



Regions of interest computed by SVM wrapped method for Alzheimer's disease examination from segmented MRI

Antonio R. Hidalgo-Muñoz*, Javier Ramírez, Juan M. Górriz and Pablo Padilla

Department of Signal Theory, Networking and Communications, University of Granada, Granada, Spain

Edited by:

Hari S. Sharma, Uppsala University, Sweden

Reviewed by:

Hari S. Sharma, Uppsala University, Sweden

John Suckling, University of Cambridge, UK

Manuel Grana, Universidad del Pais Vasco, Spain

*Correspondence:

Antonio R. Hidalgo-Muñoz,
Department of Signal Theory,
Networking and Communications,
University of Granada, C/ Periodista
Daniel Saucedo Aranda s/n, 18071
Granada, Spain
e-mail: arhidalgom@gmail.com

Accurate identification of the most relevant brain regions linked to Alzheimer's disease (AD) is crucial in order to improve diagnosis techniques and to better understand this neurodegenerative process. For this purpose, statistical classification is suitable. In this work, a novel method based on support vector machine recursive feature elimination (SVM-RFE) is proposed to be applied on segmented brain MRI for detecting the most discriminant AD regions of interest (ROIs). The analyses are performed both on gray and white matter tissues, achieving up to 100% accuracy after classification and outperforming the results obtained by the standard *t*-test feature selection. The present method, applied on different subject sets, permits automatically determining high-resolution areas surrounding the hippocampal area without needing to divide the brain images according to any common template.

Keywords: Alzheimer's disease, gray and white matter, image segmentation, MRI, SVM

INTRODUCTION

Alzheimer's disease (AD) is a progressive, neurodegenerative disorder that gradually impairs memory and other cognitive skills, such as spatial orientation, judgment or language, preventing a healthy aging. Currently, scientists are interested in researching different kinds of brain imaging to detect possible dementia at a very early stage, when medical and psychological treatments are more effective. For Alzheimer's disease, it is necessary to determine the brain regions of interest (ROIs); magnetic resonance imaging (MRI) provides valuable information on that matter. Several useful atlases and templates are reported in the literature (Shen et al., 2012). However, choosing the most appropriate one is difficult and it is not possible to use only one brain map to cover every specific characteristic from each neurological disorder. From a clinical standpoint, doctors generally require rigorous technical basis or mathematical methods (Rosenberg and Hillis, 2009). Therefore, a straightforward method is needed for discriminating the most relevant regions related to neurodegenerative diseases (López et al., 2011; Shen et al., 2011). In this study, an approach combining image segmentation and a wrapped classification algorithm is proposed for computing the ROIs from MRI, which differ meaningfully between AD patients and healthy elderly people.

Structural MRI has been widely explored in AD, giving valuable information about its underlying anatomical progression. For instance, the Alzheimer's disease neuroimaging initiative (ADNI) has compiled an MRI database of AD, and mild cognitive impairment (MCI) that has been widely employed (Jack et al., 2010; Weiner et al., 2010). MRI and AD research goals vary. Several researchers in this field have focused on volume and integrity measurement of different brain tissues (Shen et al., 2011; Zhang et al.,

2013) in order to find relevant early AD biomarkers or MCI onset (Desikan et al., 2009). Furthermore, much research attempts to achieve more efficient automated classification of AD patients as compared to MCI patients or healthy aging (Klöppel et al., 2008; Aguilar et al., 2013; Ortiz et al., 2013a). Overall, most of these studies determine the ROIs before the classification block according to previously delimited regions, which have been split based on any standard template (Mazziotta et al., 1995; Shattuck et al., 2008; Cuingnet et al., 2011).

Determining adequate brain ROIs is an important topic in medical image processing and computer-aided diagnostics (CAD) with many applications like morphology detection (Kapur et al., 1996) or 3-D visualizations for surgical planning (Suetens et al., 1993; Clarke et al., 1995). Usually, MRI is segmented into three quantitatively distinct tissues, that is, gray matter (GM), mainly linked to the cortex, white matter (WM), mainly composed by neuronal axons, and cerebrospinal fluid (CSF; Salas-Gonzalez et al., 2011). Despite the recent development of new MRI segmentation methods (Ortiz et al., 2013b; Salas-Gonzalez et al., 2013), the most recognized approach models intensity value distribution by a mixture of Gaussian distributions (MOG; Ashburner and Friston, 2003). This method is implemented in commonly used software like statistical parametric mapping (SPM; Frackowiak, 2004), which is utilized in diverse clinical protocols, and obtains the probability of each image voxel belonging to any tissue according to location and intensity level in a gray scale.

In general, machine-learning and classification techniques are increasingly used as an alternative to other multivariate statistical approaches. The aim of these pattern recognition techniques is not limited to achieving good results in classification tasks

or for artificial intelligence applications, but rather to gage the relevance of some extracted features and search for differences between experimental conditions (Shieh and Yang, 2008; Hidalgo-Muñoz et al., 2013a,b; Tomé et al., 2013). Following this line, methods where the feature selection algorithm is wrapped around the classification algorithm recursively to identify the least relevant features constitutes a good option (Kohavi and John, 1997). These algorithms are suitable for dealing with high-dimensional data like medical images, since the parameters of the classifier serve as scores to select the ROIs and the corresponding classification performance guides the iterative procedure. On the other hand, generalizing results will depend on the size of the dataset, and the cross-validation (CV) method used to evaluate the classification accuracy (Burgess, 1998). Either way, these methods permit covering a whole set of initial features without being restricted to any specific region to check its relevance. In this work, the recursive feature elimination (RFE) algorithm, proposed by Guyon et al. (2002), and based on the support vector machine (SVM; Ben-Hur et al., 2008), is used. Support vector machine recursive feature elimination (SVM-RFE) has been successfully implemented in various neuroscience applications (De Martino et al., 2008; Chu et al., 2012; Hidalgo-Muñoz et al., 2013a); nevertheless, it has hardly been used for image analyses.

The presented work focuses separately on GM and WM tissues to delimit the most discriminant brain ROIs for examining AD from MRI. This paper presents an innovative and effective method for feature selection, the SVM-RFE technique, that has never been used before for this purpose as far as authors are aware. This affordable and intuitive method, easily implementable in medical apparatuses, intends to contribute to a complete diagnosis and examination of AD and its progression.

MATERIALS AND METHODS

DATASET

Data used in the preparation of this article was obtained from the (ADNI) database (<http://adni.loni.usc.edu/>). The ADNI was launched in 2003 by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies and non-profit organizations, as a \$60 million, 5-year public-private partnership. The primary goal of ADNI has been to test whether serial MRI, positron emission tomography (PET), other biological markers, and the progression of MCI, and early AD. Determining sensitive and specific markers of very early AD progression is intended to aid researchers and clinicians to develop new treatments, as well as reduce the time and cost of clinical trials. The Principal Investigator of this initiative is Michael W. Weiner, MD, VA Medical Center and University of California, San Francisco. ADNI is the result of efforts of many co-investigators from a broad range of academic institutions and private corporations, and subjects have been recruited from over 50 sites across the U.S. and Canada. The initial goal of ADNI was to recruit 800 adults, ages 55 to 90, to participate in the research: approximately 200 cognitively normal older individuals to be followed for three years, 400 people with MCI to be followed for three

years and 200 people with early AD to be followed for two years. For up-to-date information, see www.adni-info.org.

In this article, only the data from T1-weighted MR images was considered. The participants were separated into two different classes:

- *Normal*. Control subjects. Clinical Dementia Rating (CDR; Morris, 1993) of zero. They were non-depressed, non-MCI and non-demented.
- *AD*. CDR of 0.5 or 1, met NINCDS/ADRDA criteria for probable AD (McKhann et al., 2011).

