UNIVERSIDAD NACIONAL DE COLOMBIA

SEDE BOGOTÁ

Morphometric Data Fusion for Early Detection of Alzheimer's Disease

Diana Lorena Giraldo Franco

Universidad Nacional de Colombia
Facultad de Medicina, Departamento de Imágenes Diagnósticas
Bogotá D.C., Colombia
2015

Morphometric Data Fusion for Early Detection of Alzheimer's Disease

Diana Lorena Giraldo Franco

Tesis o trabajo de grado presentada(o) como requisito parcial para optar al título de:
Magíster en Ingeniería Biomédica

Director(a):
Eduardo Romero Castro, Ph.D.

Línea de Investigación:
Imágenes Médicas
Grupo de Investigación:
Computer Imaging and Medical Applications Laboratory - Cim@lab

Universidad Nacional de Colombia
Facultad de Medicina, Departamento de Imágenes Diagnósticas
Bogotá D.C., Colombia
2015

A mis padres

Agradecimientos

A mi director, Eduardo Romero por su dedicación, orientación y ejemplo, su visión fue fundamental para el desarrollo de este trabajo. Quiero agradecerle por la paciencia y la confianza que ha depositado en mi; por todas las conversaciones que muchas veces incluyeron regaos y tambin felicitaciones, éstas han sido importantísimas en mi formacin profesional y personal. También quiero agradecerle por el apoyo incondicional que me ha brindado durante estos años en todos los aspectos posibles.

A mis padres, Ruby y Rafael por todo el amor, apoyo y empeo por darme las mejores oportunidades posibles; su ejemplo de rectitud, solidaridad y trabajo honesto han sido vitales para alcanzar este logro. También por enseñarme, a veces de manera involuntaria, a enfrentar los retos y asumir las dificultades de la vida con fuerza, valentía y convicción.

A los integrantes del grupo de investigación, por el tiempo compartido, por sus opiniones y aportes, y por constituir un maravilloso entorno de trabajo. A Juan David y a Ricardo quienes aportaron ideas fundamentales, guiaron mi trabajo y me dieron su apoyo, tiempo y colaboración.

A todas las personas que de alguna manera se involucraron en el desarrollo de este trabajo, a todos los que me colaboraron, apoyaron e inspiraron quiero expresarles mi profundo agradecimiento.

Abstract

We present a morphometry method which uses brain models generated using Nonnegative Matrix Factorization (NMF) characterized by signatures calculated from perceptual features such as intensities, edges and orientations, of some regions obtained by comparing the models. Two different measures are used to calculate volume-models distances in the regions of interest. The discerning power of these distances is tested by using them as features for a Support Vector Machine classifier.
This work shows the usefulness of both measures as metrics in medical image applications when they are used in binary classification tasks. Our methodology was tested with two experimental groups extracted from a public brain MR dataset (OASIS), the classification between healthy subjects and patients with mild AD reveals an equal error rate (EER) measure which is better than previous approaches tested on the same dataset (0.1 in the former and 0.2 in the latter). When detecting very mild AD, our results (near to 75% of sensitivity and specificity) are comparable to the results with those approaches.
Keywords: Alzheimer's Disease, MRI, Morphometry, NMF, Pattern Recognition, Kullback-Leibler Divergence, Earth Mover's Distance.

Resumen

Presentamos un método de morfometrí que usa modelos de cerebro que se generan usando factorización de matrices no-negativas (NMF por su nombre en inglés) y se caracterizan por firmas calculadas de rasgos perceptules como las intensidades, bordes y orientaciones de algunas regiones del cerebro obtenidas de la comparación entre modelos. Dos medidas, la divergencia de Kullback-Leibler y la "Earth Mover's Distance", son usadas para calcular la distancia entre volúmenes y modelos en las regiones de interés. Probamos el poder discriminante de estas distancias usándolas para construir los vectores de características para una máquina de soporte vectorial.
Este trabajo muestra la utilidad de ambas medidas en tareas de clasificación binaria. Nuestra metodología fue probada con dos grupos experimentales extraídos de la base de datos OASIS, la clasificación entre sujetos sanos y pacientes con Alzheimer leve revela un EER que mejora los resultados obtenidos por trabajos publicados previamente con los mismos grupos experimentales. Cuando se trata de detectar Alzheimer muy leve, los resultados (cercanos a 75% de sensibilidad y especificidad) son comparables con los resultados obtenidos en dichas publicaciones.
Palabras Clave: Enfermedad de Alzheimer, IRM, Morfometría, NMF, Reconocimiento de Patrones, Divergencia de Kullback-Leibler, EMD.

Contents

1 Introduction

1.1 Alzheimer's Disease

Alzheimer's disease (AD) is a neurodegenerative diseas that affects higher brain functions like memory, thinking and behavior. The neuropathological changes asociated with AD include intraneuronal lesions (neurofibrillary tangles) and extracellular parenchymal lesions (senile plaques), these changes are often acompanied by synaptic loss and vascular amyloid deposits [1, 2].

Although the disease progression has been traditionally assessed under the Braak and Braak staging scheme [3], several reports [4, 5, 6, 7] have demonstrated a very variable AD clinical picture: Neither the progression patterns nor the same anatomical areas are involved or follow a reproducible anatomic sequence, even in series of patients belonging to comparable social and cultural environments. Approximately 25% of AD brains show atypical patterns of structural damage, usually classified as hippocampal sparing and limbic predominant AD [8]. Furthermore, AD can also manifest different clinical pictures, such as the posterior, logopenic and frontal variants (IGW-2).

There are a few drugs that temporarily improve the symptoms of Alzheimers disease by increasing the amount of neurotransmitters in the brain. However, no treatment is available to slow or stop the deterioration of brain cells in Alzheimer's disease, but researches around the world are studying numerous treatment strategies that may have the potential to change the course of the disease, approximately 75 to 100 experimental therapies aimed at slowing or stopping the progression of Alzheimer's are in clinical testing in human volunteers [9].

Excluding the rare cases of AD caused by genetic mutations (less than 1% of cases), experts believe that Alzheimer's develops as result of multiple factors rather than a single cause. The greatest risk factor for AD is age, but the disease is not a normal part of the ageing process, other known risk factors for AD are: family history, $APOE\epsilon - 4$ gene, Mild Cognitive Impairment (MCI), cardiovascular diseases, traumatic brain injury and cognitive engagement [10].

The AD is the most common type of dementia and the fifth leading death cause of people over 65 years [11], the increased life expectancy and aging of the general population have

made of Alzheimer's Disease (AD) a growing public health concern. In 2013, over 35 million people worldwide were living with the disease, this number is expected to double by 2030 and more than triple by 2050 to 115 million [12]. Other reports calculate that by the year 2050 there will be approximately 135 million patients suffering from different stages of AD [13].

As the global population ages, the number of deaths related to Alzheimer's disease are growing in almost every country [14]. As example, between 2000 and 2013, deaths attributed to Alzheimer's disease increased 71 percent in United States [10]. Figure **1-1** shows the increasing number of deaths related to AD and other dementias in Colombia between 1998 and 2012.

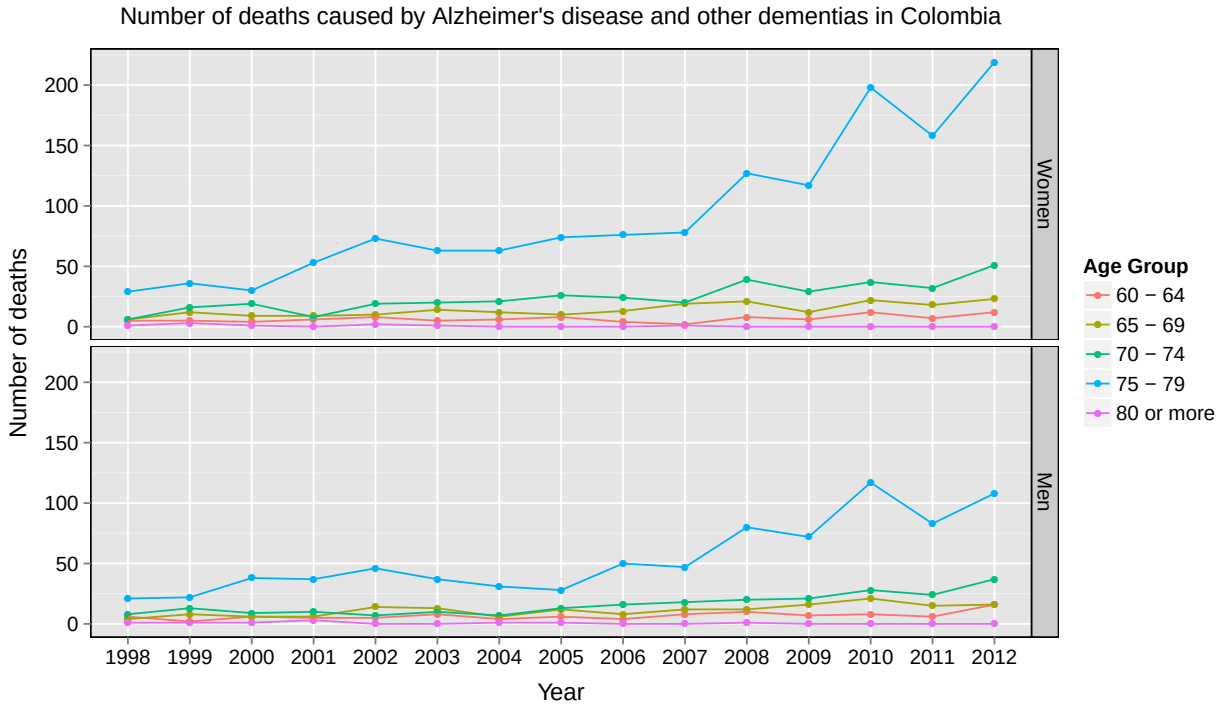

Figure **1-1**: Number of deaths realted to AD and other dementias of people over 60 years. This data was taken from the death records of DANE (Departamento Administrativo Nacional de Estadistica).

It is important to note that even in countries with complete medical certification of causes of death, the number of deaths due to AD and other dementias has been systematically underestimated because the lack of accurate diagnosis and theses deaths were associated to broad causes such as senility.

1.2 Diagnosis of AD

The clinical diagnosis of probable AD has been based on the criteria established in 1984 by the National Institute of Neurological and Communicative Disorders and Stroke-Alzheimers Disease and Related Disorders Association (NINCDS-ADRDA) [15] and the criteria from the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR) [16], these criteria specify eight cognitive domains that may be impaired in AD: memory, language, perceptual skills, attention, constructive abilities, orientation, problem solving and functional abilities. In 2010 the National Institute on Aging and the Alzheimers Association proposed recommendations to update the diagnostic criteria for Alzheimers dementia and MCI [17]. All of the recommendations incorporate the use of biomarkers for diagnosis, Among the biomarkers being considered are brain volume, level of glucose metabolism in the brain, presence of beta-amyloid in the brain and levels of beta-amyloid and tau in cerebrospinal fluid.

