an automatic approach which classifies structural Magnetic Resonance images into pathological or healthy controls. A classification model was trained to find the boundaries that allow to separate the study groups. The method uses the deformation values from a set of regions, automatically identified as relevant, in a process that selects the statistically significant regions of a t-test under the restriction that this significance must be spatially coherent within a neighborhood of 5 voxels. The proposed method was assessed to distinguish healthy controls from schizophrenia patients. Classification results showed accuracy between 74% and 89%, depending on the stage of the disease and number of training samples.

**Key words:** Neurodegenerative disease, structural MRI, pattern classification, SPM, VBM, DARTEL, support vector machines.

#### RESUMEN

Este artículo presenta un método automático para la clasificación de individuos en grupos patológicos o controles sanos haciendo uso de imágenes de resonancia magnética. El método propuesto usa los valores de deformación del sujeto analizado a un cerebro plantilla, para entrenar un modelo de clasificación capaz de identificar las fronteras que separan los grupos de estudio en un espacio de características dado. Con el fin de reducir la dimensionalidad del problema, un conjunto de regiones relevantes es automáticamente extraído en un proceso que selecciona las regiones estadística-

mente significativas en una prueba *t-student*, con la restricción de mantener coherencia en dicha significancia en una vecindad de 5 voxeles. El método propuesto fue evaluado en la clasificación de pacientes con esquizofrenia y sujetos sanos. Los resultados mostraron un desempeño entre el 74 y el 89%, el cual depende principalmente del número de muestras empleadas para el entrenamiento del modelo.

**Palabras clave:** enfermedades neurodegenerativas, resonancia magnética estructural, reconocimiento de patrones, SPM, VBM, DARTEL, máquinas de vectores de soporte (SVM).

## INTRODUCTION

Neurodegenerative diseases are characterized by a progressive loss of neurons of the central nervous system, or their synaptic function. It is estimated that more than 24 million elderly adults are affected by these disorders at the world, and this number will increment with the development of aging population. Schizophrenia is a particular case of neurological disease in which there is no neuronal death, but neuron connection loss, so that it can be considered as a neurodegenerative disease (Marengo and Weinberger, 2000).

Currently, there are no approaches or biomarkers that allow to make a definitive diagnosis of most of these diseases. Possible pathological cases are identified based on symptoms and signs that are observed into thorough interviews of the patient and relatives, neuropsychological tests and clinical examinations that rule out other diseases (Chaves *et al.*, 2009). These strategies are not capable of identifying the early stage of the diseases, because they are applied when the patient presents symptoms that are observable after the neuron death has months or years of evolution. Therefore, the development of tools for diagnosis of early stages of neurodegenerative diseases is an important research topic.

Recent studies suggest that morphological changes, produced by a number of neurodegenerative diseases, could be used as early diagnosis information (Mechelli *et al.*, 2005). But, the enormous inter and intra-subject variability and noise leads to a non trivial establishment of the membership group for unclassified individuals. Figure 1. shows images of different healthy and pathological subjects. Although all of them corresponding to the same slice of the MRI volume, there is not easy to identify visible differences associated to pathology. Moreover, differences intra groups can be more visible than differences between groups.

Voxel-based morphometry (VBM) is an automatic objective whole brain analysis method that allows to identify inter-subject brain differences. This method provides a statistical estimation of inter-group brain density or volume differences using a voxel-by-voxel basis in a standardized space. Overall, this method computes statistical parametric maps (SPM) for localizing significant differences between two or more experimental groups using a general linear model (GLM; Ashburner, 2009).

On the other hand, Machine learning classifiers had shown that they can learn complex patterns and trends from sample data for creating accurate decision surfaces. However, a suitable selection of the features used for describing each instance is necessary for

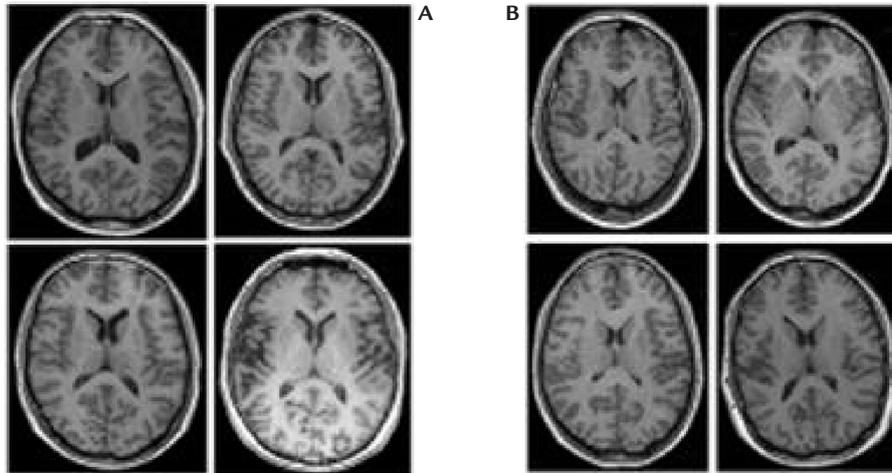


Figure 1. MRI images of healthy and pathological subjects. A. Healthy subjects. B. Pathological subjects with different schizophrenia grades.

obtaining good performance. The aim of this work was to evaluate a machine learning approach for classifying T1-weighted MRI images of pathological patients and normal controls, using VBM to identify relevant brain regions that are then used as region of interest in the feature extraction process. Mean, variance and standard deviation of those regions are used as features for the classification stage. A SVM-RFE approach is used for selecting the most discriminative features that are finally used for training a SVM learning model, which classifies individual subjects.