Table 1 shows the demographic details of the subjects who compose the dataset used in this work.

IMAGE PRE-PROCESSING

The SPM software was originally designed for analyzing functional brain images. The package also contains routines for realignment, smoothing, and spatial normalization into a standard space of T1-weighted MR images. To this end, the template implemented within the VBM8 Toolbox was used, specifically DARTEL, to achieve an accurate realignment of the images and a good normalization (<http://dbm.neuro.uni-jena.de/vbm/>). It is worthwhile to stress that spatial normalization or the registration algorithm is always a critical component to any classifier that uses voxel-wise features (Cuadra et al., 2005). Within these routines, a modulation step was implemented in order to conserve the amount of tissue and not the intensities (see Ashburner et al., 2012; p. 192). After the transformation of the images from the ADNI database, they were resized to the dimensions $121 \times 145 \times 121$ with voxel sizes of 1.5 mm (sagittal) \times 1.5 mm (coronal) \times 1.5 mm (axial). Neither smoothing nor dimension reduction were performed after that. One high-quality image per subject was chosen.

IMAGE SEGMENTATION

In this work, GM and WM image segmentation was implemented in SPM, which models the intensity value distribution of the T1-weighted MRI by a MOG (Ashburner and Friston, 2003, 2005) and takes voxel location into consideration via a tissue probability map (TPM). Using this methodology, which is described, for example, in Ashburner et al. (2012), it has been possible to overcome the partial volume effect (PVE), such that a voxel may not be purely of one tissue class, but can contain signals from a number of different tissues.

Within this work the central features are the probability values for GM or WM in a given voxel, not the intensity values *per se*. Once the images have been segmented, the resulting dataset is ready for further processing and analysis. The subsequent analyses were carried out with MATLAB™.

SVM WRAPPED METHOD

As mentioned, medical images may provide clinicians valuable information about disease status, diagnosis and prognosis. However, extracting significant features from high-dimensional data as an image is always a complex task. Usually, some reduction methods, and their subsequent feature selection, greatly transform the original data, making eventual clinical interpretations difficult. The aim of this paper is to select a reasonable number of

ROIs with high predictive value and whose aspect could then be medically meaningful.

In the wrapper approach, the feature selection algorithm wraps around the classification algorithm. The feature selection consists of searching high-dimensional data sets using the induction algorithm itself as part of the evaluating function (Kohavi and John, 1997). Hence the parameters of the classifier serve as scores to select (or to eliminate) features; the consequent classification performance guides an iterative procedure. When this recursive feature elimination strategy uses a linear SVM-based classifier, the resulting method is known as support vector machine-recursive feature elimination (SVM-RFE; Guyon et al., 2002). The validation method will be crucial to avoiding creating a system that is over-trained, that is, that fits well only to the experimental data, losing generality and providing misleading results.

An SVM-based classifier separates a given set of binary-labeled training data with what is known as the maximal margin hyperplane, which is maximally distant from two classes (for example, *AD* and *Normal* classes). The objective is to build a function that will correctly classify new examples (for example, *MRI-segmented images*).

Linear SVM parameters define a decision hyperplane in the multidimensional feature space (Burges, 1998; Ben-Hur et al., 2008; Illán et al., 2011), that is:

$$g(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + b = 0$$

where \mathbf{x} denotes the feature vector, \mathbf{w} is known as the weight vector and b is known as the threshold. The decision hyperplane position is determined by vector \mathbf{w} and b : the vector is orthogonal to the decision plane and b determines its distance to the origin. For linear SVM, the vector \mathbf{w} can be explicitly computed. The design of the classifier consists of finding the unknown parameters, that is, w components of \mathbf{w} (w_n , $n = 1 \dots M$, where $M =$ number of features) and b , which allows building a hyperplane that separates the two classes optimally.

Figure 1 illustrates a 2D toy-example of a binary classification problem, where the points $\mathbf{x}=[x_1 \ x_2]$, marked like red circles, belong to one class, and the ones marked like blue crosses belong to the other one. The problem is not linearly separable since it is not possible to find a line (2D hyperplane) that perfectly separates all training instances of the two classes. However, if a small number of misclassifications are tolerated, the problem becomes linearly separable. The figure shows the result of three training sessions with the same data, but different misclassification margins (C parameter). The vector $\mathbf{w} = \sum_1^{N_s} \gamma_i \lambda_i \mathbf{x}_i$ is a weighted sum of the *support vectors* which are the N_s elements, inside the margin, chosen from the set used during the training phase. In **Figure 1**, these support vectors are marked with circles around the training data points. λ_i are the corresponding Lagrangian parameters which are also optimized ($0 < \lambda_i < C$). Finally, the value of the threshold b is estimated by solving the equations related to the hyperplanes that define the margin. In Ben-Hur et al. (2008), an extensive algebraic explanation of SVM applied to biological sciences is reported.

The value of C must be assigned to run the optimization algorithm and represents the weight of the penalty term of the function that is related with the training set misclassification error. It is a

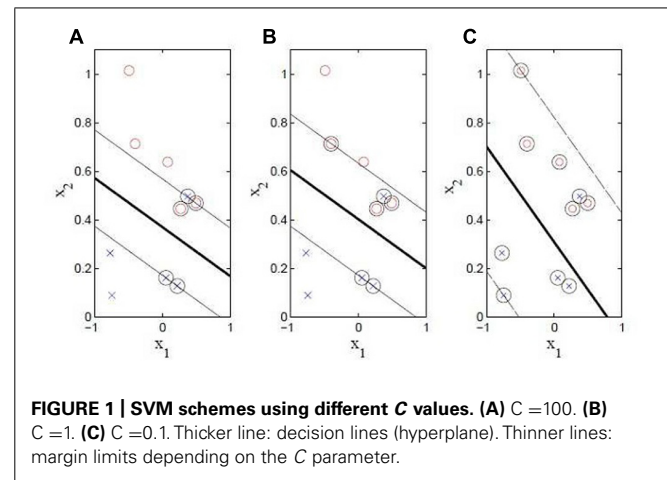


FIGURE 1 | SVM schemes using different C values. (A) $C = 100$. (B) $C = 1$. (C) $C = 0.1$. Thicker line: decision lines (hyperplane). Thinner lines: margin limits depending on the C parameter.

parameter that indirectly controls the margin width of the classifier (see **Figure 1**). A trade-off exists between the width of the margin and the number of accepted misclassifications. There is no optimal procedure to assign this parameter, but it has to be expected that if C is large, the misclassification errors are relevant during the optimization function and the margin should be narrow. On the other hand, if C is small, the misclassification errors are not relevant and a large margin has is expected.

According to SVM-RFE algorithm, the relevance of the feature vector's n -th entry is determined by the corresponding value w_n in the weight vector. In particular, if $|w_n| \simeq 0$, the corresponding feature does not contribute significantly for the value of $g(\mathbf{x})$. Then, sorting these absolute values, the relevance of the features is determined. Therefore, in each loop of the algorithm, a concrete number of features (τ) can be discarded following this criterion (see **Figure 2**).

The method explained above is closely linked to the principle of the nested CV approach for classifier optimization (Varma and Simon, 2006; Gallix et al., 2012). The nested CV techniques use loops into the training blocks for designing and tuning the classifier's parameters (see **Figure 2**). Therefore, following the recommendations reported in the literature (Varma and Simon, 2006), validating the complete algorithm was carried out on different subsets from the initial whole sample in order to diminish biased results (see section 2.5).