All the clinical criteria proposed allow diagnosis of probable disease, a definitive diagnosis of AD can only be made via histopathological confirmation of amyloid plaques and neurofibrillary tangles but these confirmations studies are usually performed *post mortem*. Although precise guidelines for AD diagnosis exist [18] the final responsibility of deciding whether the patient is suffering from AD or not falls upon the physician.

AD is commonly detected when a physician is consulted due to cognitive impairment or memory complaints. The initial diagnosis is based on the patient's clinical history and a battery of neuropsychological tests measuring different disease aspects such as symptom severity, interference with the patient activities, or cognitive impairment, among others. Certain routine laboratory test and routine brain imaging are also recommended, but these are to rule out other conditions that can cause cognitive dysfunction. The accuracy of the diagnosis is highly dependent on the examiner's skills and on the evolution of a variable clinical frame. Studies report that only 50% of the cases of probable dementia are correctly diagnosed [19].

1.2.1 Neuropsychological Tests

The neuropsychological tests used in the AD diagnosis process include the following:

- The Clinical Dementia Rating (CDR) is used to quantify the severity of symptoms of dementia. The test consists on a semi-structured interview with an informant and the patient to measure the cognitive and functional performance in six areas: memory, orientation, judgment and problem solving, community affairs, home and hobbies, and finally, personal care. In the scoring rules memory is considered the primary category

and all others are secondary. The current version of this test was published by Morris in 1993 [20].

- The MiniMental State Examination (MMSE) or Folstein Test is used to estimate the severity of cognitive impairment evaluating the following aspects: orientation to time and place, registration, attention and calculation, recall, language, repetition and complex commands. This test was published in 1975 by Folstein *et al.* [21].
- The Alzheimers Disease Assessment Scale - Cognition (ADAS-Cog) is used to evaluate cognitive impairment in the assessment of Alzheimers disease, its primary purpose was to be an index of global cognition in response to antidementia therapies. This scale was proposed by Rosen *et al.* in 1984 [22].

1.2.2 Early Diagnosis

The brain changes underlying Alzheimer's disease probably develop over a period of at least 10-20 years prior to the onset of symptoms [23], and researches believe that this may be the period where future Alzheimer drugs and treatments will be most effective.

The primary purpose of early diagnosis is timely access to information, advice and support for the patients, their relatives and caregivers. Other important benefits of early detection and diagnosis are [9]:

- Prompt evaluation and treatment of reversible or treatable causes of cognitive impairment.
- Enables potential inclusion in Alzheimer clinical trials.
- Helps prevent prescription of medications for coexisting conditions that worsen cognitive function.
- Aids management of possible behavioral symptoms.
- Allows planning for the future.

The increasing burden of the disease due to population growing and ageing is becoming a worldwide public health concern, researchers around the world are working in the development of medical treatments that should be given when extensive and irremediable damage has not occurred yet, this is the period when brain changes are beginning but major symptoms as memory loss have not appear.

Given that the first AD symptoms may be confused with other conditions, there is a strong interest in developing objective tools that allow the early clinical detection and the recognition of the prodromal expression of AD, making this a burgeoning area of research.

1.3 Medical Imaging in AD

Due to their noninvasive and harmless nature, neuroimages constitute a potential source of information, but as mentioned before, their utility remains limited.
The prospect of using neuroimaging techniques as both an early detection and a confirmation tool is very appealing and has generated a large body of research. In the clinical practice, however, neuroimaging techniques have had, until now, only a marginal role: Their main use is to exclude other pathological conditions or to visualize the anatomic pattern of neurodegeneration. Recent reviews on the NINCDS-ADRDA Alzheimer's Criteria [18] recommend using MR neuroimages as a supportive diagnosis tool while the IWG-2 criteria suggest the use of amyloidPET as evidence of Alzheimer's pathology. Volumetry of the hippocampus and the medial temporal lobe (MTL) has gained wide acceptance as a diagnostic tool and biomarker of the disease progression, yet the disease management is independent of these measures. It should be noted that the presence of atypical AD pathology and other diseases, such as the hippocampal sclerosis, compromise the reliability of the data obtained with this technique.

Structural Magnetic Resonance Imagin (MRI) provide precise anatomical information but are rarely used to detect AD as the anatomical changes it induces tend to be slow, subtle and hard to differentiate from normal brain aging. Despite this, recent articles have used morphometry based approaches to identify the anatomical differences between groups of healthy and AD patients [24, 25, 26].

1.4 State-of-the-Art

1.4.1 Brain Morphometry

Several studies report the use of sophisticated measurement techniques that assess anatomical changes in areas compromised by AD such as the intracranial volume, the cortical thickness or volumes of subcortical structures [27]. The computational methods that perform these measurements are collectively known as *Brain Morphometry* (BM) or, in some cases, computational neuroanatomy or neuromorphometry [28]. As the names suggest, these methods are only based on form, size and/or shape derived features extracted from the brain structures.

Features such as locations (landmarks), voxel intensities, template deformations or surface representations have been frequently used in morphometrical studies, giving place to specific morphometric techniques such as landmark-based morphometry [29], voxel-based morphometry [30], deformation-based or tensor-based morphometry [31] and surface-based morphometry [32].

Voxel-Based Morphometry (VBM), proposed by Ashburner and Friston in 2000 [30], is by far the most common morphometric approach used by the neuroscience research community. Based on brain tissue segmentation and voxel-by-voxel global statistics, VBM is driven by the mesoscopic differences, *i.e.* local differences, between subject volumes and the entire population. By contrast, Deformation-Based Morphometry (DBM), proposed in 1998 by Ashburner et al. [31] models macroscopic anatomical differences among brains using the deformation fields that warp individual brains to a common reference space. The deformation fields are expected to encode shapes of individual brains, providing additional information to the shape analysis about lengths, areas and angles, among others.

A drawback of morphometrical analyses is the need of an accurate intersubject registration that guarantees the comparison of homologous structures across all subjects. However, this kind of one-to-one correspondence between subjects may not always be achieved, mainly because of the inherent intersubject anatomical variability and the effects of brain pathologies. To cope with this, Toews *et al.* [33] proposed in 2010 Feature-Based Morphometry (FBM) devised to find image patterns that might not occur in all subjects. Since each subject is modeled as a collection of brain features, a probabilistic framework estimates the relationship between features and subjects, thereby clustering the co-ocurrence between features and classes.

It has to be noted that several these morphometry methods used to detect AD are based or rely on segmentation techniques: Some studies use automatic segmentation of main tissues such as cerebrospinal fluid, gray and white matter [26, 34, 35, 36, 37], other studies use semi-automatic segmentation of specific areas such as the hippocampus [38], or combinations of them [39, 40, 19, 41, 42].

1.4.2 Pattern Recognition Techniques

The use of machine learning and elaborate data interpretation techniques has become common for the analysis of neurological images due to their high dimensionality and complexity. These approaches combine different sets of features obtained from neurological volumes as input to supervised classification algorithms that learn how to assign predefined labels, *e.g.* AD or NC, to previously unseen volumes [26]. Part of the appeal of these methods comes from their increased discriminating power due to the combination weighted sets of features instead of having to rely on a single one. Machine learning in this context works by learning a set of bases from the feature space in which the brains are encoded and finding relevant relationships among them. The choice of these bases is dependent on an output variable describing a particular condition to be explained, *e.g.* the presence or absence of particular pattern in neurodegenerative disorders, or the identity of a recorded stimulus or task.

A common approach in the analysis of neuroimages is reduced to a classification problem, that is, the objective of the methods is to extract features from the images that can be used to assign labels related to clinical conditions, i.e. AD/nonAD or MCI/AD. Typical examples of this approach are Support Vector Machines (SVM) and ensemble classifiers. The former are binary classifiers broadly used in different artificial vision problems. SVM assume that there is a feature space in which the two classes are separable by a hyperplane. The use of SVM in neuroimaging research has been reported in various publications[24, 43].

1.5 Proposed Approach

In the present work we propose a methodology that extracts characteristic morphological information of groups of Magnetic Resonance Images (MRI) from diverse information channels (e.g. intensities, edges and orientations) and then condenses this information into characteristic brain models.

We use Non-negative Matrix Factorization (NMF) to obtain brain volumes representing different stages of the disease, a comparison between these models gives us the most discriminant regions between classes for each information channel. Note that these regions are important not only because of their use as classification features but, more significantly, because they blend intuitively with the doctors' diagnostic thought process.

To validate our models, we calculate a signature for each one of the discriminant regions in the MRIs and the models and compare these regions in the MRIs against the regions in the brain models by using two similarity measures: the Kullback-Leibler divergence and the Earth Mover's Distance.

With the obtained measures, we train a SVM classifier where each feature vector was composed by all the similarity measures between the corresponding MRI and both brain volumes.

In this work we present an automatic methodology to learn a brain model for each stage of the disease and obtain discerning regions from them. With the regions and the models we obtain useful classification features that allow high sensitivity and specificity with lower dimensionality in the detection of AD.

2 Methods and Materials

2.1 Methods

The overall methodology is presented schematically in Figure **2-1**. The generation of the model volumes for each class is done using Nonnegative Matrix Factorization (NMF), RoI extraction is based on the thresholding of the difference between the model volumes, and two similarity measures are used to compare the signatures of corresponding regions.