#### RELATED WORK

At the last years, advances on machine learning and pattern recognition methods have been used for predicting the diagnosis of individuals from morphological analysis of structural and functional MRI (Pereira *et al.*, 2009). From a pattern recognition point of view, the main challenge is to identify signatures of disease in the structural images, named feature vectors, which allow to discriminate pathological from healthy patients. Machine learning algorithms have been used to discriminate pathological patients from controls based on Positron Emission Tomography (PET) or Single-Photon Emission Tomography (SPECT) functional volumes (Johnson *et al.*, 1998; Stoeckel and Fung, 2005; Ramirez *et al.*, 2009; Horn *et al.*, 2009), perfusion fMRI scans (Wang *et al.*, 2007), fMRI brain activation maps (Ford *et al.*, 2003, Demirci *et al.*, 2008; Tripoliti *et al.*, 2008), Diffusion Tensor Images (Kloppel *et al.*, 2007; Caprihan *et al.*, 2008; Hua *et al.*, 2009; Robinson *et al.*, 2010), among others. Moreover, advances in medical imaging acquisition are able to generate high-resolution volumetric MRI, which improve the morphological features that are captured. This fact, offer opportunities to develop computer aided diagnosis tools based on this information.

Most of the approaches proposed as classification system for the diagnosis of neurodegenerative diseases based on structural MRI fall on one of two categories: 1. Classification based on shape descriptors of specific anatomical regions, known to be

affected by the disease, which are segmented manually (Wang *et al.*, 2009; Hua *et al.*, 2008; Chetelat and Baron, 2003, Ford *et al.*, 2002) or automatically (Gerardin *et al.*, 2009; Xu *et al.*, 2001; Lao *et al.*, 2004; Li *et al.*, 2007; Shen *et al.*, 2005; McEvoy *et al.*, 2009; Magnin *et al.*, 2009) from the original brain volume. 2. Classification based on unbiased analysis of the whole brain that requires no prior assumptions to be made (Fan *et al.*, 2007; Chu *et al.*, 2009; Costafreda *et al.*, 2009; Fan and Shen, 2009; Misra *et al.*, 2009). In the former case, results are limited by the accuracy of the segmented structure or the reliability of a priori knowledge used for defining the affected brain areas, while in the latter case statistics inferred from observed data is used as unique knowledge about the classification problem and no prior regions are necessarily considered. The framework proposed in this paper falls in the second category, so a more extensive review of former methods is beyond the scope of this paper.

The first step in a whole brain based classification approach is to infer the voxels in which morphological structures differ between groups, in order to include only disease discriminative information to train the learning model. This is commonly achieved by a voxel-wise statistical analysis of structural MRI images. Moreover, because the morphological changes of brain structures resulting from pathological processes usually do not occur in isolated regions (Fan *et al.*, 2007), most of the proposed approaches cluster the discriminative voxels in irregular regions which are used to characterize the whole brain changes.

Fan *et al.*, (2007) proposed the COMPARE (Classification of Morphological Patterns Using Adaptive Regional Elements) method for classification of structural brain magnetic resonance images by combining deformation-based morphometry with support vector machines (SVM). Density maps, extracted for each individual using the RAVENS approach (Davatzikos *et al.*, 2001), are segmented by a watershed algorithm according to the similarity of discriminative measure of near voxels. This measure results from the combination of the Pearson correlation measure between voxel tissue density value and classification labels, and their spatial consistency. A volume increment algorithm, similar to a region growing strategy, is applied to grouping voxels that show similar relationships to the discrimination measure. Finally, a support vector machine model is trained to classify between schizophrenia patients and healthy controls using the mean tissue density value of all selected regions as feature vector. Although the reported results are optimistic (accuracy upper to 90%), the dataset used to evaluate the algorithms does not include patients at early stages of the disease. This approach has been widely used as diagnostic tool in other neurodegenerative disease (Davatzikos *et al.*, 2008; Davatzikos *et al.*, 2008a, Fan *et al.*, 2008; Misra *et al.*, 2009). However, its main drawbacks are the high computational cost of the region extraction process and the use of proprietary methods for generating the tissue density maps.

As was mentioned in the introduction, Statistical Parametric Mapping (SPM) and Voxel Based Morphometry (VBM) are increasingly used and actually can be considered as standard tools in the study of neurodegeneration. So, some approaches for pathologic classification of brain volumes using SPM and SVM have been lately reported (Costafreda *et al.*, 2009). Costafreda *et al.*, (2009), propose using whole-brain ANOVA filtering of tissue density maps computed using SPM to select the areas of maximum group differences between depressed patients and healthy controls, which were then

used for predicting both the diagnostic classification and the clinical response to antidepressant medication. A two-sample t-test was applied on grey matter density maps for extracting regions which were used for classify Alzheimer disease patients (Zhang *et al.*, 2008). Similarly, Savio *et al.*, (2009) use voxel clusters detected by applying VBM to detect Alzheimer's disease on MRI. In this case three descriptors were evaluated: grey matter proportion, mean and standard deviation of VBM detected clusters and all the grey matter segmentation values for voxels belonging to it. Best performance was obtained with statistical measures of detected regions.

## MATERIALS AND METHODS

A statistical machine learning classifier model is a function that predicts the class of an unclassified instance based on the information provided by features. A model is constructed based on a set of training instances that belong to a known class. Figure 2 illustrates the main stages of the proposed approach. In an off-line process healthy and pathological images are processed in order to automatically extract relevant regions which were significantly related to the particular disease. Then, a feature extraction and selection process is applied for generating descriptors of these regions, which are then used for training a learning model able to separate the feature space into the two classes. Finally, when a new subject's brain volume is presented, the relevant regions are located and characterized with the same descriptors used in the training stage, and the trained model is used to classify the individual as healthy or pathological.