As described in the previous section, the feature values after segmentation represent the probability of belonging to the tissue (GM or WM) of every voxel of the image. Therefore, the initial set of features, using a voxel-as-feature (VAF) approach, is composed of $121 \times 145 \times 121 = 2,122,945$ voxels from each original image. Then, the number of features is reduced for each classification task, since only those from the separated segments were used as algorithm inputs. As a result, the number of initial features for the GM class ($M = 568,273$ voxels) was slightly higher than for the WM class ($M = 504,329$ voxels).

Assuming the trade-off between resolution and computational cost, after empirical trials, a step (τ) of 20,000 features to be eliminated in each SVM-RFE loop (approximately 3.5–4% of the total number of features) was considered as a proper option in the experiments.

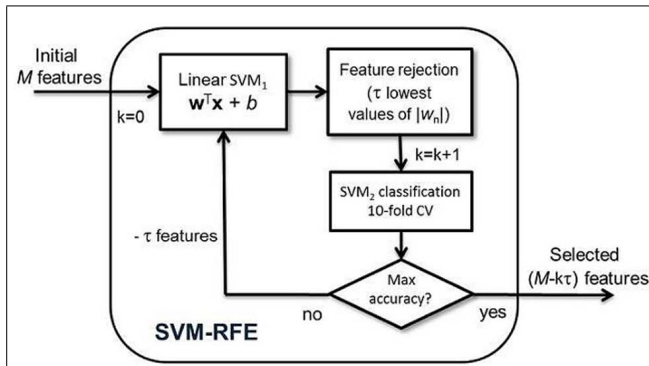


FIGURE 2 | SVM-RFE scheme. SVM₁ parameters are computed taking all the available samples for the training block and determining the relevance of the features. SVM₂ classification is performed in order to check the suitability of the previous selected features by SVM₁ and serves as a guide to achieve the optimal number of features. CV: Cross-validation.

PERFORMANCE MEASUREMENT

As indicated in **Table 1**, a total of 370 subjects (185 *AD* and 185 *Normal*) were employed in the experiment. However, in order to reduce the computational cost and get a more reliable generalization, several subsets were formed from the original whole set to carry out selection and classification (both GM and WM alternatives). The sets were composed of 60 subjects (30 *AD* and 30 *Normal*).

The subjects were organized randomly and the first 30 subjects belonging to each group (*AD* or *Normal*) were taken as the first set. Then, a “sliding window” with an overlap of 25 subjects for each group went across all the subjects, forming new groups. Thus, a total of 32 different sets were constituted to carry out the mentioned selection and classification tasks. Finally, the global accuracy of the method was computed by averaging the partial accuracies obtained using each set. Additionally, by applying SVM-RFE on different sets, the feature selection permits quantifying more exhaustively the extent to which the results are generalizable. Obviously, there were subjects that belonged to several groups and some redundancy was unavoidable. Once the main features were selected by the first blocks of SVM-RFE, an independent SVM classification task was performed using 10-fold as a CV strategy for evaluating the classifier accuracy (see SVM₂ in **Figure 2**), which is a suitable method for diminishing image variability or peculiarity influence and has been strongly suggested in machine-learning applications (Kohavi, 1995; Ortiz et al., 2013a). The 10-fold CV

Table 1 | Sociodemographic data.

Group	Subjects	Sex: M/F	Age: μ (SD)	MMSE: μ (SD)
Normal	185	95/90	75.85(5.11)	29.15(0.97)
AD	185	98/87	75.39(7.56)	23.28(2.05)
<i>p</i>	—	—	0.489	<0.001

MMSE: mini-mental state examination (Folstein et al., 1975). μ : mean; SD: standard deviation; *p*: significance of *t*-test contrast between-groups.

consists of using all the samples in each subset for training the system except 10, which is used as a test. This procedure is repeated *S* times, *S* being the number of 10-sample groups in the dataset, after which a global accuracy value is computed.

STATISTICAL ANALYSIS

In order to check the statistical significance of the differences between voxels belonging to the averaged images from the *AD* and *Normal* groups, Student's *t*-tests for independent samples were performed, assuming a significance level of $p = 0.05$, which is widely used in brain image processing (Du et al., 2001; Liu et al., 2010; Hölzel et al., 2011; Li et al., 2011).

Moreover, a study on the different sets' global accuracy was carried out, computing the mean and standard deviation (SD) for each loop of the algorithm.

RESULTS

REGIONAL DIFFERENCES IN GM

Figure 3 shows the differences in representative axial sections from GM segmented images between *AD* and *Normal* groups in order to get a rough idea about the location of the most evident ROIs. Only GM segmentation is represented, which is emphasized in this kind of research, since it is only for visualization purposes.

ROIs ACCORDING TO *t*-TEST (GM AND WM)

Using the significant features determined by the *t*-test threshold, the obtained accuracy applying the same procedure yielded 78.66% for GM and 77.8% for WM. Applying the *t*-test on the complete sample (a total of 185 subjects for each group to compute the contrast and performing only one global classification task after the feature selection, 10-fold validation), the accuracy improved to 89.46% for GM and 93.24% for WM, although these values are still lower than those obtained by SVM-RFE, as described in the next section. Furthermore, if only the top 20% of the relevant features according to the *t*-test (and corresponding *p* values) was employed, the accuracy decreased to 63.98% and 74.59% for GM, applying the same CV method explained in section 2.5 (32 sets) and using the complete sample, respectively, and 77.85% and 93.24% for WM. Note that in the case of WM, the results are similar taken specifically the significant features or only the 20% because of the number of significant features ($p < 0.05$) matched up approximately to the 20% of the total features.

ROIs ACCORDING TO SVM-RFE (GM AND WM)

Figure 4 shows the most relevant ROIs (GM and WM segmented images) according to the remaining features after applying SVM-RFE using multiple groups for the algorithm training (see section 2.5). According to the accuracy values obtained in both classification tasks, and intending to get a representative amount of features to make illustrative comparisons, only the top 20% most-relevant features pointed by SVM-RFE were chosen as ROIs.

The number of features considered relevant by diverse number of sets is represented in **Figure 5**. In **Figure 5**, two binomial probability density functions (PDFs) are represented as well, whose parameters are $n = 32$ (number of sets) and $p = 0.2$ (success probability being considered as relevant). If the feature selection by every set had been random and independent of the rest

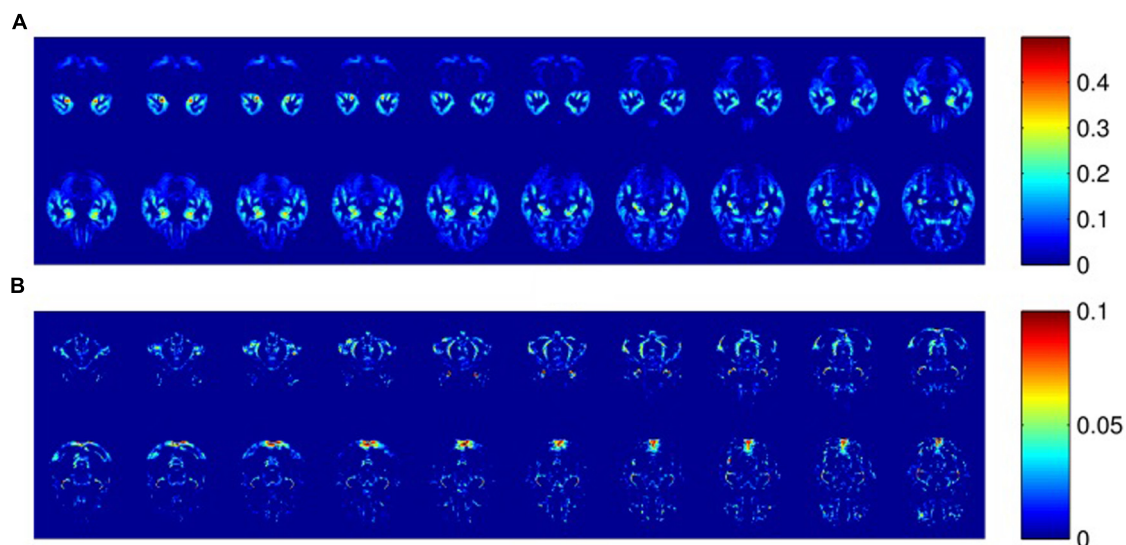


FIGURE 3 | Averaged differences in the probabilities of belonging to GM between AD and Normal groups. Horizontal sections (z axis, $23 < z < 42$). Color bars represent these differences. **(A)** ROIs where the GM probability is higher in Normal than AD (maximal difference = 0.5). **(B)** ROIs where the GM probability is higher in AD than Normal (maximal difference = 0.1).