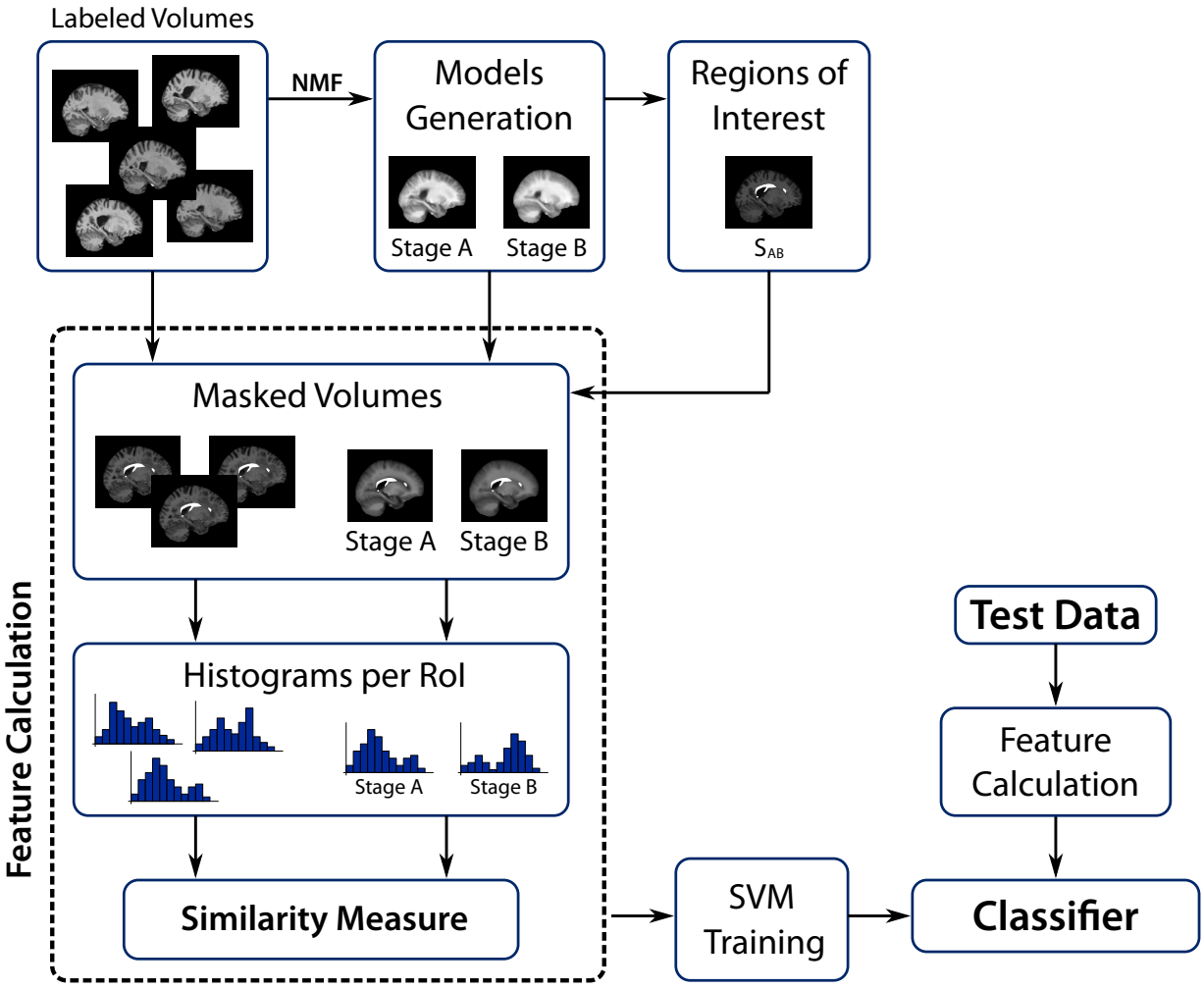

Figure **2-1**: Graphical flow representation of the proposed method.

2.1.1 Generation of Models

NMF is a dimensionality reduction technique which may be used to learn the most representative "parts" that compose the elements of a set, e.g. NMF has been used to learn individual parts of the face from a database and then use them to reconstruct and identify any individual face [44].
Given a matrix X, where each column is a d-dimensional vector of positive observations for each one of the N subjects, the NMF algorithm factorizes the matrix X into two non-negative matrices W and H:

$$X_{d\times N} = W_{d\times r} H_{r\times N} \tag{2-1}$$

with the following constraints:

$$w_{ik} \geq 0 \quad \text{and} \quad h_{kj} \geq 0, \qquad \forall i,k,j \tag{2-2a}$$

$$\sum_{i=1}^{d} w_{ik} = 1, \qquad \forall k \in \{1,\ldots,r\} \tag{2-2b}$$

Where r is the rank of the decomposition corresponding to the number of columns in W and the number of rows in H. The columns of W are called the basis vectors and the columns of H are the projections of the data into the sub-space spanned from W.
Each column in X corresponds to a brain volume I_j and H is fixed using class membership functions for each volume. Matrix W is found by solving the optimization problem of minimizing the Kullback-Leibler divergence (KL) between X and WH as in [44]. The rank of the decomposition r is the number of classes used in H, and it will be the number of columns in the obtained matrix W which are the brain models for the corresponding classes. This data arrangement is illustrated in Figure **2-2**

2.1.2 Obtaining Models by Fusing with Clinical Data

In the present application have been used three stages as classes: Normal Control subjects (NC), patients with very mild Alzheimer's disease (or Mild Cognitive Impairment) and patients with mild Alzheimer's disease (mAD). The matrix H is fixed using membership functions for each class that takes into account the neuropsychological information, in this work we have used the Clinical Dementia Rating to establish the membership of each subject to each class and posed binary functions as follows:

$$H_{jk} = f_j(I_k) = \begin{cases} 1 & I_k \in \text{Class } j \\ 0 & \text{otherwise} \end{cases} \tag{2-3}$$

With these values, the resulting matrix W is composed by the basis vectors V_{mAD}, V_{MCI} and V_{NC} which, in this case, tend towards the average of the class members.

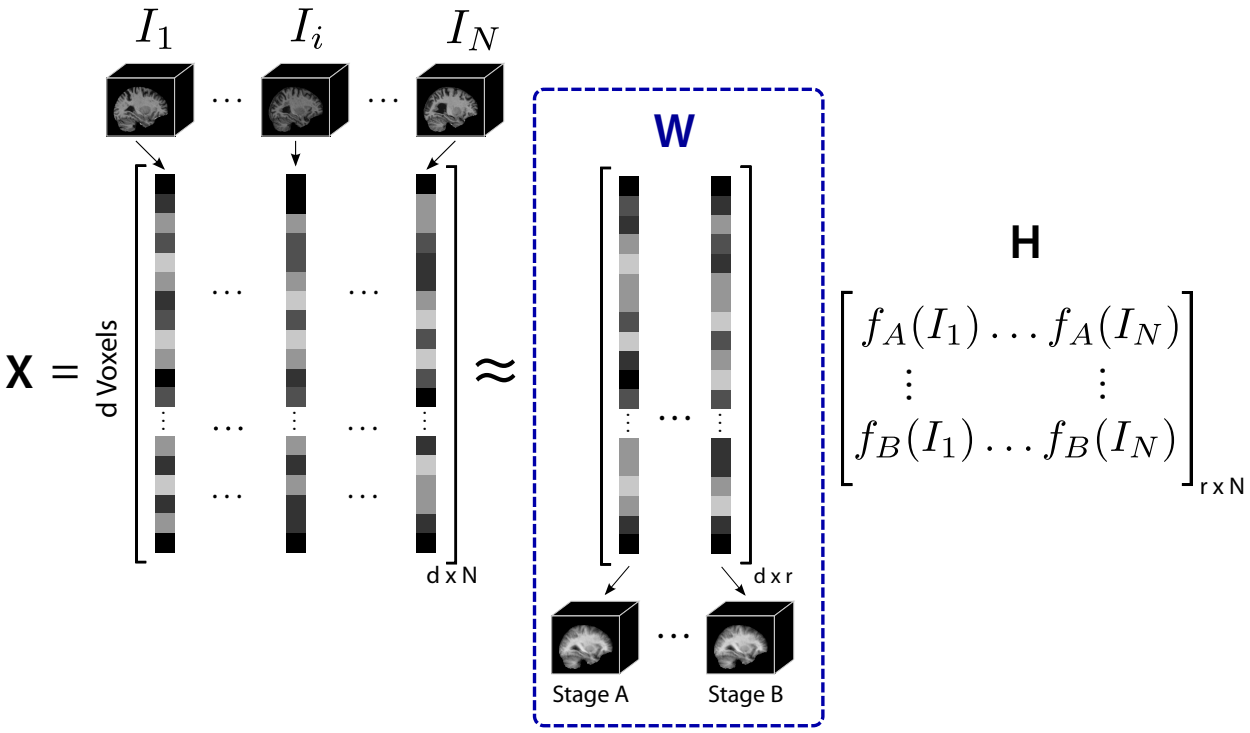

Figure **2-2**: Data arrangement for the NMF decomposition.

2.1.3 Regions Of Interest

A difference volume $D_{AB} = |V_A - V_B|$ is calculated to locate the regions with the most discriminant information between class A and class B. Then, a set of regions of interest is extracted from the difference volume D_{AB} by applying two thresholds: the first one binarizes D_{AB} and the second one is a size filter that removes very small areas that are not representative. The process is repeated for each information channel f, as result we obtain a set of RoI for each pair of classes and morphological information channel $\mathcal{S}^f_{AB} = \{S^1_{AB}, \dots, S^{R_f}_{AB}\}$, Figure **2-3** illustrates how this process works with different morphological information.
The set $\mathcal{S}^f_{AB}$ is used to mask the brain volumes $\{I^f_j\}^n_{j=1}$ and the model volumes $\{V^f_A, V^f_B\}$. Then, each of the regions is represented by a signature, in this case, a 256 or 64 bin histogram.

2.1.4 RoI Comparison

The histogram $h^f_{j,r}$ summarizes the information of the subject j in the region S^r_{AB}, that belongs to the set $\mathcal{S}^f_{AB}$, for the channel f. In a similar way, the histograms $\{h^f_{A,r}, h^f_{B,r}\}$ represent the information of the model volumes in the region S^r_{AB} for the channel f. Then, the region in the subject j is compared to the regions in the models using a similarity measure between the histograms, in this work we have used two measures: the Kullback-Leibler divergence (KL-div) and the Earth Mover's Distance (EMD).

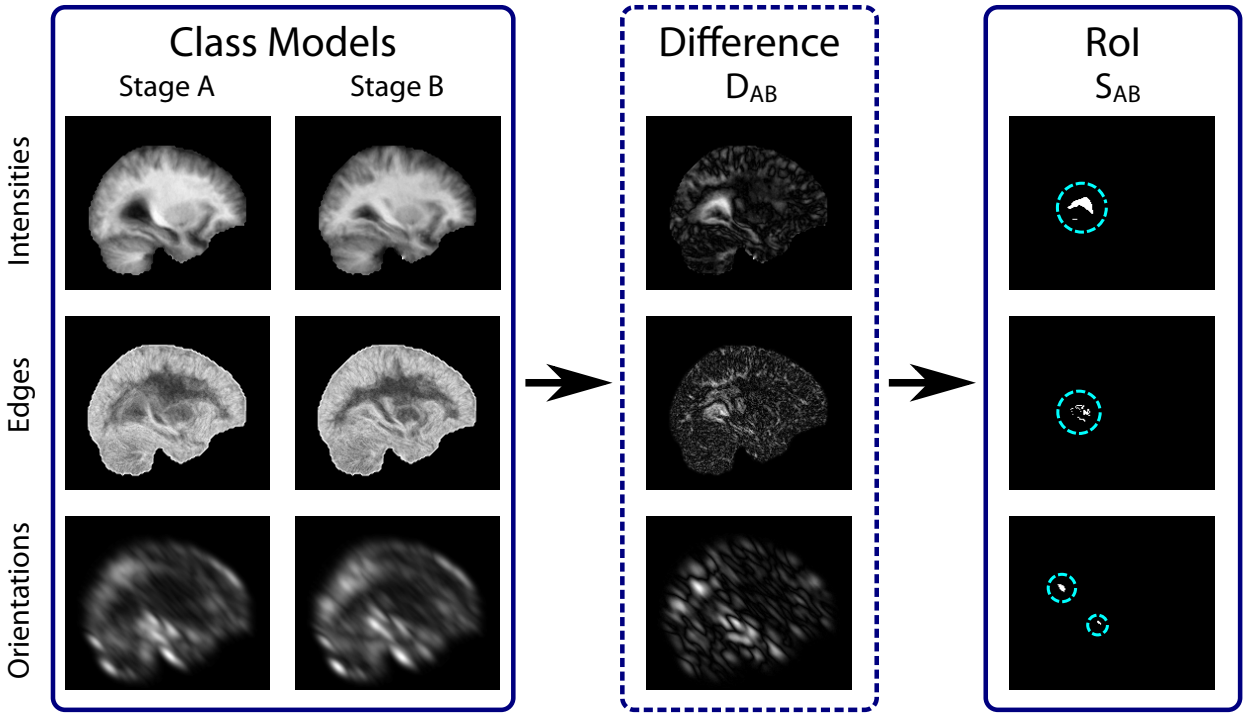

Figure **2-3**: An example of how the set of Regions of Interest (RoI) is obtained for some morphological information channel.