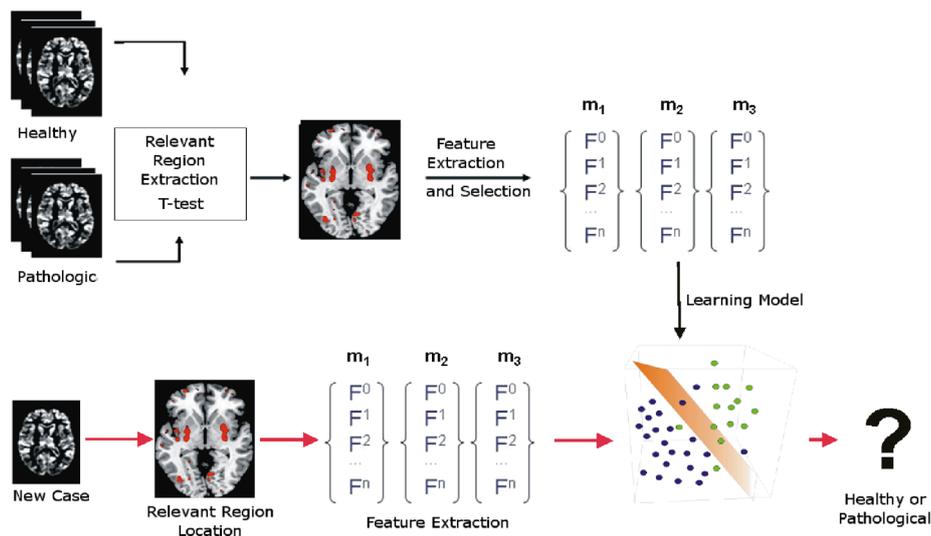


Figure 2. Overview of proposed approach.

### IMAGE PREPROCESSING

From a morphometrical analysis standpoint, diseases should have measurable manifestations, that allow to establish hyper/atrophies related to the disease under

study by comparing the different three-dimensional T1 weighted MRI among different subjects. However, acquired studies can not be directly compared voxel-by-voxel because of large anatomical variability which results in unbearable amounts of noise. Therefore, the main goal of the deformation stage is to provide a standard spatial space in which brain structures can be compared.

Standard neurological image analysis methods provided by the SPM8 software<sup>1</sup> were used for processing images (Ashburner, 2009). First, a manual alignment to the AC/PC line was performed for achieving proper correction of the brain orientation differences. Then the cerebral hemispheres were segmented into grey (GM) and white matter (WM) tissues, using the SPM8 segmentation approach (Ashburner, 2009). Segmented tissue maps were then warped to a template space via an image registration approach known as DARTEL (Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra; Ashburner *et al.*, 2007). In order to obtain deformation measures associated principally to diseases, the brain template should be very approximate to all healthy brains in the study. For this reason, the template used in this study was computed from the all control samples of the same study using also the DARTEL algorithm. The gray level of the spatially normalized images was rescaled in order to maintain the same quantity of tissue as in the original images, a process known as normalization. Finally, deformed tissue maps were smoothed applying an 8 mm FWHM Gaussian kernel for reducing differences between brain anatomies. So, in the resulting pre-processed images, the intensity value of every voxel encoded the deformation needed in that voxel to fit the template.

#### RELEVANT REGION LOCATION

This phase attempted to locate areas with significant morphological differences, between the groups that we want to classify. In the pre-processing stage all images were aligned, so that a tissue density comparison at each voxel between groups could be carried out. The general linear model (GLM) was used to identify regions of grey and white matter in which each tissue density was significantly different for the two groups (control and pathological) in the training data set. A student t-test, which basically established significant inter-group mean differences, was used to compute a statistical map that was thresholded for selecting voxels which met these two conditions: their value was statistically significant and they could be grouped together within regions with more than five neighbor voxels with similar statistical properties. Therefore a set of regions scattered through the two brain tissue maps, grey and white matter, were identified as regions of interest and used in the classification task.

#### FEATURE EXTRACTION

The aim of feature extraction step was to obtain a set of descriptors, and will be further separated in different classes by the learning model. The feature extraction can be formally defined as a function  $f$ , which maps the original tissue density maps onto a

---

<sup>1</sup><http://www.fil.ion.ucl.ac.uk/spm/>

feature vector  $x$ , i.e.  $f: I \rightarrow \mathbf{x} = (x_1, x_2, \dots, x_d)$ , where  $d$  is the number of features used to characterize the image.

Our main hypothesis is that the classification of brain MRI studies is possible by evaluation of the tissue density values on selected region of interests. This feature extraction step computes measures which allow an objective description of the tissue density variation. One efficient way to describe these variations is to model each region as a random variable, described by its corresponding probability density function and compute the first order moments for each. In this work we evaluate the effect of using uniquely the first order moment and the concatenation of the first two and first three order moments as region descriptors. Order moments were computed as described by Equation (1).

$$M_n(R_i) = \frac{1}{N(R_i)} \sum_{x \in R_i} (x - \mu(R_i))^n \quad (1)$$

where  $N(R_i)$  is the number of voxels belonging to region  $R_i$  and  $\mu(R_i)$  is the mean tissue density value in the region  $R_i$ .

#### FEATURE SELECTION

The problem of the high number of dimensions was partially solved with the relevant region extraction process. However, the remaining feature vector was also large because it was composed by statistical measures of all located regions for both density tissues: grey and white matter. So, a feature selection process was carried out in order to reduce the final feature vector as much as possible.

Feature selection is a process commonly used in pattern recognition, which allows determining the most relevant, discriminating and uncorrelated descriptors for a classification task, while reducing the dimensionality of vectors associated to instances. Recursive Feature Elimination (RFE) is a simple and effective method for reducing the feature vector size, which iteratively removes the descriptors with the smallest ranking criterion, whilst optimize the classification rate for a specific learning model. In this work the SVM-RFE algorithm was used, which found the features which reported the largest margin of class separation, using the square weight values of the support vectors as ranking criteria (Guyon *et al.*, 2004).