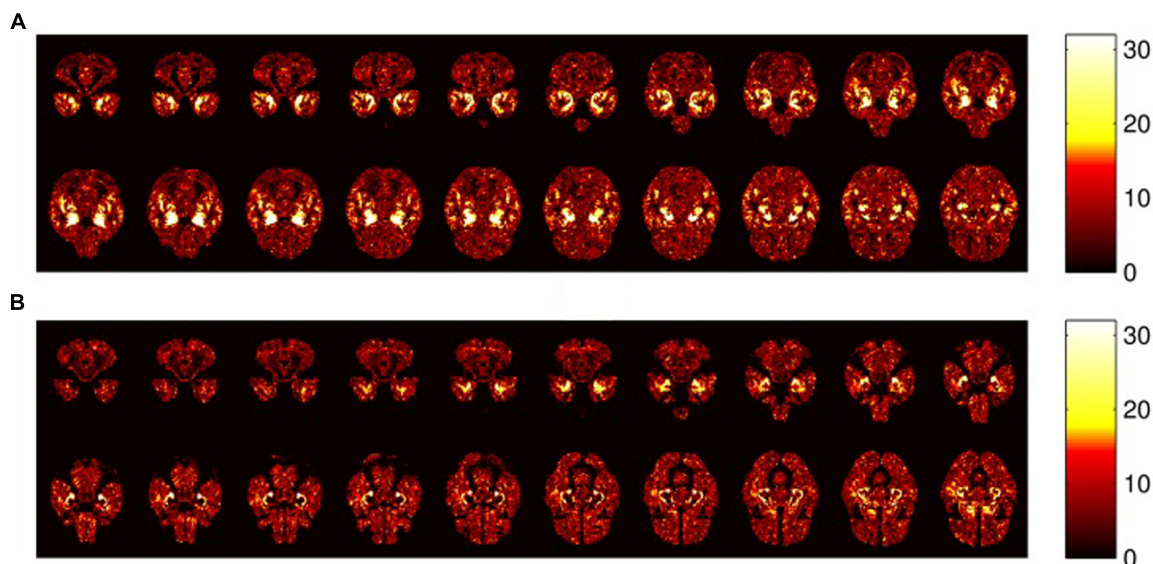


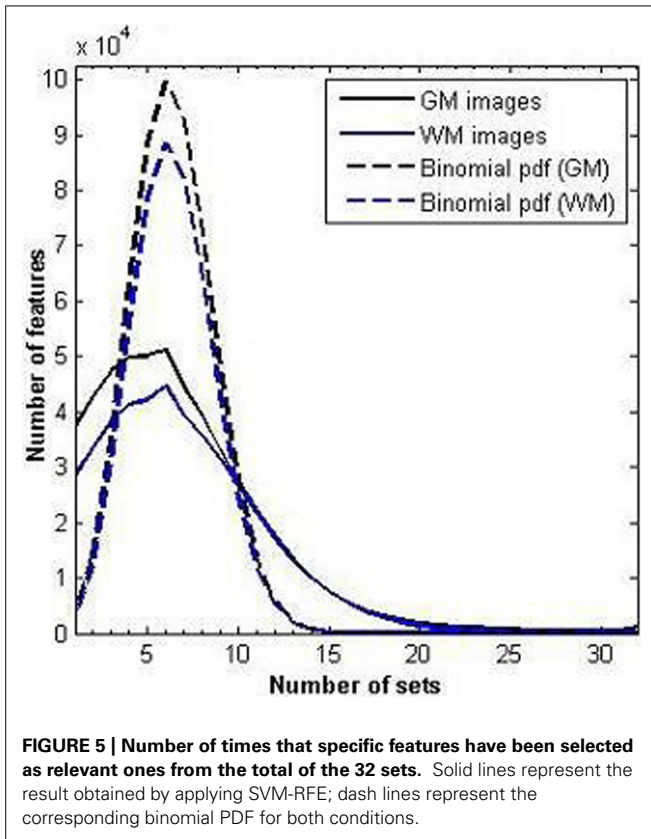
FIGURE 4 | Representative axial sections (z axis, $23 < z < 42$) where relevant differences have been found that delimit ROIs by SVM-RFE. Only the 20% most-relevant features are taken into account. Color bars represent the number of training sets that select each feature. **(A)** GM segmentation. **(B)** WM segmentation. Note that the most brilliant regions are the most relevant.

of the sets' decisions and the features had the same probability of being chosen and were statistically independent as well, these binomial PDF shapes would have been expected in the graphs.

Note that SVM-RFE leads to a substantial reduction in the number of relevant features. Either way, if the target of this work had been to design a potential classifier, instead of computing ROIs, other external data should have been tested.

Figure 6 shows the accuracy values and standard deviations of the AD vs. Normal classification tasks, selecting the relevant features based on SVM-RFE. The maximum averaged accuracy values were 99.64% using GM images and 99.74% using WM images.

Since the objective of the method is to achieve maximum accuracy (practically 100%) in order to discriminate the ROIs, it is not necessary to include either sensitive or specificity measures,



in contrast to other applications whose goals focus on designing the best classifiers and clinical repercussions of their errors.

In order to check again the method's generalization capability, a global classification task was performed after the previous feature selection by SVM-RFE. In this classification task, different amounts of features were chosen based on different required percentages of sets that indicated features as belonging to the 20%

most-relevant ones. In other words, the features were grouped depending on how many times they were selected by all the sets. In this case, the complete sample was considered (185 subjects per group) and 10-fold validation method applied. The results suggested an acceptable generalization, since 94.32% accuracy was reached for GM images taking into consideration only those features selected as relevant by at least 18 of the 32 sets (56% of the sets). However, the accuracy value decreases to 82.16% if only those features selected by 100% of the sets are chosen because too few features are reckoned in that case. Similar results occur with WM segmented images, where 95.14% accuracy is achieved when the features selected by at least 14 sets (44% of the sets) are chosen. However, it is important to note that, in this case, the 10-fold validation method is not a strict CV, since the previous ROIs were selected using the complete sample. Either way, the results can give an idea about potential reproducibility.