KL Divergence

The Kullback-Leibler divergence, also known as the relative entropy, is a widely used measure of the difference between probability density functions. Given two discrete probability distributions p and q, the KL divergence between p and q is a measure of the information lost when q is used to approximate p and is defined by:

$$D_{KL}(p\|q) = \sum_m p_m \ln \frac{p_m}{q_m} \tag{2-4}$$

Although $D_{KL}(p\|q) \neq D_{KL}(q\|p)$, so this measure is not symmetric and cannot be a true metric, it satisfies two important properties:

$$D_{KL}(p\|q) \geq 0 \quad \forall p, q \tag{2-5a}$$

$$D_{KL}(p\|q) = 0 \quad \text{if and only if} \quad p = q \tag{2-5b}$$

Note that this measure has problems when there are zero values in the probability distributions, to prevent this issue we have replaced the zeros in the histograms with a very small value just before the histogram normalization.
Then, for the subject j, the region S^r_{AB} and the channel f, we calculate the following pair

$$\left(D_{KL}(\hat{h}^f_{j,r}\|\hat{h}^f_{A,r}), D_{KL}(\hat{h}^f_{j,r}\|\hat{h}^f_{B,r})\right) \tag{2-6}$$

Where $\hat{h}^f_{j,r}$, $\hat{h}^f_{A,r}$ and $\hat{h}^f_{B,r}$ are the corresponding normalized histograms. This pair of numbers are measures of the amount of information lost when the regions in the models are used to represent the regions in the subjects.

Earth Mover's Distance

EMD is a similarity measure that calculates the minimum cost of transforming one histogram into another [45]. Finding the EMD is equivalent to solving an instance of the transportation measurement problem[46] in which n suppliers with fixed offers $\mathcal{S} = \{s_1, s_2, \ldots, s_n\}$ have to cover the demands $\mathcal{C} = \{c_1, c_2, \ldots, c_m\}$ of m consumers to which they are connected through paths with a positive given cost $\{p_{ij}\}_{i,j=1}^{n,m}$. The solution satisfies the consumers' demand while minimizing the transportation costs. The general transportation problem is illustrated in Figure **2-4**.

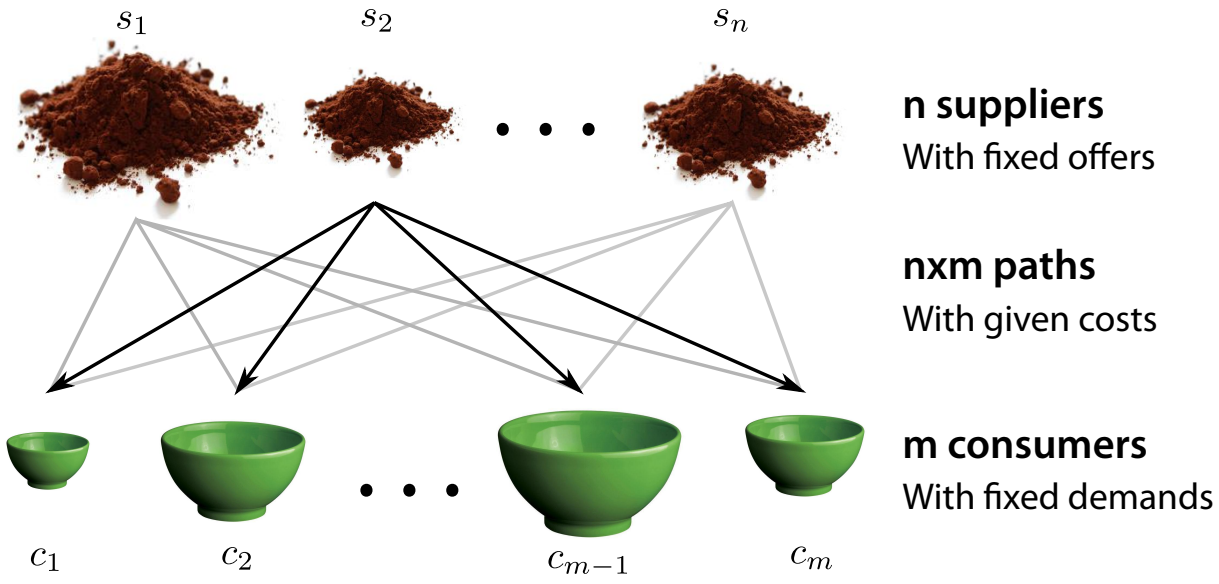

Figure **2-4**: Graphical explanation of the Transportation Problem.

This problem can be written in terms of the amount of "earth" x_{ij} that is moved from supplier i to consumer j:

$$\underset{X}{\text{minimize}} \quad \sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}x_{ij} \tag{2-7}$$

$$\text{subject to} \quad \sum_{j=1}^{m} x_{ij} \leq s_i, \qquad \text{for } i \in \{1,\ldots,n\} \tag{2-8a}$$

$$\sum_{i=1}^{n} x_{ij} \geq c_j, \qquad \text{for } j \in \{1,\ldots,m\} \tag{2-8b}$$

$$x_{ij} \geq 0, \qquad \text{for } i \in \{1,\ldots,n\} \text{ and } j \in \{1,\ldots,m\} \tag{2-8c}$$

If the solution of this problem is the set of values $\{x^*_{ij}\}^{n,m}_{i,j=1}$, one can say that the earth mover's distance between $\mathcal{S}$ and $\mathcal{C}$ is the total cost divided by the amount of mass that is moved:

$$EMD(\mathcal{S},\mathcal{C}) = \frac{1}{\sum_{ij} x^*_{ij}} \left(\sum_{i=1}^{n} \sum_{j=1}^{m} p_{ij} x^*_{ij} \right) \tag{2-9}$$

When comparing two histograms or signatures (a generalized form of histograms) these can be seen as the sets of offers and demands, if the histograms have the same mass, as in our case, the problem is symmetric and the EMD is a metric equivalent to the Wasserstein's distance [45]. Figure **2-5** shows a simple example of EMD between two histograms p and q.

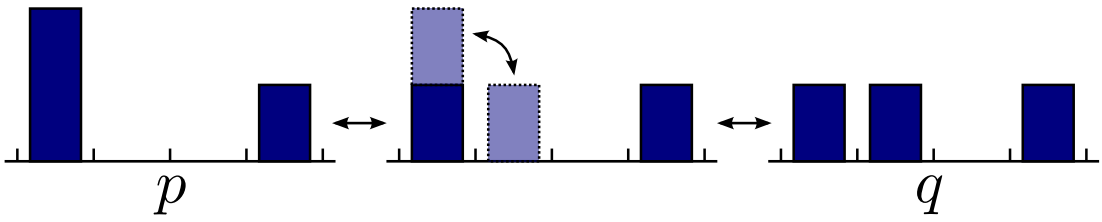

Figure **2-5**: In this case the EMD between p and q is the cost of moving one unit from one bin to the next divided by the total mass: 1/3.

As with the KL-div, for each subject, region and channel we calculated the following pair:

$$\Big(EMD(h^f_{j,r}, h^f_{A,r}), EMD(h^f_{j,r}, h^f_{B,r}) \Big) \tag{2-10}$$

For each subject, all pairs of measures along all the regions and channels are concatenated to make up a feature vector for the classification task. The length of these vectors are dependent on the number of RoI obtained for each channel and is given by the expression $\sum_f 2R_f$

2.2 Materials

2.2.1 MRI Data

In this work we have used T_1 weighted magnetic resonance images of the brain, the images come from the Open Access Series of Imaging Studies (OASIS) database [47], the images in this study were aquired with 1.5 T Vision scanner.

Pre-processing

As described in [47] the postprocessing includes: removal of facial features, correction for head movement, location of all volumes in the same coordinate system, affine registration to the Talairach and Tournoux atlas [48], skull removal and intensity inhomogeneity correction. All these steps were already for the images in the OASIS database.
We have registered the brain volumes to the MNI152 space using the affine registration tool FLIRT [49, 50] in the FSL library (`http://fsl.fmrib.ox.ac.uk/fsl/`), This procedure was done with the purpose of being able to identify anatomical regions using the structural atlases provided by the FSL library.
The result of this process is a single 1-mm isotropic volume with $176 \times 208 \times 176$ voxels.

2.2.2 Clinical Data

Dementia status of the subjects in OASIS was established and staged using the CDR scale, this scale allows the clinical diagnosis of probable Alzheimer's Disease in subjects with a CDR of 0.5 or greater.
We tested the region extracting method with a subset of the available data composed by individuals aged from 60 to 80 with a CDR not greater than 1. The distribution by gender and stage of the disease of this group is shown in Table **2-1**.

Table **2-1**: Distribution of the individuals in our experimental group.

CDR	Stage	Women	Men	Total
1	mild AD	13	7	20
0.5	very mild AD	28	22	50
0	NC	48	18	66
	All	89	47	136

The MMSE scores for all the subjects in our experimental group are also available as a supporting measure.

2.2.3 Morphological Information Extraction

For each volume, the edge channel was obtained by applying a $3D$ Sobel filter over the whole volume. The orientation channel was obtained by applying a $2D$ Gabor filters per slice in the sagittal axis.

