

Article

Identification of Alzheimer's Disease on the Basis of a Voxel-Wise Approach

Iman Beheshti ^{1,*}, Hadi Mahdipour Hossein-Abad ^{2,*}, Hiroshi Matsuda ¹ and for the Japanese-Alzheimer's Disease Neuroimaging Initiative [‡]

¹ Integrative Brain Imaging Center, National Center of Neurology and Psychiatry, 4-1-1, Ogawahigashi-cho, Kodaira, Tokyo 187-8551, Japan

² Electrical Engineering Group, Esfarayen University of Technology, Esfarayen, North Khorasan 96619998195, Iran

* Correspondence: Beheshtiiman@gmail.com (I.B.); Mahdipourhadi@esfarayen.ac.ir (H.M.H.-A.)

† These authors contributed equally to this study.

‡ Data used in this article were obtained from Japanese-Alzheimer's Disease Neuroimaging Initiative (J-ADNI).

Received: 6 June 2019; Accepted: 25 July 2019; Published: 29 July 2019



Abstract: Robust prediction of Alzheimer's disease (AD) helps in the early diagnosis of AD and may support the treatment of AD patients. In this study, for early detection of AD and prediction of mild cognitive impairment (MCI) conversion, we develop an automatic computer-aided diagnosis (CAD) framework based on a merit-based feature selection method through a whole-brain voxel-wise analysis using baseline magnetic resonance imaging (MRI) data. We also explore the impact of different MRI spatial resolution on the voxel-wise metric AD classification and MCI conversion prediction. We assessed the proposed CAD framework using the whole-brain voxel-wise MRI features of 507 J-ADNI participants (146 healthy controls [HCs], 102 individuals with stable MCI [sMCI], 112 with progressive MCI [pMCI], and 147 with AD) among four clinically relevant pairs of diagnostic groups at different imaging resolutions (i.e., 2, 4, 8, and 16 mm). Using a support vector machine classifier through a 10-fold cross-validation strategy at a spatial resolution of 2 mm, the proposed CAD framework yielded classification accuracies of 91.13%, 74.77%, 81.12%, and 81.78% in identifying AD/healthy control, sMCI/pMCI, sMCI/AD, and pMCI/HC, respectively. The experimental results show that a lower spatial resolution (i.e., 2 mm) may provide more robust information to trace the neuronal loss-related brain atrophy in AD.

Keywords: Alzheimer's disease; mild cognitive impairment; early detection; merit-based feature selection; voxel-wise

1. Introduction

Alzheimer's disease (AD) is a neurodegenerative disorder that causes progressive problems with memory, cognition, learning, and higher executive function [1] in older people. Mild cognitive impairment (MCI), as a traditional phase between normal aging and AD, involves a slight but noticeable decline in memory and thinking [2]. Generally, MCI patients who later convert to AD are considered to have progressive MCI (pMCI), whereas MCI patients who have stable cognitive conditions are considered to have stable MCI (sMCI). One important clinical task is to distinguish between sMCI and pMCI in order to identify MCI patients who are at high risk of converting to AD in the foreseeable future. Consequently, early detection of this group of patients may help clinicians to effectively treat the condition.

To this end, different neuroimaging modalities such as magnetic resonance imaging (MRI) [3,4], positron emission tomography (PET), and single-photon emission computed tomography (SPECT) [5–9]

have been widely used in a variety of AD studies. In contrast to PET and SPECT modalities [10], MRI does not require injection of a radioactive pharmaceutical and also provides excellent spatial resolution with good contrast images. Radiological studies have confirmed that MRI measures can be used as a robust procedure to track brain atrophy due to neuronal loss [11]. In the case of AD, various MRI studies have shown that structural brain alterations due to AD can be used to distinguish AD patients from healthy controls (HCs) as well as pMCI patients from sMCI patients [12,13]. Furthermore, a recent study has documented that volumetric measures achieved from MRI data have the potential ability of detecting the biological heterogeneity patterns among amnesic MCI patients [14].