#### CLASSIFICATION

Based on the regional features selected in the previous stage, a support vector machine (SVM) learning model was trained for classifying individual MRI images. SVM is a supervised binary classification algorithm that has received increasing interest because it has outperformed other methods in several pattern recognition tasks. Intuitively, SVM produces an optimal separating hyperplane between two classes in a feature space, in which each training instance has been represented. These hyperplane corresponds to the largest separation, or margin, between the two classes. When a new instance should be classified, original feature vector is mapped to the same space and the distance to SVM hyperplane allows to decide if the new instance falls into one category or the other.

In general, the original SVM algorithm proposed by Vladimir Vapnik (Platt, 1999) was a linear discriminator, but subsequent modifications suggested that non-linear

discrimination tasks can be achieved thanks to the “kernel trick”. This is, mapping the original non-linear observations into a higher-dimensional space in which boundaries among classes become linear. In this work, the Gaussian radial basis function kernel, defined as Equation (2) was used (Platt,1999).

$$K(x, y) = e^{-\gamma \|x - y\|^2} \quad (2)$$

where  $x$  and  $y$  are two feature vectors, and  $\gamma$  controls the size of the Gaussian kernel.

## RESULTS AND DISCUSSION

### TEST SUBJECTS

The performance of the classification approach was evaluated on one dataset of T1-weighted MRI images, which included schizophrenia patients (SP), patients with a first-time seizure that developed the pathology (FTS), patients with a first-time seizure that did not develop the pathology (FT), and healthy controls (HC). Previous works have shown that there exist important differences between brains of women and men, which can introduce more noise to the classification problem. For these reasons, our dataset was divided into female and male subjects for conforming two datasets. A total of 221 brain volumes were used for evaluating the proposed approach. Distribution of this dataset corresponds to the showed in the Table 1.

	Male		Female	
	n	Age	n	Age
Healthy Controls	26	30/17-45	29	30/19-51
SP patients	60	36/20-63	30	38/19-69
Patients with a first-time seizure without posterior pathology	39	27/18-43	18	25/19-63
Patients with a first-time seizure with posterior pathology	9	33/22-49	10	27/20-34

Table 1 Data set distribution, with n: number of samples per class.

The patients, whose images were analyzed, participated in a longitudinal study of first-break schizophrenia. Patients were include in the protocol after having a first psychotic episode with symptoms for longer than 1 week, not attributable to organic or toxic causes, and not related to other axis I disorders. The absence of marked stressors, according to the Diagnostic and Statistical Manual (A.P.M,2000), clearly related to the episode was also required for inclusion. These criteria were aimed at avoiding the inclusion of patients with transient psychotic symptoms. The imaging studies took place within 2 weeks following enrolment. The FE was defined as the first contact with psychiatric services. Information from patients and their families was used to estimate the length of the episode and thus the onset of symptoms. After inclusion, all patients were followed up, in monthly visits, for 2 years on an outpatient basis to confirm or rule out a diagnosis of schizophrenia at the end of this period. After this follow-up, two experienced psychiatrists, who were blind to the results of the MRI scans, diagnosed each patient and decided whether the index episode was a first break of schizophrenia

or, instead, a single psychotic episode that had not evolved into schizophrenia 2 years after the enrolment. For this final diagnosis they used all the available follow-up information (including the total duration of significant symptoms), the data from a semi-structured interview (SCID, Clinical Version), and the information provided by the families and clinical personnel. For diagnostic purposes, the duration of symptoms was calculated covering the period before and after initiation of treatment.

#### MODEL EVALUATION

SVM learning models were trained through an exhaustive search of their learning parameters. The regularization parameter  $C$  was varied from 1 to 10 with increment steps of 1, while the parameter  $\gamma$  that defines the nonlinear mapping from input space to some high-dimensional feature space, was varied from 0.01 and 1 with increment steps of 0.02. Finally, the number of features selected via SVM-RFE, was varied from 10 to 100 with a step of 10.

Parameter evaluation was carried out through a conventional 10-fold cross validation approach, in which the original sample set is randomly partitioned into 10 subsamples. Nine of these subsamples are used as training data whilst the remaining subsample is used as the validation data. This process is repeated 10 times, so that each subsample is used exactly once as the validation data. On the other hand, performance of the classification tasks was quantified in terms of its predictive precision, sensitivity and specificity. Effectiveness, computed as  $F\beta$ -measure, was used to combine sensitivity and precision measures. Performance measures were computed as shown in equation (3).

$$\begin{aligned} \gamma = 0.1 \\ \text{Precision}(PR) &= \frac{TP}{TP + FP} \\ \text{Recall / Sensitivity}(RC) &= \frac{TP}{TP + FN} \\ \text{Specificity} &= \frac{TN}{TN + FP} \\ F\beta &= \frac{PR * RC}{\beta * RC + (1 - \beta) * PR} \end{aligned} \quad (3)$$

where TP stands for the true positives, FN for the false negatives, and FP for the false positives. The  $\beta$  coefficient ( $0 < \beta < 1$ ) allows to assign relative weights to both the precision and recall measures (Daskalaki, 2006). In this work,  $F\beta$  was set at 0.4, so the search was addressed to detection of TP.

#### REGIONS DETECTED AND SELECTED IN THE PROCESS

Figures 4 and 5 show the regions included in the classification process. Regions considered as significant by the SPM analysis appear in red color, and the regions among those that were considered significant for classification appear in pink color. Among the selected regions were the thalamus, the caudate nucleus, the basis pedunculi, the parahippocampal gyrus and the superior frontal gyrus. In the case of female subjects (shown on the right of figures 4 and 5) the superior vertebral peduncles and the amigdala are also considered as important features for classification.