2.3 Evaluation

To test the discriminant nature of the obtained regions and the usefulness of the chosen similarity measures, we have used the inter-region measures to build features to train a Support Vector Machine classifier (SVM). The training features for each subject are the measures calculated in 2.1.4, these features are concatenated and then they are used to train an SVM binary classifier using the libSVM library [51].
The validation was performed using a leave-one-out cross-validation. In this scheme, a single MRI is classified using the classifier trained with the remaining observations. The calculation of the feature vector for the test subject is based on the model volumes and sets of RoI obtained from the training data. Note that each iteration requires re-calculating a new set of model volumes and RoI to effectively isolate the test subject.
The validation of the whole methodology was performed for two experimental groups, the first one consisted of 66 control subjects and 20 patients diagnosed with mild AD, the distribution of data by gender is shown in Table **2-2**

Table **2-2**: Subjects distribution by gender.

	mild AD	NC	Total
Women	13	48	61
Men	7	18	25
Total	20	66	86

The second one was the whole group described in 2.2.1.
Classification with the first group aims to distinguish between two separated classes, when including the subjects with $CDR = 0.5$ we have 3 classes or stages of the diseases that are not easily separable, to distinguish between these classes we divided the task in two steps as shown in Figure **2-6**:

- Step I: in this step, all the subjects are taken and classified as NC or AD, the AD class includes patients with mild and very mild AD, this first step can be viewed as a diagnosis step.

- Step II: in this step we only included the patients diagnosed with probable AD (mild

or very mild) and performed the validation scheme to classify them as mild AD or very mild AD, this second step can be viewed as the grading the disease process.

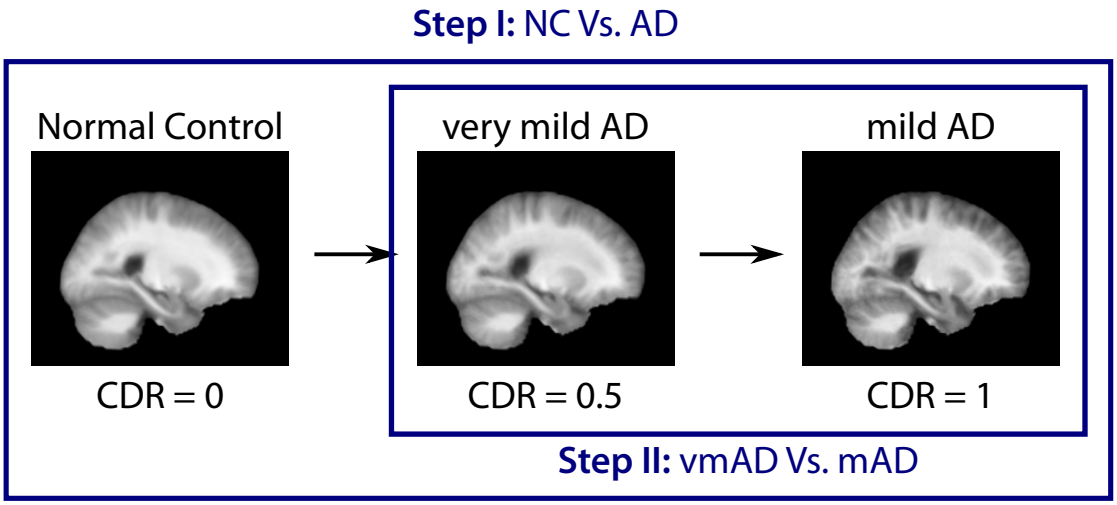

Figure **2-6**: Two steps classification

To account for the morphometric variability between genders, we repeated the validation scheme separately by gender. Cross validation was also done using leave-one-out.

3 Results

For each iteration of the validation procedure, the SVM classifier returns a value that is often interpreted as the probability of the test subject belonging to the positive class. With the values collected from the whole experimental group, a receiver operating characteristic (ROC) curve was constructed since this is the metrics more used in case of binary classification problems. The area under the curve (AUC), a global measure of the classification performance, and the equal error rate (EER), the rate for which both false positive rate (type I error) and false rejection rate (type II error) are equal, were calculated. The EER value means that there is a decision threshold in which it is possible to achieve simultaneously sensibility and specificity rates of $1 - \text{EER}$.

3.1 First Group: NC and mild AD

The first experimental group is described in Table **2-2**, it has two different classes: Normal Controls and patients diagnosed with mild Alzheimer's Disease (the positive class). The features for the SVM were calculated using the two similarity measures described in subsection 2.1.4.
Firstly, the complete validation scheme was performed in the whole group, and then we repeated the process but splitting the group by gender.

3.1.1 With Kullback-Leibler Divergence

In the experiments in which this metrics was used, histograms were divided into 256 bins. At each iteration of the validation process for this set of experiments, the SVM was trained with a polynomial kernel. Figure **3-1** and Figure **3-2** show the obtained ROC curves when using the whole group and when splitting the group by gender, respectively.

The curve in Figure **3-1** has an AUC of 0.96 and an EER of 0.10 these figures demonstrate the good performance of the classifier even when our experimental group is relatively small and has an important imbalance of classes, likewise the aging distribution is quite heterogeneous in the two classes. The fact that the experiment has shown so clear differences between the two classes suggests that definitely there are morphological patterns differently distributed in the two groups.

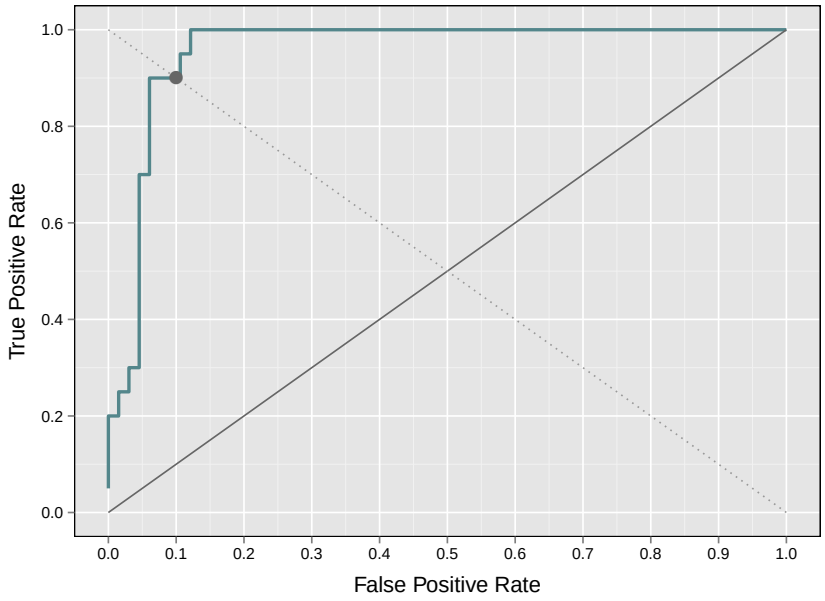

Figure **3-1**: ROC curve from the classification between mild AD and NC using the KL-Divergence as the similarity measure

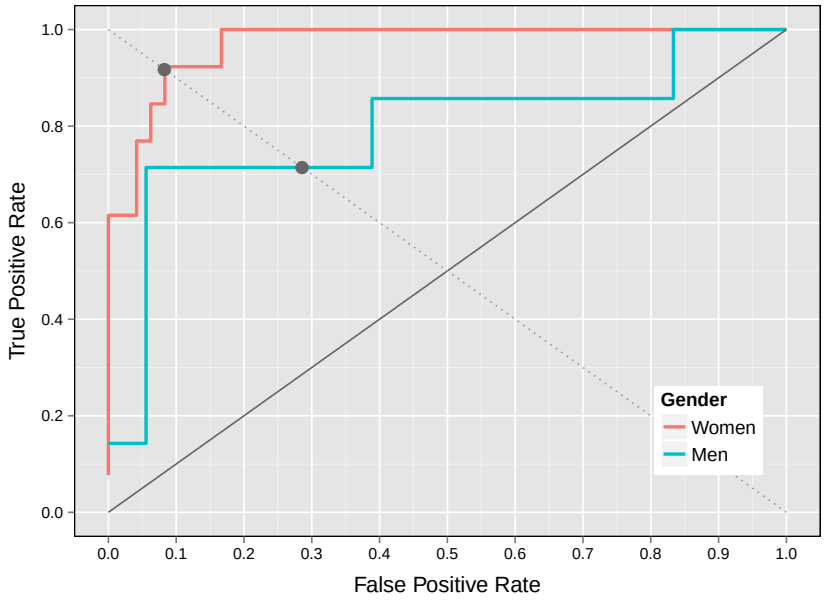

Figure **3-2**: ROC curve from the classification between mild AD and NC using the KL-Divergence as the similarity measure, separately by gender.

When the experimental group is divided by gender, the classification performance improves

for women, as shown in Figure **3-2** where the pink curve is closer to the top-left corner, this curve has an AUC of 0.97 and an EER of 0.08. On the contrary, the classification performance for men decreases, with an AUC of 0.79 and an EER of 0.29. This suggests that the patterns that might discriminate between AD and NC are more outstanding in women than in men. However, the decrease in the performance for men may be due to the low total number of cases (7 AD and 18 NC).

3.1.2 With Earth Mover's Distance

The histograms that were compared using EMD have only 64 bins. The SVM kernel used in this set of experiments is linear, i.e. a first degree polynomial kernel, meaning that the separation in the feature space is given by a simple hyperplane. The obtained ROC curves when using the whole group and when splitting the group by gender are shown in Figure **3-3** and Figure **3-4** respectively.

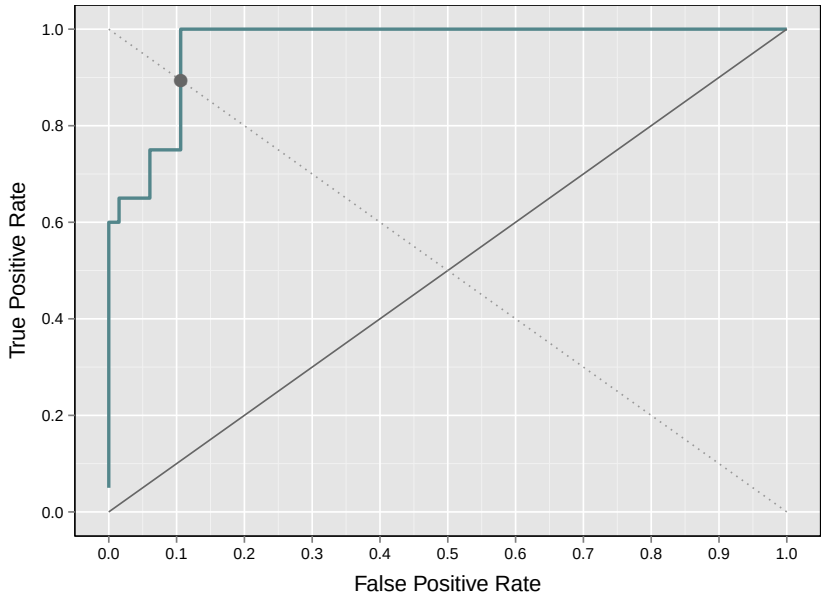

Figure **3-3**: ROC curve for the classification between mild AD and NC using the EMD as the similarity measure

The curve in Figure **3-3** has an AUC of 0.97 and an EER of 0.11, similar to the curve in Figure **3-1**, this shows that using the EMD reduces the dimensionality of the characterization of RoI without affecting the classification performance, which is an aside benefit in terms of the computation. It is noticeable that there is a point in this ROC curve where it is possible to achieve a sensitivity of 100% with a specificity of 90%.