Generally, in the case of feature extraction from MRI, the AD classification studies can be categorized into three main groups: (1) Vertex-wise metrics use cortical and subcortical features, such as volumetric, thickness, and area information. The performance of this method is highly reliant on surface registration [15]. (2) Region of interest (ROI)-based metrics use a nonlinear registration followed by a template to extract the features from each ROI [16]. Practically, the dimension of features in this method is a function of the number of ROIs considered. As a drawback, this method has difficulty in reflecting the minor abnormalities associated with brain diseases [17]. (3) Voxel-wise metrics use the voxel intensities extracted from MRI scans as a feature. Although the voxel-wise metric robustly reflects the minor changes occurring due to brain disease, this approach suffers from the curse of dimensionality, in which the number of extracted features (i.e., voxel intensities) is markedly higher than the number of samples. With respect to this point, a robust feature selection for high-dimensional space plays a crucial role in voxel-wise approaches.

In the present study, we aimed to use a voxel-wise metric followed by advanced pattern recognition methods to develop a robust computer-aided diagnosis (CAD) system for AD classification. We introduce a merit-based feature selection (MFS) method as part of a proposed CAD framework to select the most informative features (i.e., voxel intensities) from high-dimensional space extracted from whole-brain analysis. We will show how, by using the proposed MFS, we are able to classify AD patients from healthy individuals and predict MCI conversion with robust performance. We additionally investigate the effect of different MRI spatial resolutions on the proposed CAD framework. The efficiency of the proposed CAD framework is assessed using a support vector machine (SVM) classifier followed by 10-fold cross-validation.

In summary, we, therefore, propose the following contributions to four pairs of clinical groups (i.e., AD/HC, pMCI/sMCI, pMCI/HC, and sMCI/AD patients) based on whole-brain voxel-wise MRI features obtained at baseline:

- (a) An automatic CAD framework based on a new feature selection method, namely the MFS, to identify the most informative features (i.e., voxel intensities) from high-dimensional space.
- (b) Investigation of the impact of different isotropic spatial resolutions on the voxel-wise AD classification and MCI conversion prediction.

2. Similar Studies

Researchers have widely investigated advanced neuroimaging and machine learning techniques for the early detection of AD and prediction of MCI conversion. Along these lines, the authors in Reference [16] proposed a multitask learning method based on multimodality data including MRI, fluorodeoxyglucose positron emission tomography (FDG-PET), and cerebrospinal fluid (CSF). They reported accuracies of 84.8%, 73.9%, and 62% for MRI with AD/HC, MCI/HC, and pMCI/sMCI samples, respectively. Moreover, they achieved accuracies of 93.30%, 83.2%, and 73.9% for the combination of MRI, CSF, and PET data with AD/HC, MCI/HC, and pMCI/sMCI samples, respectively. In Reference [18], the authors constructed structural connectivity networks of cortical morphometry based on individual gray matter images for AD classification and MCI conversion prediction. Using a histogram-based feature generation technique of structural connectivity networks, they achieved accuracies of 84.17%, 61.05%, 67.59%, and 70.38% with AD/HC, pMCI/sMCI, AD/sMCI,