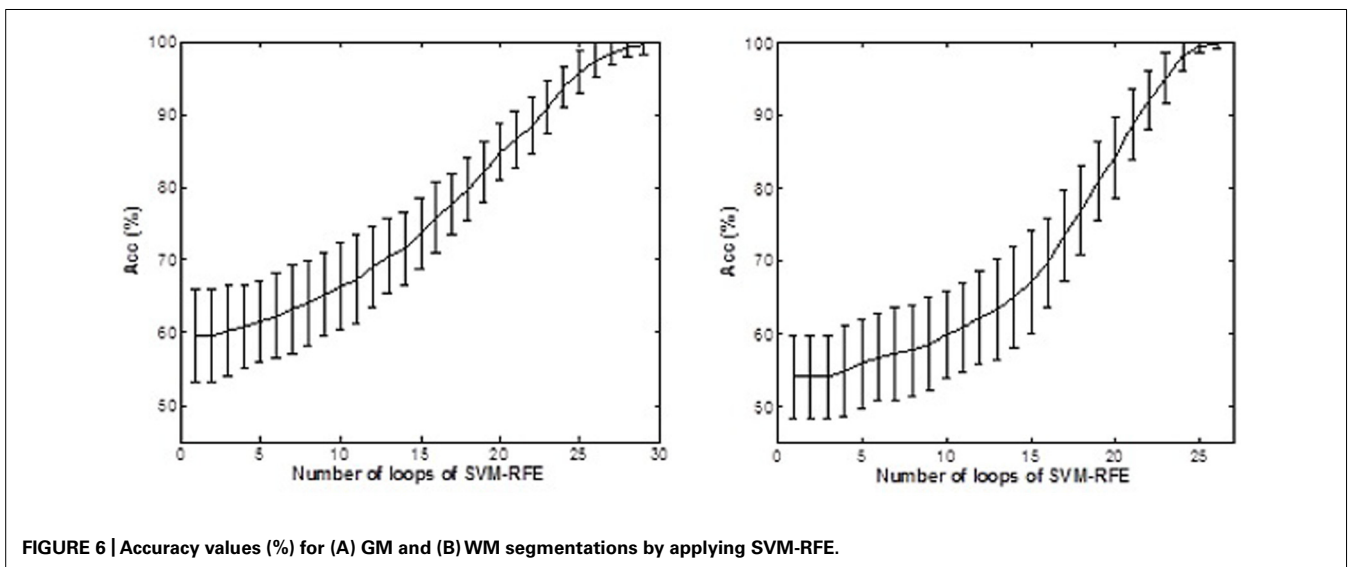
3-D ILLUSTRATIONS

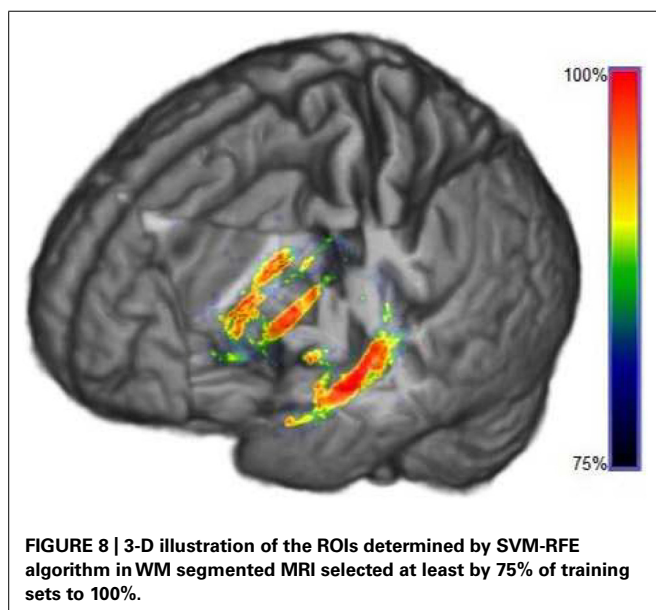
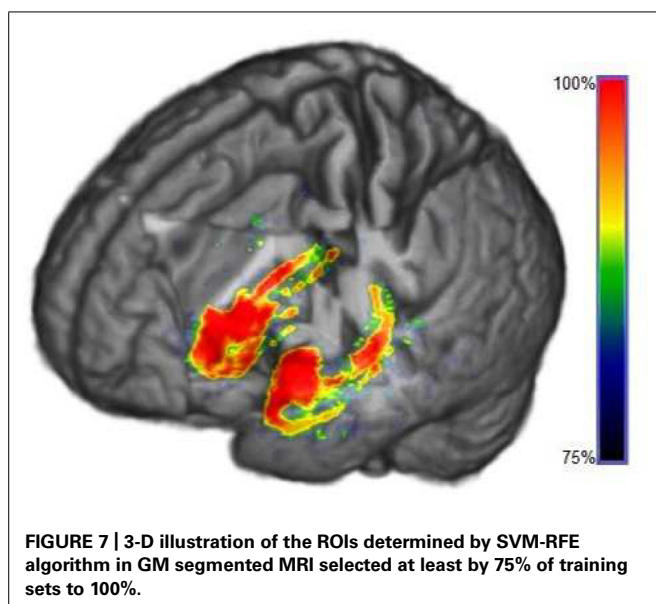
Figures 7 and 8 show three-dimensional illustrations of the ROIs computed by the SVM-RFE method for GM and WM, respectively. In the graphs are marked only those voxels that were selected by the majority of the training sets (more than 75% of the sets' total) in order to show the most robust results.

DISCUSSION

In this work, a novel application of a wrapper method based on SVM was performed to determine the ROIs from MRI, belonging to both GM and WM tissues, in order to better detect AD. Neither reduction of the data dimensionality nor previous voxel selection was required, thereby keeping the highest possible resolution to delimit the mentioned ROIs (see Figures 7 and 8).

From the beginning, the SVM-RFE algorithm was widely used in studies in order to select important genes (Guyon et al., 2002; Tang et al., 2007; Mundra and Rajapakse, 2010). Nonetheless, its potential applications have hardly been considered for image processing and specifically for ROIs computing from brain images. In





this paper, SVM-RFE has been shown to be able to select relevant ROIs that may aid an early diagnosis of AD. This is achieved by means of using the probabilities of belonging to GM or WM tissues of each voxel as features, which composes a segmented MRI. Methods based on SVM classifiers are suitable for analyzing big data, such as image processing, where the curse of dimensionality is a very common concern, since the SVM algorithm is not sensitive to over-fitting thanks to its margin definition (C parameter). In any case, it is fair to reiterate this concern to be cautious for the generalization of this kind of results.

The results obtained in the presented experiments have suggested that SVM-RFE selects discriminant features more efficiently than t -test significance for classification purposes (see section 3.4), as other research works report (Hidalgo-Muñoz et al., 2013a), specifically in digital signal processing. Furthermore,

since wrapped algorithms are based on the same classification technique, the complete method becomes more simple and intuitive than other multivariate statistical approaches. Moreover, the SVM-RFE permits choosing features from a huge set, without needing the multiple comparison corrections that are still a controversial issue in statistics (Perneger, 1998). Either way, the CV method into the wrapper algorithm is a key block for evaluating the accuracy and testing the capability of classifier generalization.

Regarding region morphology, ROIs delimited by SVM-RFE are located mainly in specific regions where more differences are evident by comparing the MRI averages, matching up considerably within the expected regions that are statistically different between groups (see **Figures 3** and **4**). However, the ROIs are not limited to the statistically significant regions, but also cover voxels surrounding these regions and disregard numerous voxels inside them. These regions' borders could be particularly important, since they could be related to the cortical atrophy suffered by AD patients in early stages (Quiroz et al., 2013). However, it would be necessary to check individually whether the ROI differences are mainly due to tissue deterioration (especially GM) or if they are influenced by some internal heterogeneity of the matter. For this reason, maintaining high MRI resolution is essential. This fact could question the suitability of some predefined ROIs by general brain templates (Magnin et al., 2009; Shen et al., 2011) in machine-learning applications that deal with such concrete neurological disorders as AD, whereas SVM-RFE takes all the features from MRI as a whole set without preconceived areas of interest.