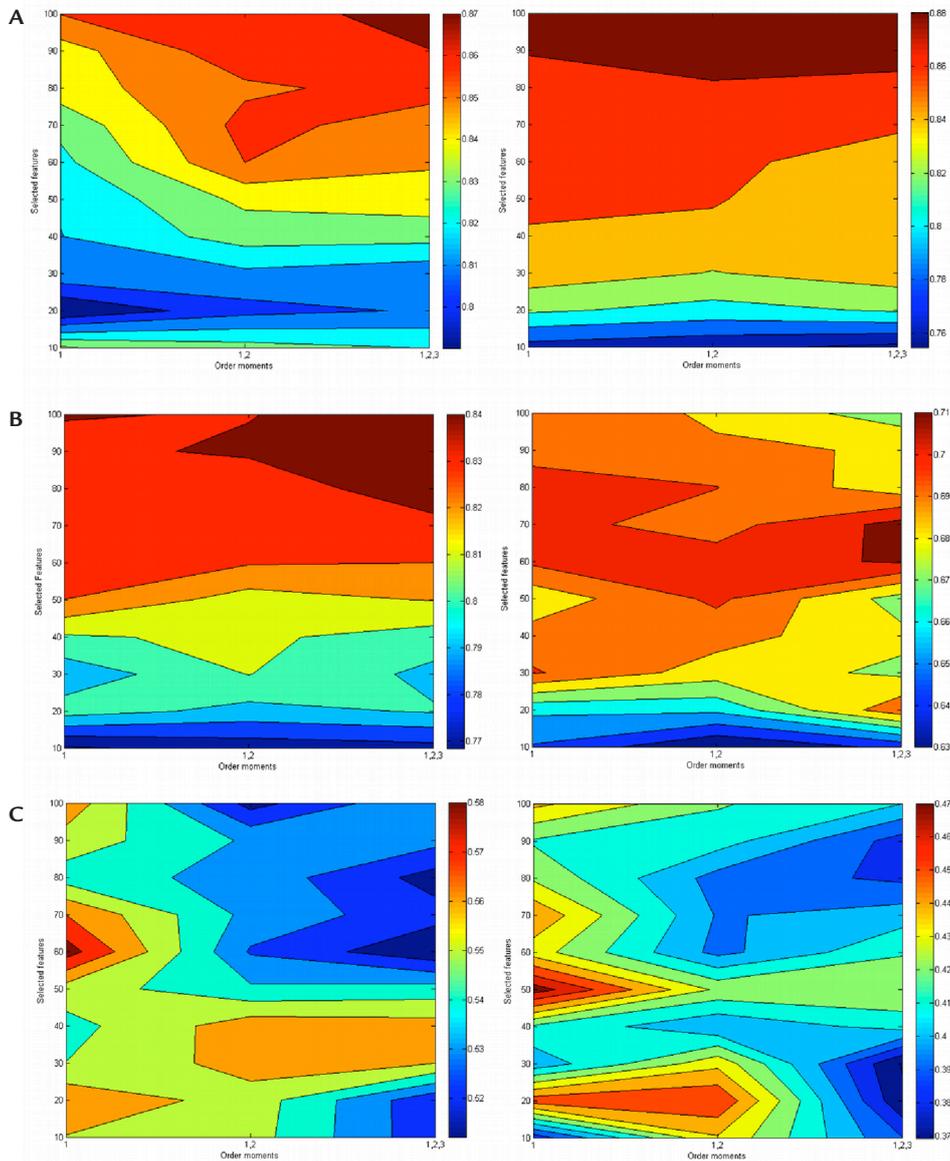


Figure 3. Contour plots of the  $F\beta$  measure for male (left) and female (right) classification between healthy controls and schizophrenia patients (row A), healthy controls and patients with a first-time seizure that did not develop the pathology (row B), and healthy controls and patients with a first-time seizure that developed the pathology (row C). x-axis corresponds to the order moments used as feature vector. The y-axis corresponds to number of selected feature.

### CLASSIFICATION PERFORMANCE

Six SVM learning models were created for classifying the different combinations of binary healthy-vs-pathological problems i.e HC-vs-SP, HC-vs-FTS and HC-vs-FT for both male

and female datasets. Learning model effectiveness was assessed using the  $F\beta$  measure varying the learning parameters as was described in section 4.3. A comprehensive analysis of all parameter combination was assessed using many contour plots as shown in Figure 3. We found that the number of selected feature was the parameter that has larger effect on the performance. On the other hand complexity ( $C$ ) and  $\gamma$  parameters of the SVM learning model did not have relevant effect. Figure 3 shows the contour plots of  $F\beta$  measure for the six learning models, using a SVM with  $C$  and  $\gamma$  fixed and varying number of selected features and order moments used as descriptor. Each  $F\beta$  value corresponds to the average of 10 experiments from 10-fold cross validation process i.e. a total of 100 experiments were carried out for each classification task. Dark red regions correspond to the values of parameters that report best performance. Note that, using the third first order moments as region feature descriptor improve the classification performance, with respect to use of mean value used in previous work. Interestingly, classification of healthy controls-vs-patients with a first-time seizure that developed the pathology was improved when selected features was around of 50, which mean that discriminative information is only in a few of specific regions.

Results show that the proposed approach achieves different performances on each subset of samples. That can be explained because it is expected that the brain structural differences between healthy controls and schizophrenia patients are be greater than with the other patient sets, and therefore selected features should be more discriminative. However, the unbalanced number of training samples for each dataset might be an important issue that affects the performance of the method.

In the HC-vs-SP classification task the effectiveness was always above 0.76 in both cases male and female datasets. The best performance is reported when great values of selected features and three order moments are used. In the case of males, a better performance was reached for complexity and  $\gamma$  parameters set to 1, and order moments and selected features set to 3 and 100, respectively, which presented an average  $F\beta$  of 0.86. In the female case better performance was reported by the SVM with complexity 1,  $\gamma = 0.1$  and order moments and selected features set at 3 and 90, which presented a  $F\beta$  measure of 0.89.