and MCI/HC samples. The authors in Reference [19] proposed a semi-supervised SVM for predicting sMCI from pMCI using MRI. They achieved a sensitivity (SEN) of 78.8%, specificity (SPE) of 51%, and area under the curve (AUC) of 61% with a supervised method and SEN of 79.4%, SPE of 51.9%, and AUC of 69% with a proposed semi-supervised method. The authors in Reference [20] investigated a multivariate data analysis method using multimodality data (i.e., MRI and CSF). They reported accuracies of 87.0%, 71.8%, and 61.5% for AD/HC, MCI/HC, and pMCI/sMCI using MRI only and accuracies of 91.8%, 77.6%, and 66.7 for AD/HC, MCI/HC, and pMCI/sMCI by combining MRI and CSF. In Reference [21], the authors proposed the use of Spatial Pattern of Abnormalities for Recognition of Early AD (SPARE-AD) based on MRI. They reported accuracies of 55.8% and 61.7% in the prediction of sMCI from pMCI by MRI and MRI + t-tau, respectively. In Reference [22], the authors presented a multimodality classification framework based on pairwise connectivity measures derived by random forest classifiers. Using MRI, they reported accuracies of 81.6% in the classification of AD/HC, 66.9% in the classification of MCI/HC, and 55.1% for the prediction of pMCI/sMCI. In addition, they reported accuracies of 89.0%, 74.6%, and 58% for the classification of AD/HC, MCI/HC, and sMCI/pMCI groups by combining MRI with CSF, FDG-PET, and genetic data. The authors in Reference [23] investigated a Gaussian process approach using several combined multimodality data sources (i.e., MRI, PET, CSF, and APOE genotype). They stated balanced accuracies of 61.5% using MRI and of 74.1% using several combined multimodality data sources (i.e., MRI, PET, CSF, and APOE genotype) for sMCI/pMCI groups. Using voxel-based morphometry (VBM) analysis of gray matter (GM) as a feature followed by semi-supervised learning, the authors in Reference [13] presented a robust framework for MCI conversion prediction with accuracies of 74.74% based on MRI and of 82% by combining MRI features with age and cognitive measures in pMCI/sMCI samples. In another study [24], the authors calculated 55 volumetric variables from MRI data using FreeSurfer software (<http://surfer.nmr.mgh.harvard.edu/>, Version 5.1.0, Athinoula A. Martinos Center at the Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA). They reported accuracies of 78%, 66.7%, and 62.1% in the classification of AD/HC, amnesic-MCI/HC, and nonamnesic-MCI/HC using MRI data. In addition, they achieved accuracies of 92.4%, 74.9%, and 74.1% by combining MRI data with the Mini-Mental State Examination (MMSE) score in a classification of AD/HC, amnesic-MCI/HC, and nonamnesic-MCI/HC, respectively. The use of multiscale feature extraction based on baseline MRI data was presented in Reference [25] with accuracies of 84.13% and 76.69% for AD/HC and pMCI/sMCI, respectively. One study [3] used VBM and deformation-based morphometry (DBM) methods via MRI data for AD/HC classification with accuracies of 85%, 84%, and 79% on the basis of DBM, displacement magnitude, and Jacobian determinant feature extraction methods, respectively.

The researchers in Reference [26] presented a multivariate data analysis for AD detection using MRI data followed by independent component analysis, cognition data, and cerebrospinal fluid biomarkers. They stated accuracies of 84.7%, 94.8%, 83.1%, and 80.3% for HC/MCI, HC/AD, MCI/AD, and aMCI/pMCI tasks, respectively. A multimodal imaging technique on the basis of multi-kernel learning followed by a Dartel algorithm was proposed for early detection of AD by means of MRI and PET data [27], with prediction accuracies > 0.95% for AD/HC, MCI/HC, and AD/MCI tasks, respectively.

3. Materials and Methods

3.1. Image Acquisition and Participants

The T1-weighted MRI scans and clinical information were acquired from the Japanese-Alzheimer's Disease Neuroimaging Initiative (J-ADNI) dataset. All participants were assessed using neuropsychological assessment tools for AD, including the MMSE [28] and the Clinical Dementia Rating (CDR) [29], and then divided into four groups according to the results: (1) HC individuals (N: 146, mean age \pm sd: 68.28 \pm 5.61, mean MMSE \pm sd: 29.11 \pm 1.20, mean CDR \pm sd: 0.00 \pm 0.00); (2) sMCI individuals who were stable in MCI stage at least for 36 months (N: 102, mean age \pm sd: 73.44 \pm 5.97,

mean MMSE \pm sd: 26.55 ± 1.81 , mean CDR \pm sd: 0.50 ± 0.00); (3) pMCI individuals who converted from MCI to AD was within 6–36 months after baseline (N: 112, mean age \pm sd: 73.62 ± 5.57 , mean MMSE \pm sd: 26.08 ± 1.53 , mean CDR \pm sd: 0.50 ± 0.00); and (4) AD individuals (N: 147, mean age \pm sd: 74.07 ± 6.57 , mean MMSE \pm sd: 22.55 ± 1.80 , mean CDR \pm sd: 0.67 ± 0.23). Our study population included 507 participants at baseline. More details of J-ADNI dataset is described elsewhere [30].