The ROIs computed by SVM-RFE correspond intimately to the regions pointed in the literature as most affected by AD: hippocampus, entorhinal cortex, parahippocampal region (Du et al., 2001; Velayudhan et al., 2013) and insular cortex (Xie et al., 2012) among other regions like the amygdala, lenticular nucleus or fusiform gyrus (Tzourio-Mazoyer et al., 2002). As is well known, the hippocampus is a basic subcortical structure involved in declarative memory consolidation and spatial orientation (Squire, 1992; Tsien et al., 1996), which are skills severely affected by AD. These regions show an evident inter-hemisphere symmetry, particularly for GM segments (see **Figures 4** and **7**).

In addition, **Figures 4** and **7** show that although neocortex deterioration is often evident in AD (Quiroz et al., 2013), it is possible to point out a more general pattern for all patients in temporal lobes and insular cortex, whereas frontal lobe damage seems to be more diffuse in this inter-subject study. This fact can be influenced by the images' high-resolution, the necessary pre-processing and the specific and different development of the frontal lobes in people, depending on their experiences and even their educational level (Stuss and Knight, 2013) suggesting a remarkable contribution of individual differences.

The 3-D illustration helps to understand the possible generalization of the results for aid diagnosis or for clinical evaluations. Once the first ROIs have been determined, it would be possible to focus on these regions exclusively and separately in order to enclose and analyze in detail the pertinent anatomical regions (Ahmed et al., 2013). For instance, some research suggests that distinct regions of the hippocampus are affected differently in

AD (Burger, 2010; Lee et al., 2012). Therefore, a rigorous study using SVM-RFE might be recommended for exhaustive anatomical inspections and analyzing particular cases. On the other hand, by maintaining the high-resolution, it would be possible to examine some regions that are difficult to delimit in averaged images due to their location and size, such as the *locus coeruleus* or the fornix.

Regarding the comparison between GM and WM, the former tissue provides more relevant information and delimits more properly the important ROIs, as is suggested in much research that analyzes the GM directly (Thompson et al., 2003; Karas et al., 2004). Whereas for GM segmentation there are more highlighted regions using SVM-RFE, for WM segmentation the ROIs are distributed vaguely similarly to the results obtained with *t*-test. In any case, by means of the proposed methodology, it is possible to achieve practically 100% accuracy on average for both GM and WM segmentation options (Figure 6) after running a complete SVM-RFE in different feature sets.

In future works, it could be suggested to apply the same method to GM and WM segments obtained by other sophisticated MRI processing approaches. In addition, it is possible to employ the method in artificial intelligence applications in medicine for investigating diverse neurological disorders linked to senescence, such as fronto-temporal lobar degeneration, different forms of dementia or Parkinson's disease among others, where an accurate MRI information management is crucial.

CONCLUSION

In this paper, the main region of brain interest involved in Alzheimer's disease has been delimited by means of a SVM-based wrapper method applied on structural images, that is, MRI. The proposed method, which recursively eliminates the least-relevant features from the initial set (SVM-RFE), has proven to outperform *t*-test selection in terms of accuracy, achieving practically 100%. The high-resolution ROIs have been computed for both gray and white segmented matters, matching up with recent research that designates the hippocampal region as one of the most important in Alzheimer's disease development. In addition, 3-D illustrations of the regions have been provided in order to better understand the anatomical morphology linked to AD. Furthermore, this method, previously unexplored for MRI, could give valuable information about brain structures in other clinical applications on aging research.

AUTHOR CONTRIBUTIONS

Data used in preparation of this article was obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (<http://adni.loni.usc.edu/>). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data, but did not participate in the analysis or writing of this report.

ACKNOWLEDGMENTS

This work was partly supported by the MICINN under the TEC2012-34306 project and the Consejería de Innovación, Ciencia y Empresa (Junta de Andalucía, Spain) under the Excellence

Projects P09-TIC-4530 and P11-TIC-7103. Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI; National Institutes of Health Grant U01 AG024904). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: Abbott, AstraZeneca AB, Bayer Schering Pharma AG, Bristol-Myers Squibb, Eisai Global Clinical Development, Elan Corporation, Genentech, GE Healthcare, GlaxoSmithKline, Innogenetics, Johnson and Johnson, Eli Lilly and Co., Medpace, Inc., Merck and Co., Inc., Novartis AG, Pfizer Inc., F. Hoffman-La Roche, Schering-Plough, Synarc, Inc., as well as non-profit partners the Alzheimer's Association and Alzheimer's Drug Discovery Foundation, with participation from the U.S. Food and Drug Administration. Private sector contributions to ADNI are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro-Imaging at the University of California, Los Angeles. This research was also supported by NIH grants P30 AG010129, K01 AG030514, and the Dana Foundation.

REFERENCES

- Aguilar, C., Westman, E., Muehlboeck, J., Mecocci, P., Vellas, B., Tsolaki, M., et al. (2013). Different multivariate techniques for automated classification of MRI data in Alzheimer's disease and mild cognitive impairment. *Psychiatry Res.* 212, 89–98. doi: 10.1016/j.psychres.2012.11.005
- Ahmed, O. B., Benois-Pineau, J., Allard, M., Amar, C. B., and Catheline, G. (2013). "Alzheimer Disease detection on structural MRI," in *Proceedings of ESMRMB 2013 Congress*, Toulouse.
- Ashburner, J., and Friston, K. (2003). *Human Brain Function*, 2nd Edn, eds R. Frackowiak, K. Friston, C. Frith, R. Dolan, K. Friston, C. Price, et al. (San Diego, CA: Academic Press).
- Ashburner, J., and Friston, K. J. (2005). Unified segmentation. *Neuroimage* 26, 839–851. doi: 10.1016/j.neuroimage.2005.02.018
- Ashburner, J., Barnes, G., Chen, C., Daunizeau, J., Flandin, G., Friston, K., et al. (2012). "SPM8 manual," in *Functional Imaging Laboratory, Institute of Neurology*, London.
- Ben-Hur, A., Ong, C. S., Sonnenburg, S., Schölkopf, B., and Rätsch, G. (2008). Support vector machines and kernels for computational biology. *PLoS Comput. Biol.* 4:e1000173. doi: 10.1371/journal.pcbi.1000173
- Burger, C. (2010). Region-specific genetic alterations in the aging hippocampus: implications for cognitive aging. *Front. Aging Neurosci.* 2:140. doi: 10.3389/fnagi.2010.00140
- Burges, C. (1998). A tutorial on support vector machines for pattern recognition. *Data Min. Knowl. Disc.* 2, 121–167. doi: 10.1023/A:1009715923555
- Chu, C., Hsu, A. L., Chou, K. H., Bandettini, P., and Lin, C. (2012). Does feature selection improve classification accuracy? Impact of sample size and feature selection on classification using anatomical magnetic resonance images. *Neuroimage* 60, 59–70. doi: 10.1016/j.neuroimage.2011.11.066
- Clarke, L. P., Velthuisen, R. P., Camacho, M. A., Heine, J. J., Vaidyanathan, M., Hall, L. O., et al. (1995). MRI segmentation: methods and applications. *Magn. Reson. Imaging* 13, 343–368. doi: 10.1016/0730-725X(94)00124-L
- Cuadra, M. B., Cammoun, L., Butz, T., Cuisenaire, O., and Thiran, J. P. (2005). Comparison and validation of tissue modelization and statistical classification methods in T1-weighted MR brain images. *IEEE Trans. Med. Imaging* 24, 1548–1565. doi: 10.1109/TMI.2005.857652
- Cuingnet, R., Gerardin, E., Tessieras, J., Auzias, G., Lehéricy, S., Habert, M. O., et al. (2011). Automatic classification of patients with Alzheimer's disease from structural MRI: a comparison of ten methods using the ADNI database. *Neuroimage* 56, 766–781. doi: 10.1016/j.neuroimage.2010.06.013