In the HC-vs-FT classification task the effectiveness was lower than HC-vs-SP. Besides major difference was present between male and female classification performance, which shows the relevance of number of training samples to improve the results. On the other hand, in this classification task, the Gaussian kernel size had a major effect on the overall performance. For both male and female best performances were reported for  $\gamma$  lower than 0.06., this mean that the instances of the same group are most scattered in the feature space. Likewise than HC-vs-SP classification task, best effectiveness was achieved when selecting higher number of features and order moments. In the male case better performance was reported for complexity = 1,  $\gamma = 0.06$ , selected features = 80 and order moments = 3, which presented an average  $F\beta$  of 0.85. In the female case a  $F\beta$  measure of 0.72 was achieved with complexity,  $\gamma$ , selected features and order moments set at 1, 0.01, 70 and 3, respectively.

Finally, the HC-vs-FTS classification task reports the worse performance. Effectiveness does not exceed the 0.59 value for male and 0.47 for females. So, the results can't be considered reliable due that the small number of training images in two datasets.

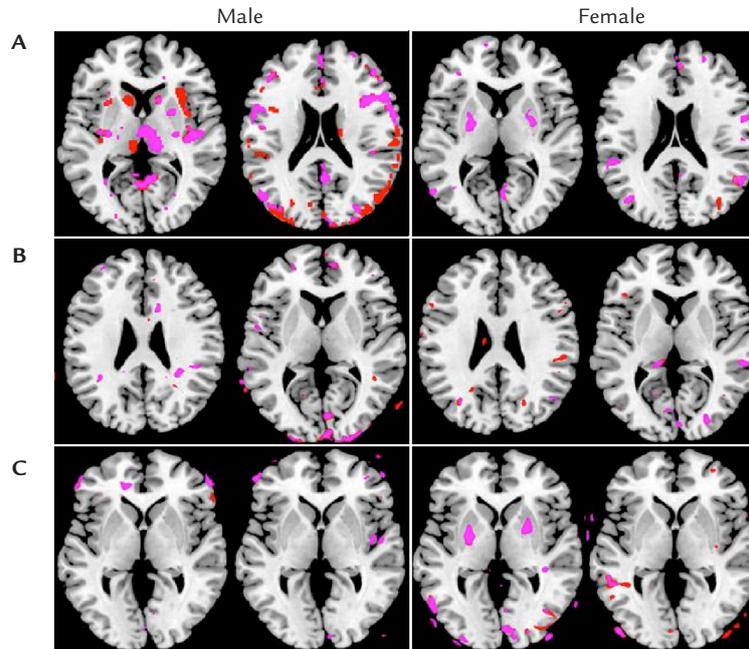


Figure 4 Local regions extracted and selected as input of classification process. Pink regions correspond to remaining regions after to feature selection stage, which were more discriminative for classification task. Red regions correspond to regions detected by SPM analysis that were removed after selection stage. In A row, regions found for HC-vs-SP classification, in B row, regions found for HC-vs-FTS classification and in C row, regions found for HC-vs-FT classification.

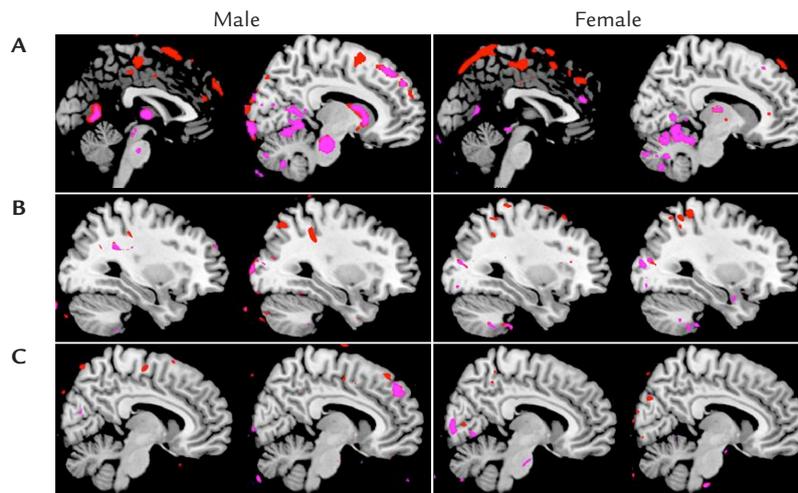


Figure 5. Local regions extracted. Pink correspond to remaining regions after to feature selection stage, which were more discriminative for classification task. Red regions correspond to regions detected by SPM analysis that were removed after selection stage. In A row, regions found for HC-vs-SP classification, in B row, regions found for HC-vs-FTS classification and in C row, regions found for HC-vs-FT classification.

Table 2 presents the overall performance of the proposed approach for each classification task, using the selected parameters. Note that the precision measure of all cases is relative worse (between 47.9% and 89.6%). This can be explained by the fact that the parameters selected considered preference by the correct classification of the pathological samples, by which miss classification of healthy controls are most accepted than if selection was based on overall accuracy or mean error. Recall measure, also named sensitivity, reports how the proposed approach is able to identify pathological patients, which in this case is high for HC-vs-SP and HC-vs-FT tasks. On the other hand, specificity, the approach capacity to distinguish healthy controls, was between 71.8% and 89.4%.

	Male			Female		
	PR	RC	SP	PR	RC	SP
C-vs-SP	80.4	92.4	75.2	89.6	88.9	89.4
C-vs-FT	91	81	87.8	62.2	79	71.8
C-vs-FTS	59.5	58.8	84.2	47.9	47.3	83.2

Table 2 Average Accuracy (ACC), Precision (PR), Recall (RC) and specificity (SP) measures of proposed approach achieved by the learning models with the selected parameters for each classification task.

## CONCLUSIONS

In this paper a pattern classification approach for predicting individual categories according to morphological brain differences was studied. This approach looks for relevant regions from T1 brain images, which are identified as discriminative to the classification task without a priori information about regions affected by specific disease. The approach has been tested on a study on schizophrenia, providing promising results on the classification between controls and pathological brains, which should be used as baseline to recognition of patients with other neurodegenerative diseases.