3.2. MRI Preprocessing

Using the voxel-based morphometry technique implemented in statistical parametric mapping software (<https://www.fil.ion.ucl.ac.uk/spm/software/spm8/>, University College London, London, UK), all MRI images were bias-corrected, spatially normalized, and segmented into GM, white matter, and CSF components. Only the GM images were used. As per the pipeline proposed in [31,32], we used affine registration on GM images and then spatially smoothed them with an 8-mm full width at half maximum Gaussian kernel. To assess the impact of different resolutions on the voxel-wise AD classification, the smoothed GM images were resampled to 2-mm, 4-mm, 8-mm, and 16-mm isotropic spatial resolutions. The 1-mm spatial resolution was excluded from the analysis because of the high computational cost. The voxel intensities extracted from the smoothed GM images with different spatial resolutions were used as raw features for the classification.

3.3. Proposed MFS

In the proposed MFS, all merit features that are not similar to other ones are selected. For the input data $X = \{x_i\}_{i=1}^n = \{x_1, x_2, \dots, x_n\}$, where each input has p features (dimensions), $x_i = [x_{i,1}, x_{i,2}, \dots, x_{i,p}]^T$, $i = 1, 2, \dots, n$, the classification performance of each feature is evaluated separately. This means that when the input data are considered a training set for the classification, each feature (j -th feature) is considered the sole data feature and the classification is performed based on this single feature (classification using $\{x_{i,j}\}_{i=1}^n$). This process is performed for all features ($j = 1, 2, \dots, p$) and the corresponding classification accuracy of each feature is evaluated as CA_j , $j = 1, 2, \dots, p$. Then, all features are sorted based on classification accuracy. After this descending-based sorting, the features of each input data x_i are rearranged to $x_{(i)} = [x_{i,(1)}, x_{i,(2)}, \dots, x_{i,(p)}]^T$, where $CA_{(1)} \geq CA_{(2)} \geq \dots \geq CA_{(p)}$. Then, as the first selected feature, the feature with the best classification accuracy $x_{i,(1)}$ is selected. Therefore, the number of selected features so far (named nsf) is one ($nsf = 1$) and the selected features and data can be noted by $X_{fs} = \{x_{(i)}\}_{i=1}^n$, $x_{(i)} = [x_{i,(1)}]^T$, where $x_{i,(1)} = x_{i,(1)}$. In the next step and where the (m)-th sorted feature is examined for selection or not ($m = 2, \dots, p$), a threshold value (named T) is defined and used. After a comparison of the (m)-th feature with all previously selected nsf features one by one (e.g., let the ((1))-th selected feature, where l varies from 1 to nsf), if the (m)-th feature has at least T new correctly classified data where they are classified incorrectly using the ((1))-th feature, then the (m)-th feature has merit and is selected; otherwise, it is not selected.

Finally, the proposed MFS for p dimensional input data $= \{x_i\}_{i=1}^n = \{x_1, x_2, \dots, x_n\}$, $x_i = [x_{i,1}, x_{i,2}, \dots, x_{i,p}]^T$ can be summarized to a semi-code as follows:

- Step 1.** Set $j = 1$
- Step 2.** Perform classification using only the j -th feature (i.e., fulfilling the classification where the input data are $\{x_{i,j}\}_{i=1}^n$) and get the corresponding classification accuracy CA_j .
- Step 3.** If $j < p$, $j \rightarrow j + 1$ and go to **Step 2**; else, go to **Step 4**.
- Step 4.** Sort $\{CA_j\}_{j=1}^p$ in a descending manner and get the corresponding descended input data $X_d = \{x_{(i)}\}_{i=1}^n$, $x_{(i)} = [x_{i,(1)}, x_{i,(2)}, \dots, x_{i,(p)}]^T$, where $CA_{(1)} \geq CA_{(2)} \geq \dots \geq CA_{(p)}$.