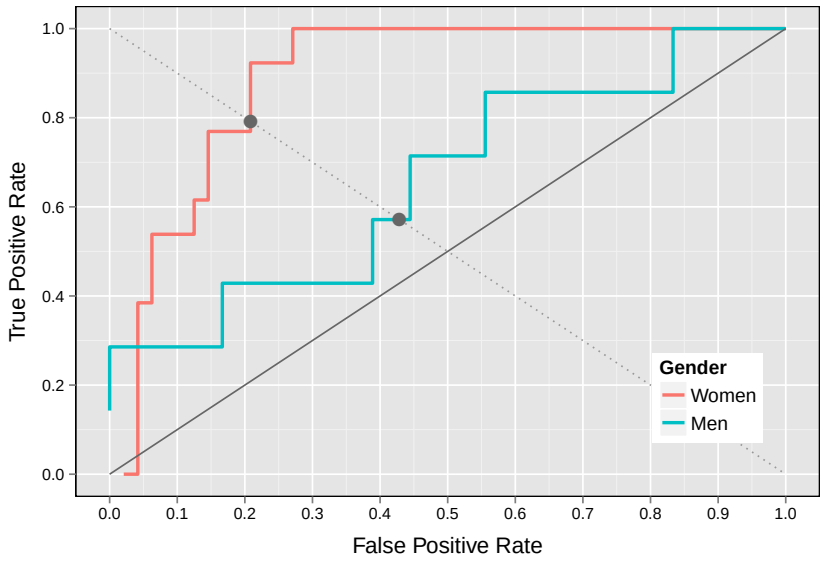

Figure **3-4**: ROC curve for the classification between mild AD and NC using the EMD as the similarity measure, separately by gender.

When splitting the group by gender, the classification performance for both women and men decreases as shown in Figure **3-4**, the ROC curve for women shows an AUC of 0.89 and an EER of 0.21 while the ROC curve for men has an AUC of 0.66 and an EER of 0.43, as illusrated in Figure **3-2** the classification performance for men is worst than for women, this observation reinforces the idea that the NC and mild AD patterns differ more in women than in men.

The Area Under the Curve and the Equal Error Rate of these ROC curves for both sets of experiments are shown in Table **3-1**.

Table **3-1**: Classification performance for both sets of experiments. The sensitivity and sensibility of them is given by $1 - \text{EER}$.

	All		Women		Men	
	EER	AUC	EER	AUC	EER	AUC
Using KL-Div	0.10	0.96	0.08	0.97	0.29	0.79
Using EMD	0.11	0.97	0.21	0.89	0.43	0.66

3.2 Second Group: NC, very mild AD and mild AD

The second group is the same described in Table **2-1** and is composed of three classes: Normal Controls, patients diagnosed with mild Alzheimer's Disease and an intermediate group, composed of patients diagnosed with very mild AD, which is the Mild Cognitive Impairment. With this group, the experiments were also conducted using both similarity measures, the KL-Divergence and the EMD.
As described in Section 2.3, the experimentation with this group consists of two binary classification tasks:

1. AD Detection: This step distinguishes between normal controls and patients with AD (mild or very mild).

2. Grading the Stage: This phase discriminates between mild and very mild AD.

All the experiments were carried out by separating the group by gender.

3.2.1 With Kullback-Leibler Divergence

In this set of experiments, the histograms were compared using KL divergence, having also 256 bins and the SVM was trained with a polynomial kernel. The obtained ROC curves for steps 1 and 2 are shown in Figure **3-5** and Figure **3-6** respectively.

The AD detection is a more difficult task when the subjects with very mild AD are included in the experimental group, Figure **3-5** shows the ROC curves corresponding to this classification task when the KL-Divergence is used as a similarity measure between RoIs. The classification performance is very similar for both genders, the ROC curve for women shows an AUC of 0.82 and an EER of 0.27, while the ROC curve for men has an AUC of 0.81 and an EER of 0.21, suggesting that the descriptive patterns start to get blurred in case of the women group, while remains almost the same, or even better, for men. This could be attributed to the inclusion of more men with AD in the new experimental group.

The task of distinguish between stages of the disease is much more difficult than the aforementioned, Figure **3-6** shows that the classification performance for men, in terms of the EER, is slightly better than for women, the ROC curve for women has an AUC of 0.73 and an EER of 0.36 while the ROC curve for men shows an AUC of 0.71 and an EER of 0.32.

3.2.2 With Earth Mover's Distance

When using the EMD, the histograms are composed of 64 bins and the SVM was trained with a linear kernel. The obtained ROC curves for steps 1 and 2 are shown in Figure **3-7**

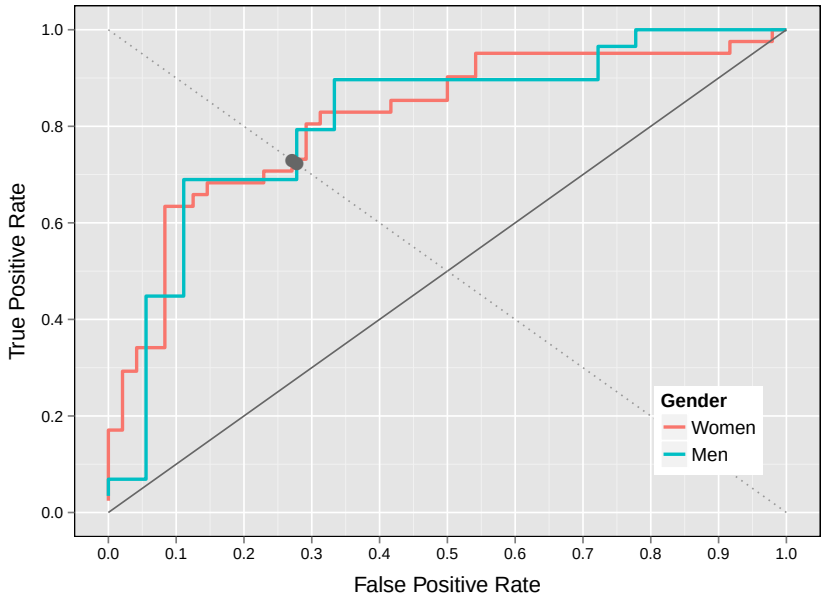

Figure **3-5**: ROC curve from the classification between AD and NC using the KL-Divergence as the similarity measure, separately by gender.

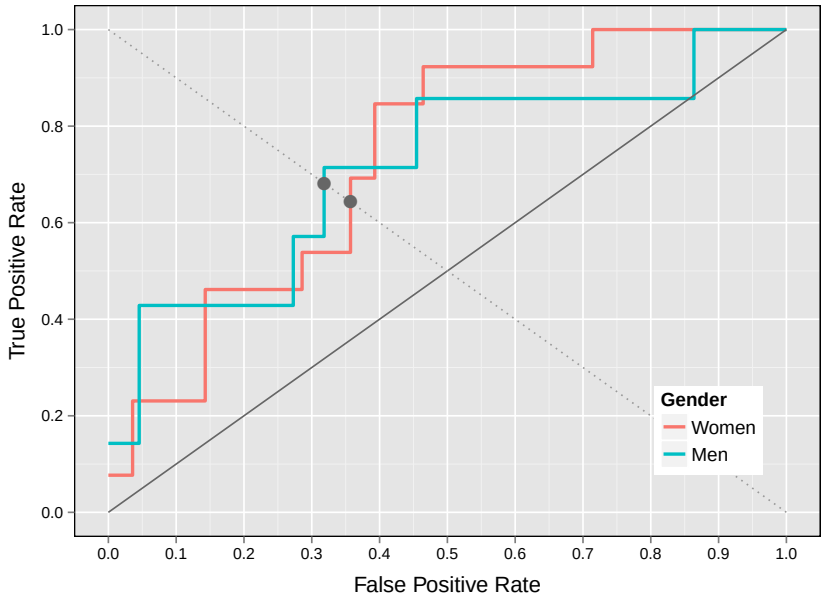

Figure **3-6**: ROC curve from the classification between very mild AD and mild AD using the KL-Divergence as the similarity measure, separately by gender.

and Figure **3-8** respectively.

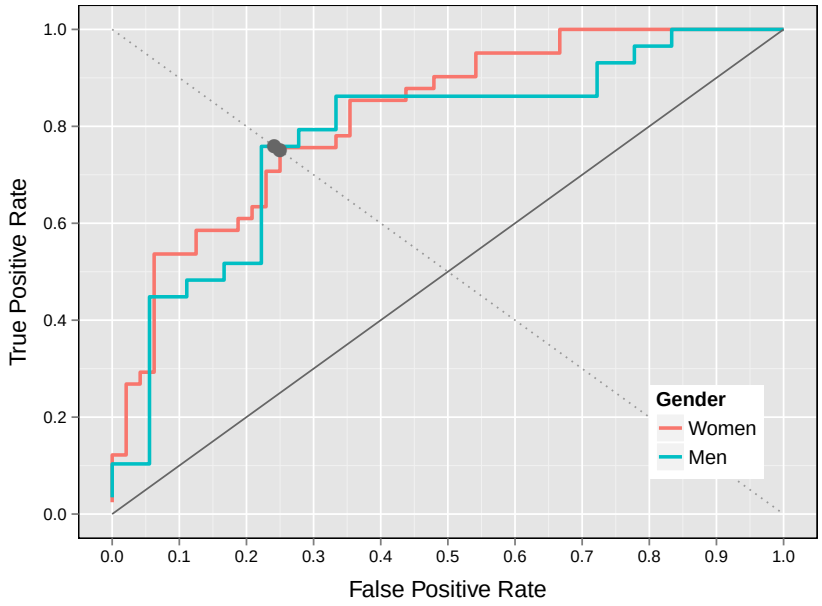

Figure **3-7**: ROC curve from the classification between AD and NC using the EMD as the similarity measure, separately by gender.

The Figure **3-7** illustrates the results for the AD detection step when using the EMD as the distance between RoI, these results are similar to the results obtained using the KL-Divergence (in Figure **3-5**), here we obtained an AUC of 0.82 and an EER of 0.25 for women, and an AUC of 0.78 and an EER of 0.24 for men. The fact that the classification performance is similar when using both measures suggests that the morphological patterns that we found can help to detect AD at early stages and can be captured independently of the dimensionality of the characterization or the similarity measure.