In this study we demonstrate that it is feasible to make use of standard tools as SPM and VBM, that are frequently applied in comparisons between pathological and healthy groups, for extracting morphological regions that come be relevant in classification tasks. However, an exhaustive learning a feature selection must be carried out in order to rise the expected results.

A conventional non linear SVM learning algorithm was built on a set of selected features that described the morphological brain deformations respect to a brain template. From the results we can observe that classification accuracy is highly dependent of sample size, which is a critical problem in neurodegenerative disease research, as it is not easy to find a highly number of patients for each pathological stage. So, initiatives as multicentre dataset for research development, as ADNI<sup>2</sup>, must be promoted. On the other hand, an important issue found in this study was the improvement achieved when statistical measures of deformation degree are used to describe the brain variations instead of using only the mean deformation as proposed in previous works.

<sup>2</sup><http://www.loni.ucla.edu/ADNI/>

In the future, we plan to deal with the multiclass classification problem, which requires more sophisticated feature extraction process and machine learning approaches to be investigated. Moreover, we plan to integrate or combine multimodal information such as fMRI, sMRI, signals and clinical data in order to look for other patterns that provide higher discrimination power for assisting in the early diagnosis of neurodegenerative diseases in which morphometrical changes may not be enough.

## BIBLIOGRAPHY

- AMERICAN PSYCHIATRIC ASSOCIATION. Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition. Text Revision. Washington D.C., American Psychiatric Association; 2000.
- ASHBURNER J. A fast diffeomorphic image registration algorithm. *NeuroImage*. 2007;38(1):95-113
- ASHBURNER J. Computational anatomy with the SPM software. *Magnetic Resonance Imaging*. 2009;27:1163-1174.
- CAPRIHAN A, PEARLSON GD, CALHOUN VD. Application of Principal Component Analysis to Distinguish Patients with Schizophrenia from Healthy Controls Based on Fractional Anisotropy Measurements. *NeuroImage*. 2008;42:675-682.
- CHAVES R, RAMIREZ J, GÓRRIZ JM, LÓPEZ M, SALAS-GONZALEZ D, ÁLVAREZ I, *et al*. SVM-based computer-aided diagnosis of the Alzheimers disease using t-test NMSE feature selection with feature correlation weighting. *Neuroscience Letters*. 2009;461:293-297.
- CHETELAT G, BARON JC. Early diagnosis of Alzheimers disease: contribution of structural neuroimaging. *NeuroImage*. 2003;18:525-541.
- CHU C, Novak M, Kloppel S, Ashburner J. Gaussian Processes Classification of Huntington's Disease with automatic feature selection. *NeuroImage*. 2009;47 S1:S39-S41.
- COSTAFREDA SG, CHU C, ASHBURNER J, FU CHY. Prognostic and Diagnostic Potential of the Structural Neuroanatomy of Depression. *PLoS ONE*. 2009;4(7):e6353. doi:10.1371/journal.pone.0006353
- DAVATZIKOS C, GENC A, XU D, RESNICK SM. Voxel-Based Morphometry Using the RAVENS Maps: Methods and Validation Using Simulated Longitudinal Atrophy. *NeuroImage*. 2001;14:1361-1369.
- DAVATZIKOS C, FAN Y, WU X, SHEN D, Resnick SM. Detection of prodromal Alzheimers disease via pattern classification of magnetic resonance imaging. *Neurobiol Aging*. 2008;29(4):514-523.
- DAVATZIKOS C, RESNICK SM, WU X, PARMPI P, CLARK CM. Individual patient diagnosis of AD and FTD via high-dimensional pattern classification of MRI. *Neuroimage*. 2008;41(4):1220-7.
- DEMIRCI O, CLARK VP, CALHOUN VD. A projection pursuit algorithm to classify individuals using fMRI data: Application to schizophrenia. *NeuroImage*. 2008;39(4):1774-1782.
- FAN Y, SHEN D, GUR RC, GUR RE, DAVATZIKOS C. COMPARE: Classification of Morphological Patterns Using Adaptive Regional Elements. *IEEE Transactions on Medical Imaging*. 2007;26:93-105.

FAN Y, BATMANGHELICH N, CLARK CM, DAVATZIKOS C. Spatial patterns of brain atrophy in MCI patients, identified via high-dimensional pattern classification, predict subsequent cognitive decline. *NeuroImage*. 2008;39(4):1731-43.

FAN Y, SHEN D. Integrated Feature Extraction and Selection for Neuroimage Classification. *Proceedings of SPIE*. 2009;10(35):72591U.1-72591U.8.

FORD J, SHEN L, MAKEDON F, FLASHMAN LA, SAYKIN AJ. Combining Structural and Functional Features for Classification of Schizophrenia. *DEVLAB-FMRI-02-01*; 2002.

FORD J, FARID H, MAKEDON F, FLASHMAN L, MCALLISTER T. Patient classification of fMRI activation maps. In: *international Conference on Medical Image Computing and Computer Assisted Intervention*; 2003.

GERARDIN E, CHETELAT G, CHUPIN M, CUINGNET R, DESGRANGES B, Kim HS, *et al*. Multidimensional classification of hippocampal shape features discriminates Alzheimer's disease and mild cognitive impairment from normal aging. *NeuroImage*. 2009;47:1476-1486.

GUYON I, WESTON J, BARNHILL S, VAPNIK V. Gene Selection for Cancer Classification using Support Vector Machines. *Machine Learning*. 2004;46:389-422.

HORN J, HABERT MO, KAS A, MALEK Z, MAKSUD P. Differential automatic diagnosis between Alzheimer's disease and frontotemporal dementia based on perfusion SPECT images. *Artificial Intelligence in Medicine*. 2009;47:147-158.