- Step 5.** $x_{i,(1)}$ has merit and is selected. Therefore, the number of selected features so far (named nsf) is one (nsf = 1) and selected features and data can be noted by $X_{fs} = \{x_{((i))}\}_{i=1}^n, x_{((i))} = [x_{i,((1))}]^T$, where $x_{i,((1))} = x_{i,(1)}$. Set $m = 2$ and determine the threshold value T practically.
- Step 6.** After a comparison of the (m) -th feature with all previously selected nsf features one by one (e.g., let the $((1))$ -th selected feature, where l varies from 1 to nsf), if the (m) -th feature has at least T new correctly classified data where they are classified incorrectly using the $((1))$ -th feature, then the (m) -th feature has merit and is selected, add the (m) -th feature to X_{fs} and set the $nsf \rightarrow nsf + 1$; otherwise, the (m) -th feature is not selected.
- Step 7.** If $m < p$, $m \rightarrow m + 1$ and go to **Step 6**; otherwise, $X_{fs} = \{x_{((i))}\}_{i=1}^n, x_{((i))} = [x_{i,((1))}, x_{i,((2))}, \dots, x_{i,((nsf))}]^T$, is the output of the proposed feature selection method.

3.4. Validation and Classification

The classification was conducted using the support vector machine (SVM) algorithm to distinguish individuals between four pairs of clinical groups (i.e., AD/HC, pMCI/sMCI, pMCI/HC, and sMCI/AD). The SVM algorithm has been widely employed in different brain imaging studies [33–38]. We qualified the SVM classifier with a linear and nonlinear (i.e., Gaussian) kernels. The linear kernel (default settings) was chosen due to its better classification performances on our dataset. To validate the proposed MFS, we used k -fold cross-validation strategy ($k = 10$), because it has been widely used as a reliable validation process in a series of AD classification studies [39–42]. Based on the 10-fold cross-validation strategy, the results of each task were averaged over 10 repetitions so that, in each iteration, one-fold (10% of samples) was considered the test set and the remaining folds (90% of samples) were considered the training set. Of note, the subjects were randomly divided into 10-folds for each task. The classification results were assessed by means of accuracy (ACC), sensitivity (SEN), specificity (SPE), and area under the curve (AUC). We additionally computed the balanced accuracy (BACC) because of imbalance in the number of samples between two classes.

4. Experimental Results

As described in Section 3.2, we resampled the smoothed GM images to 2 mm, 4 mm, 8 mm, and 16 mm isotropic spatial resolutions. This procedure generated 238,955, 29,852, 3747, and 467 voxel intensities for each individual that were used as raw features. According to the proposed MFS, we sorted the features on the basis of their classification accuracy. An example of the sorted voxels for a spatial resolution of 2 mm is shown in Figure 1.

After sorting of the features (i.e., voxel intensities), the most informative features are selected on the basis of the proposed MFS as part of the CAD framework for the four pairs of clinical groups (i.e., AD/HC, pMCI/sMCI, pMCI/HC, and sMCI/AD). The classification performance obtained through 10-fold cross-validation without feature selection (i.e., all feature vectors) and through the proposed MFS approach followed by different isotropic spatial resolutions are presented in Table 1.

As can be seen in Table 1, the performance of the proposed CAD system followed by the MFS method was superior to the use of all features for all pairs of clinical groups for different spatial resolutions.

Figures 2 and 3 show the receiver operating characteristic (ROC) curves and the classification performance of the proposed CAD framework for the four pairs of clinical groups with respect to different spatial resolutions. As can be seen in Figures 2 and 3 as well as in Table 1, the performance of the proposed CAD framework was clearly improved with a decrease in the spatial resolution. One reason for this finding may be because a lower spatial resolution (i.e., 2 mm) may contribute to a more robust tracing of atrophic regions in AD. For example, in the case of AD/HC, the AUC was increased from 0.82 to 0.91 by reducing the spatial resolution from 16 mm to 2 mm. As can be seen in Figures 2 and 3, the optimum performance of the proposed CAD framework was achieved using

a 2-mm spatial resolution versus the other spatial resolutions (i.e., 4 mm, 8 mm, and 16 mm). The ROC curves through the proposed CAD framework for the four pairs of clinical groups with a 2-mm spatial resolution are shown in Figure 4.