- De Martino, F., Valente, G., Staeren, N., Ashburner, J., Goebel, R., and Formisano, E. (2008). Combining multivariate voxel selection and support vector machines for mapping and classification of fMRI spatial patterns. *Neuroimage* 43, 44–58. doi: 10.1016/j.neuroimage.2008.06.037
- Desikan, R. S., Cabral, H. J., Hess, C. P., Dillon, W. P., Glastonbury, C. M., Weiner, M. W., et al. (2009). Automated MRI measures identify individuals with mild cognitive impairment and Alzheimer's disease. *Brain* 132, 2048–2057. doi: 10.1093/brain/awp123
- Du, A. T., Schuff, N., Amend, D., Laakso, M. P., Hsu, Y. Y., Jagust, W. J., et al. (2001). Magnetic resonance imaging of the entorhinal cortex and hippocampus in mild cognitive impairment and Alzheimer's disease. *J. Neurol. Neurosurg. Psychiatry* 71, 441–447. doi: 10.1136/jnnp.71.4.441
- Folstein, M. F., Folstein, S. E., and McHugh, P. R. (1975). "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *J. Psychiatry Res.* 12, 189–198. doi: 10.1016/0022-3956(75)90026-6
- Frackowiak, R. S. (2004). *Human Brain Function*, eds K. J. Friston, C. D. Frith, R. J. Dolan, C. J. Price, S. Zeki, J. T. Ashburner, et al. (San Diego, CA: Academic Press).
- Gallix, A., Górriz, J. M., Ramírez, J., Illan, I. A., and Lang, E. W. (2012). On the empirical mode decomposition applied to the analysis of brain SPECT images. *Expert Syst. Appl.* 39, 13451–13461. doi: 10.1016/j.eswa.2012.05.058
- Guyon, I., Weston, J., Barnhill, S., and Vapnik, V. (2002). Gene selection for cancer classification using support vector machines. *Mach. Learn.* 46, 389–422. doi: 10.1023/A:1012487302797
- Hidalgo-Muñoz, A. R., López, M. M., Santos, I. M., Pereira, A. T., Vázquez-Marrufo, M., Galvao-Carmona, A., et al. (2013a). Application of SVM-RFE on EEG signals for detecting the most relevant scalp regions linked to affective valence processing. *Expert Syst. Appl.* 40, 2102–2108. doi: 10.1016/j.eswa.2012.10.013
- Hidalgo-Muñoz, A. R., Pereira, A. T., López, M. M., Galvao-Carmona, A., Tomé, A. M., Vázquez-Marrufo, M., et al. (2013b). Individual EEG differences in affective valence processing in women with low and high neuroticism. *Clin. Neurophysiol.* 124, 1798–1806. doi: 10.1016/j.clinph.2013.03.026
- Hölzel, B. K., Carmody, J., Vangel, M., Congleton, C., Yerramsetti, S. M., Gard, T., et al. (2011). Mindfulness practice leads to increases in regional brain gray matter density. *Psychiatry Res.* 191, 36–43. doi: 10.1016/j.psychres.2010.08.006
- Illán, I. A., Górriz, J. M., López, M. M., Ramírez, J., Salas-Gonzalez, D., Segovia, F., et al. (2011). Computer aided diagnosis of Alzheimer's disease using component based SVM. *Appl. Soft Comput.* 11, 2376–2382. doi: 10.1016/j.asoc.2010.08.019
- Jack, Jr. C. R., Bernstein, M. A., Borowski, B. J., Gunter, J. L., Fox, N. C., Thompson, P. M., et al. (2010). Update on the magnetic resonance imaging core of the Alzheimer's disease neuroimaging initiative. *Alzheimer's Dement.* 6, 212–220. doi: 10.1016/j.jalz.2010.03.004
- Kapur, T., Grimson, W. E. L., Wells, W. M. III, and Kikinis, R. (1996). Segmentation of brain tissue from magnetic resonance images. *Med. Image Anal.* 1, 109–127. doi: 10.1016/S1361-8415(96)80008-9
- Karas, G. B., Scheltens, P., Rombouts, S. A. R. B., Visser, P. J., Van Schijndel, R. A., Fox, N. C., et al. (2004). Global and local gray matter loss in mild cognitive impairment and Alzheimer's disease. *Neuroimage* 23, 708–716. doi: 10.1016/j.neuroimage.2004.07.006
- Klöppel, S., Stonnington, C. M., Chu, C., Draganski, B., Scahill, R. I., Rohrer, J. D., et al. (2008). Automatic classification of MR scans in Alzheimer's disease. *Brain* 131, 681–689. doi: 10.1093/brain/awm319
- Kohavi, R. (1995). "A study of cross-validation and bootstrap for accuracy estimation and model selection," in *Proceedings of The International Joint Conference on Artificial Intelligence*, Vol. 14. (Montreal, QC: IJCAI), 1137–1145.
- Kohavi, R., and John, G. H. (1997). Wrappers for feature subset selection. *Artif. Intell.* 97, 273–324. doi: 10.1016/S0004-3702(97)00043-X
- López, M., Ramírez, J., Górriz, J. M., Álvarez, I., Salas-Gonzalez, D., Segovia, F., et al. (2011). Principal component analysis-based techniques and supervised classification schemes for the early detection of Alzheimer's disease. *Neurocomputing* 74, 1260–1271. doi: 10.1016/j.neucom.2010.06.025
- Lee, D. Y., Fletcher, E., Carmichael, O. T., Singh, B., Mungas, D., Reed, B., et al. (2012). Sub-regional hippocampal injury is associated with fornix degeneration in Alzheimer's disease. *Front. Aging Neurosci.* 4:1. doi: 10.3389/fnagi.2012.00001
- Li, X., Coyle, D., Maguire, L., Watson, D. R., and McGinnity, T. M. (2011). Gray matter concentration and effective connectivity changes in Alzheimer's disease: a longitudinal structural MRI study. *Neuroradiology* 53, 733–748. doi: 10.1007/s00234-010-0795-1
- Liu, Y., Paajanen, T., Zhang, Y., Westman, E., Wahlund, L. O., Simmons, A., et al. (2010). Analysis of regional MRI volumes and thicknesses as predictors of conversion from mild cognitive impairment to Alzheimer's disease. *Neurobiol. Aging* 31, 1375–1385. doi: 10.1016/j.neurobiolaging.2010.01.022
- Magnin, B., Mesrob, L., Kinkingnehun, S., Pélégriani-Issac, M., Colliot, O., Sarazin, M., et al. (2009). Support vector machine-based classification of Alzheimer's disease from whole-brain anatomical MRI. *Neuroradiology* 51, 73–83. doi: 10.1007/s00234-008-0463-x
- Mazziotta, J. C., Toga, A. W., Evans, A., Fox, P., and Lancaster, J. (1995). A probabilistic atlas of the human brain: theory and rationale for its development the international consortium for brain mapping (ICBM). *Neuroimage* 2, 89–101. doi: 10.1006/nimg.1995.1012
- McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C. R. Jr., Kawas, C. H., et al. (2011). The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's Dement.* 7, 263–269. doi: 10.1016/j.jalz.2011.03.005
- Morris, J. C. (1993). The clinical dementia rating (CDR): current version and scoring rules. *Neurology* 43, 2412–2414. doi: 10.1212/WNL.43.11.2412-a
- Mundra, P. A., and Rajapakse, J. C. (2010). SVM-RFE with MRMR filter for gene selection. *IEEE Nanobiosci.* 9, 31–37. doi: 10.1109/TNB.2009.2035284
- Ortiz, A., Górriz, J. M., Ramírez, J., and Martínez-Murcia, F. J. (2013a). LVQ-SVM Based CAD tool applied to structural MRI for the diagnosis of the Alzheimer's disease. *Pattern Recogn. Lett.* 34, 1725–1733. doi: 10.1016/j.patrec.2013.04.014
- Ortiz, A., Palacio, A. A., Górriz, J. M., Ramírez, J., and Salas-González, D. (2013b). Segmentation of brain MRI using SOM-FCM-based method and 3D statistical descriptors. *Comput. Math. Methods Med.* 2013, 638563. doi: 10.1155/2013/638563
- Perneger, T. V. (1998). What's wrong with Bonferroni adjustments. *Brit. Med. J.* 316, 1236–1238. doi: 10.1136/bmj.316.7139.1236
- Quiroz, Y. T., Stern, C. E., Reiman, E. M., Brickhouse, M., Ruiz, A., Sperling, R. A., et al. (2013). Cortical atrophy in presymptomatic Alzheimer's disease presenilin 1 mutation carriers. *J. Neurol. Neurosurg. Psychiatry* 84, 556–561. doi: 10.1136/jnnp-2012-303299
- Rosenberg, P. B., and Hillis, A. E. (2009). Biomarkers for Alzheimer's disease: ready for the next step. *Brain* 132, 2002–2004. doi: 10.1093/brain/awp184
- Salas-Gonzalez, D., Górriz, J. M., Ramirez, J., Schloegl, M., Lang, E. W., and Ortiz, A. (2013). Parameterization of the distribution of white and grey matter in MRI using the α -stable distribution. *Comput. Biol. Med.* 43, 559–567. doi: 10.1016/j.combiomed.2013.01.003
- Salas-Gonzalez, D., Schlögl, M., Górriz, J. M., Ramirez, J., and Lang, E. (2011). "Bayesian segmentation of magnetic resonance images using the α -stable distribution," in *Hybrid Artificial Intelligent Systems*. (Springer Berlin Heidelberg), 99–106.
- Shattuck, D. W., Mirza, M., Adisetiyo, V., Hojatkashani, C., Salamon, G., Narr, K. L., et al. (2008). Construction of a 3D probabilistic atlas of human cortical structures. *Neuroimage* 39, 1064–1080. doi: 10.1016/j.neuroimage.2007.09.031
- Shen, Q., Loewenstein, D. A., Potter, E., Zhao, W., Appel, J., Greig, M. T., et al. (2011). Volumetric and visual rating of magnetic resonance imaging scans in the diagnosis of amnesic mild cognitive impairment and Alzheimer's disease. *Alzheimer's Dement.* 7, e101–e108. doi: 10.1016/j.jalz.2010.07.002
- Shen, Q., Zhao, W., Loewenstein, D. A., Potter, E., Greig, M. T., Raj, A., et al. (2012). Comparing new templates and atlas-based segmentations in the volumetric analysis of brain magnetic resonance images for diagnosing Alzheimer's disease. *Alzheimer's Dement.* 8, 399–406. doi: 10.1016/j.jalz.2011.07.002
- Shieh, M. D., and Yang, C. C. (2008). Multiclass SVM-RFE for product form feature selection. *Expert Syst. Appl.* 35, 531–541. doi: 10.1016/j.eswa.2007.07.043
- Squire, L. R. (1992). Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans. *Psychol. Rev.* 99, 195–231. doi: 10.1037/0033-295X.99.2.195
- Stuss, D. T., and Knight, R. T. (2013). *Principles of Frontal Lobe Function*. Oxford University Press.
- Suetens, P., Bellon, E., Vandermeulen, D., Smet, M., Marchal, G., Nuyts, J., et al. (1993). Image segmentation: methods and applications in diagnostic radiology and nuclear medicine. *Eur. J. Radiol.* 17, 14–21. doi: 10.1016/0720-048X(93)90023-G
- Tang, Y., Zhang, Y. Q., and Huang, Z. (2007). Development of two-stage SVM-RFE gene selection strategy for microarray expression data analysis.