When trying to discriminate between very mild and mild AD using the EMD as the metric between RoIs, the classification performance is also better for men than women, as observed in Figure **3-8** where the blue curve is closer to the top-left corner than the pink. From these curves we obtained an AUC of 0.72 and an EER of 0.36, for women, and an AUC of 0.78 and an EER of 0.27, for men. It is a remarkable fact that, when using both similarity measures, the classification performance between stages of the disease is better for men than for women. In contrast, the classification performance between NC and mild AD is better for women than for men, as shown in Section 3.1.

Table **3-2** summarizes the classification performance, in terms of the AUC and the EER, for all the experiments with the second experimental group.

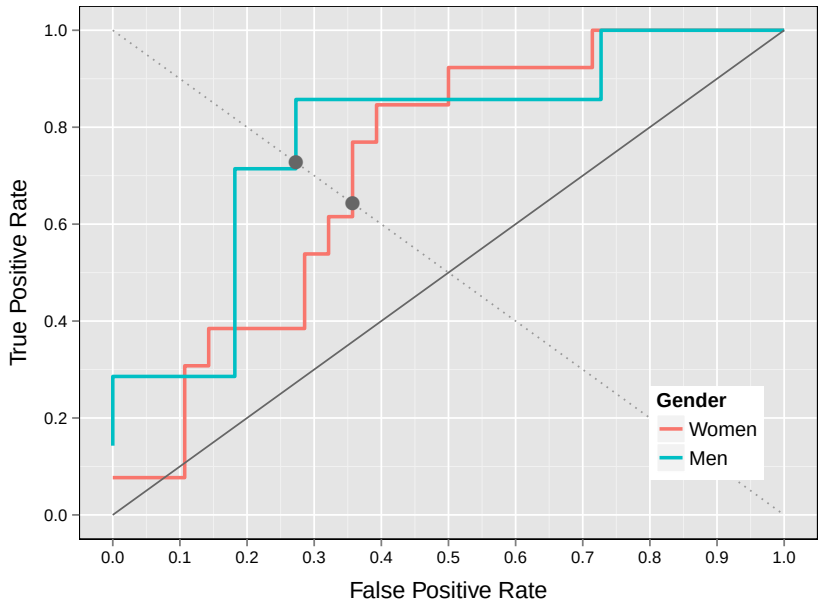

Figure **3-8**: ROC curve from the classification between very mild AD and mild AD using the EMD as the similarity measure, separately by gender.

Table **3-2**: Classification performance for the experiments with the second group. The sensitivity and sensibility are given by 1 − EER.

	Step 1: NC Vs. AD				Step 2: vmAD Vs. mAD			
	Women		Men		Women		Men	
	EER	AUC	EER	AUC	EER	AUC	EER	AUC
Using KL-Div	0.27	0.82	0.28	0.81	0.36	0.73	0.32	0.71
Using EMD	0.25	0.82	0.24	0.78	0.36	0.72	0.27	0.78

4 Discussion

In this work we have presented an automatic strategy to learn brain models for different groups based on morphological features by applying a novel fusion strategy, from these models we obtain discerning regions and then these regions may be used to extract descriptive signatures of the morphological information which it is present in normal and pathological brains. Furthermore, we have shown how it is possible to obtain useful classification features by using two different similarity measures to compare RoI in a test MRI against RoI in the brain models.

By applying the proposed strategy the classification results are an improvement with respect to previous results that use the same group as our first experimental group and the same validation scheme. As shown in Section 3.1, with this group we obtained Equal Error Rates of 0.10 and 0.11 using the KL-Divergence and EMD respectively while Toews *et al.* [33] and Rueda *et al.* [52] obtained EER of 0.20 and 0.14 respectively.

The EER that we obtained means that we can achieve about 90% of sensitivity and specificity at the same time when detecting mild AD. The same task was carried out by Cuignet *et al.* [24] using the ADNI database, but assessing 10 different methods and reporting that the best method achieved around 80% of sensitivity and 95% of specificity. Similar or slightly lower results where found using methods based or relying on segmentation of the brain main tissues, namely cerebrospinal fluid, gray and white matters [26, 34, 35, 36, 37]. Other methods with similar classificaion performances use elastic registration [53], semi-automatic segmentation of the hippocampus [38], or combine measures over the mentioned features [39, 40, 19, 41, 42].

Additionally, when comparing control with mild AD, the experimental groups was assessed by separating by gender since there exist well documented brain anatomical differences between them that might bias the model [54]. When using both metrics, the classification results for women outperform what was observed in the men case, this may be due to the low number of men cases (7 AD and 18 NC), likely a consequence of the fact that women are at a greater risk for developing Alzheimer's disease [55].

When including patients with very mild AD we obtain Equal Error Rates, 0.25 for women and 0.24 for men (see Table **3-2**). This means that we can achieve about 75% of sensitivity and specificity at the same time when detecting AD in women and 76% of sensitivity and

specificity when detecting AD in men. Toews *et al.* [33] and Rueda *et al.* [52] also used this experimental group to classify AD, obtaining EER of 0.29 and 0.24 respectively. Althought they used the same experimental group, they performed a leave-one-out validatio scheme with the whole group while we have performed a leave-one-out validation scheme separately by gender.

4.1 Products

While developing this thesis, three works were presented in international conferences:

- *"Estudio del Tamao del Campo de Anlisis en la Tarea de Clasificacin de IRM en Cerebros con la Enfermedad de Alzheimer"* in VIII International Seminar on Medical Image Processing and Analysis - SIPAIM (2012). [56]
- *"An Automatic MRI Fusion Strategy Based on Nonnegative Matrix Factorization"* in IX International Seminar on Medical Image Processing and Analysis - SIPAIM (2013). [57]
- *"Morphometry-Based Comparison of Relevant Brain Regions for Alzheimer's Disease Detection"* in X International Seminar on Medical Image Processing and Analysis - SIPAIM (2014). [58]

4.2 Future Work

Future work includes integrating prior knowledge to improve the models of the brains, this knowledge could include other neuropsychological tests and demographic data, extensive experimentation in larger databases as the Alzheimer's Disease Neuroimaging Initiative (`http://adni.loni.usc.edu/`), analizing the functional relations between the obtained regions of interest and exploiting the used metrics to explore and compare intermediate stages of the disease.

References

[1] Health and N. Human Services Department, *Alzheimer's Disease: Unraveling the Mystery*. Government Printing Office, 2003.

[2] B. Dubois, H. H. Feldman, C. Jacova, J. L. Cummings, S. T. DeKosky, P. Barberger-Gateau, A. Delacourte, G. Frisoni, N. C. Fox, D. Galasko, S. Gauthier, H. Hampel, G. A. Jicha, K. Meguro, J. O'Brien, F. Pasquier, P. Robert, M. Rossor, S. Salloway, M. Sarazin, L. C. de Souza, Y. Stern, P. J. Visser, and P. Scheltens, "Revising the definition of Alzheimer's disease: a new lexicon," *The Lancet Neurology*, vol. 9, no. 11, pp. 1118 – 1127, 2010.

[3] H. Braak and E. Braak, "Neuropathological stageing of Alzheimer-related changes," *Acta Neuropathologica*, vol. 82, no. 4, pp. 239–259, 1991.

[4] N. Janocko, K. Brodersen, A. Soto-Ortolaza, O. Ross, A. Liesinger, R. Duara, N. Graff-Radford, D. Dickson, and M. Murray, "Neuropathologically defined subtypes of Alzheimers disease differ significantly from neurofibrillary tangle-predominant dementia," *Acta Neuropathologica*, vol. 124, no. 5, pp. 681–692, 2012.

[5] R. Armstrong, D. Nochlin, and T. Bird, "Neuropathological heterogeneity in Alzheimer's disease: a study of 80 cases using principal components analysis," *Neuropathology*, vol. 20, no. 1, pp. 31–37, 2000.

[6] H. Akatsu, M. Takahashi, N. Matsukawa, Y. Ishikawa, N. Kondo, T. Sato, H. Nakazawa, T. Yamada, H. Okada, T. Yamamoto, and K. Kosaka, "Subtype analysis of neuropathologically diagnosed patients in a japanese geriatric hospital," *Journal of the neurological sciences*, vol. 196, pp. 63–69, April 2002.

[7] M. E. Murray, A. Cannon, N. R. Graff-Radford, A. M. Liesinger, N. J. Rutherford, O. A. Ross, R. Duara, M. M. Carrasquillo, R. Rademakers, and D. W. Dickson, "Differential clinicopathologic and genetic features of late-onset amnestic dementias.," *Acta Neuropathol*, 2014.

[8] M. E. Murray, N. R. Graff-Radford, O. A. Ross, R. C. Petersen, R. Duara, and D. W. Dickson, "Neuropathologically defined subtypes of Alzheimer's disease with distinct clinical characteristics: a retrospective study," *The Lancet Neurology*, vol. 10, no. 9, pp. 785–796, 2011.

[9] Alzheimer's Association, "2011 Alzheimer's disease facts and figures," *Alzheimer's & dementia : the journal of the Alzheimer's Association*, vol. 7, p. 208244, March 2011.

[10] A. Association, "2014 Alzheimer's disease facts and figures," *Alzheimer's & Dementia*, vol. 10, no. 2, pp. e47 – e92, 2014.

[11] A. D. International, "Dementia: a public health priority," tech. rep., World Health Organization, 2012.

[12] M. Prince, M. Prina, and M. Guerchet, "World alzheimer report 2013: Journey of caring, an analysis of long-term care for dementia," September 2013.

[13] Alzheimers Disease International, "The global impact of dementia 20132050," tech. rep., World Health Organization, 2013.

[14] GBD 2013 Mortality and Causes of Death Collaborators, "Global, regional, and national agesex specific all-cause and cause-specific mortality for 240 causes of death, 19902013: a systematic analysis for the global burden of disease study 2013," *The Lancet*, vol. 385, p. 117171, January 2015.

[15] G. McKhann, D. Drachman, M. Folstein, R. Katzman, D. Price, and E. M. Stadlan, "Clinical diagnosis of Alzheimer's disease report of the nincds-adrda work group* under the auspices of department of health and human services task force on Alzheimer's disease," *Neurology*, vol. 34, no. 7, pp. 939–944, 1984.

[16] A. P. Association, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition: DSM-IV-TR®*. American Psychiatric Association, 2000.

[17] the Alzheimer's Association and the National Institute on Aging (NIA), "New criteria and guidelines for alzheimer's disease diagnosis," 2010.

[18] B. Dubois, H. H. Feldman, C. Jacova, S. T. DeKosky, P. Barberger-Gateau, J. Cummings, A. Delacourte, D. Galasko, S. Gauthier, G. Jicha, *et al.*, "Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS–ADRDA criteria," *The Lancet Neurology*, vol. 6, no. 8, pp. 734–746, 2007.

[19] 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. Frackowiak, "Automatic classification of MR scans in Alzheimer's disease," *Brain*, vol. 131, no. 3, pp. 681–689, 2008.