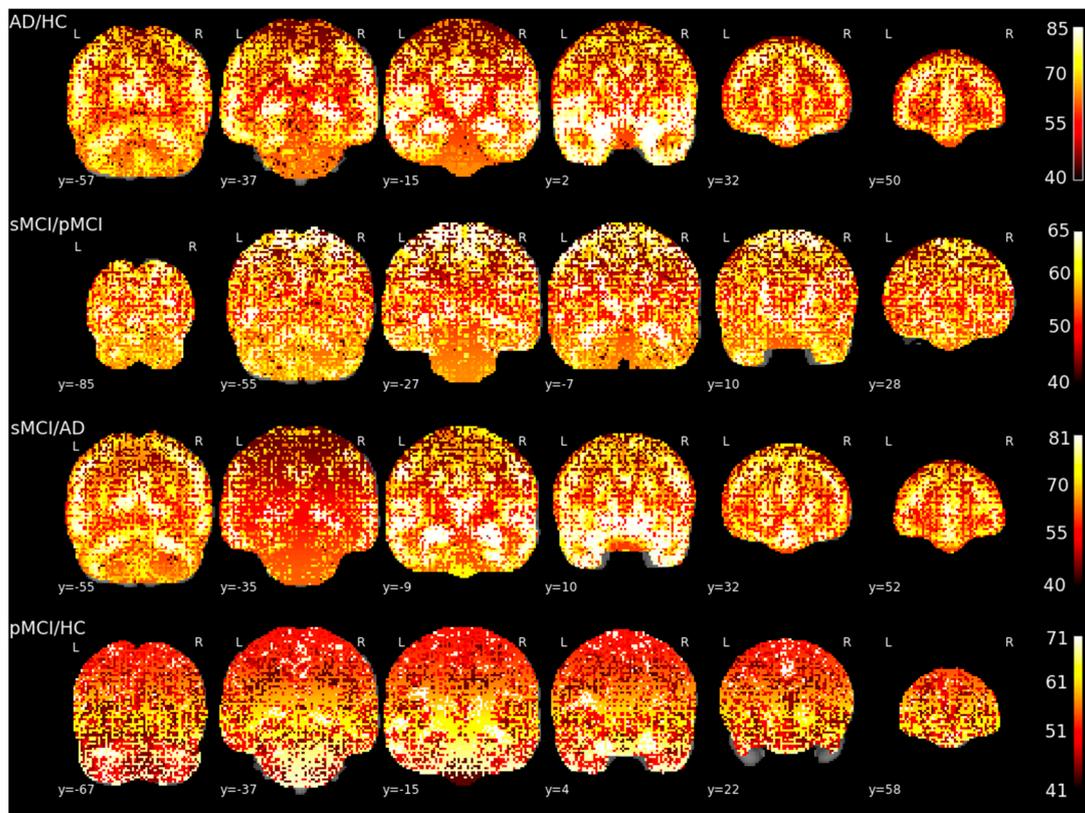


Figure 1. Sorted voxel intensities for a spatial resolution of 2 mm. The color bar shows the accuracy (%) of each individual voxel. AD: Alzheimer’s disease; HC: healthy controls; sMCI: stable mild cognitive impairment; pMCI: progressive mild cognitive impairment.

Table 1. The classification performance of the computer-aided diagnosis (CAD) system using all features and of the proposed merit-based feature selection (MFS) procedure with respect to different spatial resolutions through 10-fold cross-validation.