- IEEE/ACM Trans. Comput. Biol. Bioinform.* 4, 365–381. doi: 10.1109/TCBB.2007.70224
- Thompson, P. M., Hayashi, K. M., de Zubicaray, G., Janke, A. L., Rose, S. E., Semple, J., et al. (2003). Dynamics of gray matter loss in Alzheimer's disease. *J. Neurosci.* 23, 994–1005.
- Tomé, A. M., Hidalgo-Muñoz, A. R., López, M. M., Teixeira, A. R., Santos, I. M., Pereira A. T., et al. (2013). "Feature extraction and classification of biosignals. Emotion valence detection from EEG signals" in *Proceedings of BIOSIGNALS 2013-International Conference on Bio-inspired Systems and Signal Processing*, Barcelona. 54–60.
- Tsien, J. Z., Huerta, P. T., and Tonegawa, S. (1996). The essential role of hippocampal CA1 NMDA receptor-dependent synaptic plasticity in spatial memory. *Cell* 87, 1327–1338. doi: 10.1016/S0092-8674(00)81827-9
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15, 273–289. doi: 10.1006/nimg.2001.0978
- Varma, S., and Simon, R. (2006). Bias in error estimation when using cross-validation for model selection. *BMC Bioinformatics* 7:91. doi: 10.1186/1471-2105-7-91
- Velayudhan, L., Proitsi, P., Westman, E., Muehlboeck, J. S., Mecocci, P., Velas, B., et al. (2013). Entorhinal cortex thickness predicts cognitive decline in Alzheimer's disease. *J. Alzheimer's Dis.* 33, 755–766. doi: 10.3233/JAD-2012-121408
- Weiner, M. W., Aisen, P. S., Jack, Jr. C. R., Jagust, W. J., Trojanowski, J. Q., Shaw, L., et al. (2010). The Alzheimer's disease neuroimaging initiative: progress report and future plans. *Alzheimer's Demen.* 6, 202–211. doi: 10.1016/j.jalz.2010.03.007
- Xie, C., Bai, F., Yu, H., Shi, Y., Yuan, Y., Chen, G., et al. (2012). Abnormal insula functional network is associated with episodic memory decline in amnesic mild cognitive impairment. *Neuroimage* 63, 320–327. doi: 10.1016/j.neuroimage.2012.06.062
- Zhang, Y., Schuff, N., Camacho, M., Chao, L. L., Fletcher, T. P., Yaffe, K., et al. (2013). MRI markers for mild cognitive impairment: comparisons between white matter integrity and gray matter volume measurements. *PLoS ONE* 8:e66367. doi: 10.1371/journal.pone.0066367

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 08 October 2013; paper pending published: 06 January 2014; accepted: 02 February 2014; published online: 20 February 2014.

Citation: Hidalgo-Muñoz AR, Ramírez J, Górriz JM and Padilla P (2014) Regions of interest computed by SVM wrapped method for Alzheimer's disease examination from segmented MRI. *Front. Aging Neurosci.* 6:20. doi: 10.3389/fnagi.2014.00020

This article was submitted to the journal *Frontiers in Aging Neuroscience*.

Copyright © 2014 Hidalgo-Muñoz, Ramírez, Górriz and Padilla. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.