[20] J. C. Morris, "The clinical dementia rating (cdr): current version and scoring rules," *Neurology*, vol. 43, no. 11, 1993.

[21] M. F. Folstein, S. E. Folstein, and P. R. McHugh, "mini-mental state: A practical method for grading the cognitive state of patients for the clinician," *Journal of Psychiatric Research*, vol. 12, no. 3, pp. 189 – 198, 1975.

[22] W. Rosen, R. Mohs, and K. Davis, "A new rating scale for Alzheimer's disease," *American Journal of Psychiatry*, vol. 141, no. 11, pp. 1356–1364, 1984. PMID: 6496779.

[23] M. Prince, R. Bryce, and C. Ferri, "World alzheimer report 2011: The benefits of early diagnosis and intervention," September 2011.

[24] R. Cuingnet, E. Gerardin, J. Tessieras, G. Auzias, S. Lehéricy, M.-O. Habert, M. Chupin, H. Benali, and O. Colliot, "Automatic classification of patients with Alzheimer's disease from structural MRI: a comparison of ten methods using the ADNI database," *Neuroimage*, vol. 56, no. 2, pp. 766–781, 2011.

[25] E. E. Bron, M. Smits, W. M. van der Flier, H. Vrenken, F. Barkhof, P. Scheltens, J. M. Papma, R. M. Steketee, C. M. Orellana, R. Meijboom, M. Pinto, J. R. Meireles, C. Garrett, A. J. Bastos-Leite, A. Abdulkadir, O. Ronneberger, N. Amoroso, R. Bellotti, D. Crdenas-Pea, A. M. lvarez Meza, C. V. Dolph, K. M. Iftekharuddin, S. F. Eskildsen, P. Coup, V. S. Fonov, K. Franke, C. Gaser, C. Ledig, R. Guerrero, T. Tong, K. R. Gray, E. Moradi, J. Tohka, A. Routier, S. Durrleman, A. Sarica, G. D. Fatta, F. Sensi, A. Chincarini, G. M. Smith, Z. V. Stoyanov, L. Srensen, M. Nielsen, S. Tangaro, P. Inglese, C. Wachinger, M. Reuter, J. C. van Swieten, W. J. Niessen, and S. Klein, "Standardized evaluation of algorithms for computer-aided diagnosis of dementia based on structural MRI: The CADDementia challenge," *NeuroImage*, vol. 111, no. 0, pp. 562 – 579, 2015.

[26] C. Davatzikos, Y. Fan, X. Wu, D. Shen, and S. M. Resnick, "Detection of prodromal Alzheimer's disease via pattern classification of magnetic resonance imaging," *Neurobiology of aging*, vol. 29, no. 4, pp. 514–523, 2008.

[27] A. van der Kouwe, T. Benner, D. Salat, and B. Fischl, "Brain morphometry with multiecho MPRAGE," *NeuroImage*, vol. 40, no. 2, pp. 559–569, 2008.

[28] D. Mietchen and C. Gaser, "Computational morphometry for detecting changes in brain structure due to development, aging, learning, disease and evolution," *Frontiers in Neuroinformatics*, vol. 3, no. 25, 2009.

[29] J. R. DeQuardo, M. S. Keshavan, F. L. Bookstein, W. W. Bagwell, W. D. K. Green, J. A. Sweeney, G. L. Haas, R. Tandon, N. R. Schooler, and P. J. W., "Landmark-based morphometric analysis of first-episode schizophrenia," *Society of Biological Psychiatry*, vol. 45, pp. 1321–1328, 1999.

[30] J. Ashburner and K. Friston, "Voxel-based morphometry – The methods," *NeuroImage*, vol. 11, pp. 805–821, 2000.

[31] J. Ashburner, C. Hutton, R. Frackowiak, I. Johnsrude, C. Price, and K. Friston, "Identifying global anatomical differences: Deformation-based morphometry," *Human Brain Mapping*, vol. 6, no. 5, pp. 348–357, 1998.

[32] D. Pantazis, R. M. Leahy, T. E. Nichols, and M. Styner, "Statistical surface-based morphometry using a non-parametric approach," in *ISBI*, pp. 1283–1286, 2004.

[33] M. Toews, W. M. W. III, D. L. Collins, and T. Arbel, "Feature-based morphometry: Discovering group-related anatomical patterns," *NeuroImage*, vol. 49, no. 3, pp. 2318–2327, 2010.

[34] D. Zhang, Y. Wang, L. Zhou, H. Yuan, and D. Shen, "Multimodal classification of alzheimer's disease and mild cognitive impairment," *Neuroimage*, vol. 55, no. 3, pp. 856–867, 2011.

[35] P. Vemuri, J. L. Gunter, M. L. Senjem, J. L. Whitwell, K. Kantarci, D. S. Knopman, B. F. Boeve, R. C. Petersen, and C. R. J. Jr., "Alzheimer's disease diagnosis in individual subjects using structural MR images: Validation studies," *NeuroImage*, vol. 39, no. 3, pp. 1186 – 1197, 2008.

[36] E. Westman, C. Aguilar, J.-S. Muehlboeck, and A. Simmons, "Regional magnetic resonance imaging measures for multivariate analysis in Alzheimers disease and mild cognitive impairment," *Brain Topography*, vol. 26, no. 1, pp. 9–23, 2013.

[37] Y. Fan, S. M. Resnick, X. Wu, and C. Davatzikos, "Structural and functional biomarkers of prodromal Alzheimer's disease: a high-dimensional pattern classification study," *Neuroimage*, vol. 41, no. 2, pp. 277–285, 2008.

[38] J. Barnes, R. I. Scahill, R. G. Boyes, C. Frost, E. B. Lewis, C. L. Rossor, M. N. Rossor, and N. C. Fox, "Differentiating AD from aging using semiautomated measurement of hippocampal atrophy rates," *NeuroImage*, vol. 23, no. 2, pp. 574 – 581, 2004.

[39] S. Farhan, M. A. Fahiem, and H. Tauseef, "An ensemble-of-classifiers based approach for early diagnosis of Alzheimers disease: Classification using structural features of brain images," *Computational and Mathematical Methods in Medicine*, vol. 2014, 2014.

[40] C. Plant, S. J. Teipel, A. Oswald, C. Bhm, T. Meindl, J. Mourao-Miranda, A. W. Bokde, H. Hampel, and M. Ewers, "Automated detection of brain atrophy patterns based on MRI for the prediction of Alzheimer's disease," *NeuroImage*, vol. 50, no. 1, pp. 162 – 174, 2010.

[41] S. J. Teipel, C. Born, M. Ewers, A. L. Bokde, M. F. Reiser, H.-J. Mller, and H. Hampel, "Multivariate deformation-based analysis of brain atrophy to predict Alzheimer's disease in mild cognitive impairment," *NeuroImage*, vol. 38, no. 1, pp. 13 – 24, 2007.

[42] R. Wolz, V. Julkunen, J. Koikkalainen, E. Niskanen, D. P. Zhang, D. Rueckert, H. Soininen, J. Lötjönen, *et al.*, "Multi-method analysis of MRI images in early diagnostics of alzheimer's disease," *PloS one*, vol. 6, no. 10, p. e25446, 2011.

[43] P. Padilla, M. López, J. M. Górriz, J. Ramirez, D. Salas-Gonzalez, and I. Álvarez, "NMF-SVM based CAD tool applied to functional brain images for the diagnosis of Alzheimer's disease," *Medical Imaging, IEEE Transactions on*, vol. 31, no. 2, pp. 207–216, 2012.

[44] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, pp. 788 – 791, 1999.

[45] Y. Rubner, C. Tomasi, and L. J. Guibas, "The earth mover's distance as a metric for image retrieval," *International Journal of Computer Vision*, vol. 40, pp. 99–121, Nov. 2000.

[46] F. L. Hitchcock, "The distribution of a product from several sources to numerous localities," *J. Math. Phys*, vol. 20, no. 2, pp. 224–230, 1941.

[47] D. S. Marcus, T. H. Wang, J. Parker, J. G. Csernansky, J. C. Morris, and R. L. Buckner, "Open access series of imaging studies (OASIS): Cross-sectional MRI data in young, middle aged, nondemented, and demented older adults," *J. Cognitive Neuroscience*, vol. 19, no. 9, pp. 1498–1507, 2007.

[48] J. Talairach and P. Tournoux, *Co-Planar Stereotaxic Atlas of the Human Brain: 3-D Proportional System: An Approach to Cerebral Imaging (Thieme Classics)*. Thieme, 1988.

[49] M. Jenkinson and S. Smith, "A global optimisation method for robust affine registration of brain images," *Medical image analysis*, vol. 5, no. 2, pp. 143–156, 2001.

[50] M. Jenkinson, P. Bannister, M. Brady, and S. Smith, "Improved optimization for the robust and accurate linear registration and motion correction of brain images," *Neuroimage*, vol. 17, no. 2, pp. 825–841, 2002.

[51] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, pp. 27:1–27:27, 2011.

[52] A. Rueda, F. Gonzlez, and E. Romero, "Extracting salient brain patterns for imaging-based classification of neurodegenerative diseases," *Medical Imaging, IEEE Transactions on*, vol. 33, pp. 1262–1274, June 2014.

[53] B. Magnin, L. Mesrob, S. Kinkingnhun, M. Plgrini-Issac, O. Colliot, M. Sarazin, B. Dubois, S. Lehricy, and H. Benali, "Support vector machine-based classification of

Alzheimers disease from whole-brain anatomical MRI," *Neuroradiology*, vol. 51, no. 2, pp. 73–83, 2009.

[54] S. F. Witelson and D. L. Kigar, "Sylvian fissure morphology and asymmetry in men and women: Bilateral differences in relation to handedness in men," *Journal of Comparative Neurology*, vol. 323, no. 3, pp. 326–340, 1992.

[55] M. M. Mielke, P. Vemuri, and W. A. Rocca, "Clinical epidemiology of alzheimers disease: assessing sex and gender differences," *Clinical Epidemiology*, vol. 6, pp. 37–48, January 2014.

[56] D. L. Giraldo, R. Gutiérrez, and E. Romero, "Estudio del tamaño del campo de análisis en la tarea de clasificación de IRM en cerebros con la enfermedad de Alzheimer," in *8th International Seminar on Medical Image Processing and Analysis-SIPAIM*, (San Cristbal, Venezuela), 2012.

[57] D. L. Giraldo, J. D. García-Arteaga, and E. Romero, "An Automatic Search of Alzheimer Patterns Using a Nonnegative Matrix factorization.," in *9th International Seminar on Medical Image Processing and Analysis-SIPAIM*, (Ciudad de México, México), 2013.

[58] D. L. Giraldo, J. D. García-Arteaga, and E. Romero, "Morphometry-based comparison of relevant brain regions for Alzheimer's disease detection," in *10th International Seminar on Medical Image Processing and Analysis-SIPAIM*, (Cartagena de Indias, Colombia), 2014.