HUA X, LEOW AD, PARIKSHAK N, LEE S, CHIANG MC, TOGA AW, *et al*. Tensor-based morphometry as a neuroimaging biomarker for Alzheimer's disease: An MRI study of 676 AD, MCI, and normal subjects. *NeuroImage*. 2008;47:1476-1486.

HUA X, LEE S, YANOVSKY I, LEOW AD, CHOU YY, HO AJ, *et al*. Optimizing power to track brain degeneration in Alzheimer's disease and mild cognitive impairment with tensor-based morphometry: An ADNI study of 515 subjects. *NeuroImage*. 2009;48:668-681.

JOHNSON KA, JONES K, HOLMAN BL, BECKER JA, SPIERS PA, SATLIN A, *et al*. Preclinical prediction of Alzheimer's disease using SPECT. *Neurology*. 1998;50:1563-1571.

KLOPPEL S, DRAGANSKI B, GOLDING CV, CHU C, NAGY Z, COOK PA, *et al*. White matter connections reflect changes in voluntary-guided saccades in pre-symptomatic Huntington disease. *Brain Advance*. 2007;131(1):196-204.

LAO Z, SHEN D, XUE Z, KARACALI B, RESNICK, M S, *et al*. Morphological Classification of brains via high-dimensional shape transformations and machine learning method. *Neuroimage*. 2004;21:46-57.

LI S, SHI F, PU F, Li X, JIANG T, XIE S, WANG Y. Hippocampal shape analysis of Alzheimer disease based on machine learning methods. *American Journal of Neuroradiology*. 2007;28:1339-1345.

MAGNIN B, MESROB L, KINKINGNHUN S, PLGRINI-ISSAC M, COLLIOT O, SARAZIN M, *et al*. Support vector machine-based classification of Alzheimers disease from whole-brain anatomical MRI. *Neuroradiology*. 2009;51:73-83.

MARENCO S, WEINBERGER DR. The neurodevelopment hypothesis of schizophrenia: following a trail of evidence from cradle to grave. *Development and Psychopathology*. 2000;12:501-527.

MCEVOY LK, FENNEMA-NOTESTINE C, RODDEY JC, DONALD J HAGLER J, HOLLAND D, KAROW DS, *et al*. Alzheimer Disease: Quantitative Structural

Neuroimaging for Detection and Prediction of Clinical and Structural Changes in Mild Cognitive Impairment. *Radiology*. 2009;251:195-205.

MECHELLI A, PRINCE JC, FRISTON K, ASHBURNER J. Voxel-Based Morphometry of the Human Brain: Methods and Applications. *Current Medical Imaging Reviews*. 2005;1:105-113.

MISRA C, FAN Y, DAVATZIKOS C. Baseline and longitudinal patterns of brain atrophy in MCI patients, and their use in prediction of short-term conversion to AD: Results from ADNI. *NeuroImage*. 2009;44:1415-1422.

PEREIRA F, MITCHELL T, BOTVINICK M. Machine learning classifiers and fMRI: A tutorial overview. *NeuroImage*. 2009;45:S199-S209.

PLATT JC. Fast training of support vector machines using sequential minimal optimization. In: *Advances in kernel methods: support vector learning*. MIT Press; 1999. p. 185-208.

RAMÍREZ J, CHAVES R, GÓRRIZ JM, LÓPEZ M, SALAS-GONZÁLEZ D, ÁLVAREZ I, *et al*. SPECT Image Classification Techniques for Computer Aided Diagnosis of the Alzheimer Disease. In: *International Work-Conference on Artificial Neural Network*; 2009.

ROBINSON EC, HAMMERS A, ERICSSON A, EDWARDS AD, RUECKERT D. Identifying population differences in whole-brain structural networks: A machine learning approach. *NeuroImage*. 2010;50:910-919.

SAVIO A, GARCÍA-SEBASTIÁN M, GRAÑA M, VILLANÚA J. Results of an Adaboost Approach on Alzheimers Disease Detection on MRI. In: *Proceedings of the 3rd International Work-Conference on The Interplay Between Natural and Artificial Computation: Part II: Bioinspired Applications in Artificial and Natural Computation*; 2009.

SHEN L, SAYKIN A, MCHUGH T, WEST J, RABIN L, WISHART H, CHUNG MK, *et al*. Morphometric Analysis of 3D Surfaces: Application to Hippocampal Shape in Mild Cognitive Impairment. In: *International Conference on Computer Vision, Pattern Recognition and Image Processing*; 2005.

STOECKEL G, FUNG J. SVM feature selection for classification of SPECT images of Alzheimer's disease using spatial information. In: *Fifth IEEE International Conference on Data Mining*; 2005.

TRIPOLITI EE, FOTIADIS DI, ARGYROPOULOU M. An automated supervised method for the diagnosis of Alzheimers disease based on fMRI data using weighted voting schemes. In: *International Workshop on Imaging Systems and Techniques*; 2008.

WANG Z, CHILDRESS AR, WANG J, DETRE JA. Support vector machine learning-based fMRI data group analysis. *NeuroImage*. 2007;36:1139-1151.

WANG J, de HAAN G, UNAY D, SOLDEA O, EKIN A. Voxel-based Discriminant Map Classification on Brain Ventricles for Alzheimers Disease. *Proceedings of SPIE*; 2009;10(35):72591W.1-72591W.8.

XU Y, JACK J, O'BRIEN PC, KOKMEN E, SMITH GE, IVNIK RJ, *et al*. Usefulness of MRI measures of entorhinal cortex versus hippocampus in AD. *Neurology*. 2000;54:1760-1767.

ZHANG J, YAN B, HUANG X, YANG P, HUANG C. The diagnosis of Alzheimers disease based on voxel-based morphometry and support vector machine. In: *Fourth International Conference on Natural Computation*; 2008.