Task	Resolution	Raw Data (No Selection)				Followed by Proposed MFS Procedure					
		ACC	BACC	SEN	SPE	AUG	ACC	BACC	SEN	SPE	AUG
AD/HC	2 mm	83.28	83.28	85.62	80.95	0.83	91.13	91.12	91.10	91.16	0.91
	4 mm	82.93	82.94	85.61	80.27	0.82	88.73	88.73	89.04	88.43	0.89
	8 mm	83.96	83.96	86.30	81.63	0.84	88.05	88.06	91.78	84.35	0.87
	16 mm	81.23	81.23	82.88	79.59	0.80	83.28	83.28	84.93	81.63	0.82
pMCI/sMCI	2 mm	50.93	50.89	51.79	50.00	0.51	74.77	74.84	73.21	76.47	0.74
	4 mm	50.93	50.89	51.78	50.00	0.51	71.96	71.90	73.21	70.58	0.70
	8 mm	48.60	48.62	48.21	49.02	0.50	70.56	70.69	67.86	73.53	0.72
	16 mm	49.07	49.06	49.11	49.02	0.48	62.15	62.13	62.50	61.76	0.60
sMCI/AD	2 mm	59.84	59.08	54.90	63.27	0.54	81.12	80.41	76.47	84.35	0.77
	4 mm	55.82	54.03	44.12	63.95	0.50	71.48	71.49	71.56	71.42	0.69
	8 mm	52.61	50.71	40.19	61.22	0.47	71.89	70.48	62.75	78.23	0.74
	16 mm	52.61	50.56	39.22	61.90	0.46	65.46	64.44	58.82	70.07	0.64
pMCI/HC	2 mm	79.84	79.28	83.56	75.00	0.79	81.78	81.40	84.25	78.57	0.80
	4 mm	78.68	78.04	82.87	73.21	0.79	79.45	79.04	82.19	75.89	0.79
	8 mm	78.68	78.04	82.88	73.21	0.79	81.40	81.27	82.19	80.36	0.80
	16 mm	69.77	69.44	71.92	66.96	0.68	77.91	77.25	82.19	72.32	0.77

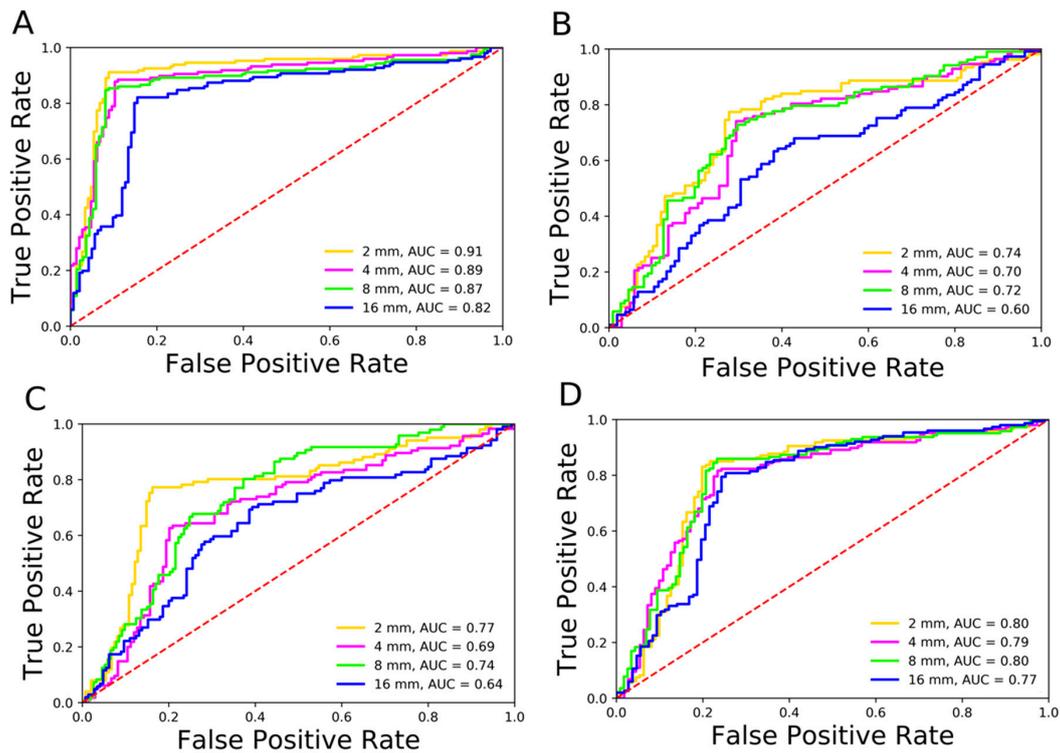


Figure 2. The receiver operating characteristic (ROC) curves of the proposed CAD framework with respect to the different isotropic spatial resolutions. (A) Alzheimer’s disease (AD)/healthy control (HC), (B) stable mild cognitive impairment (sMCI)/progressive MCI (pMCI), (C) sMCI/AD, and (D) pMCI/HC.

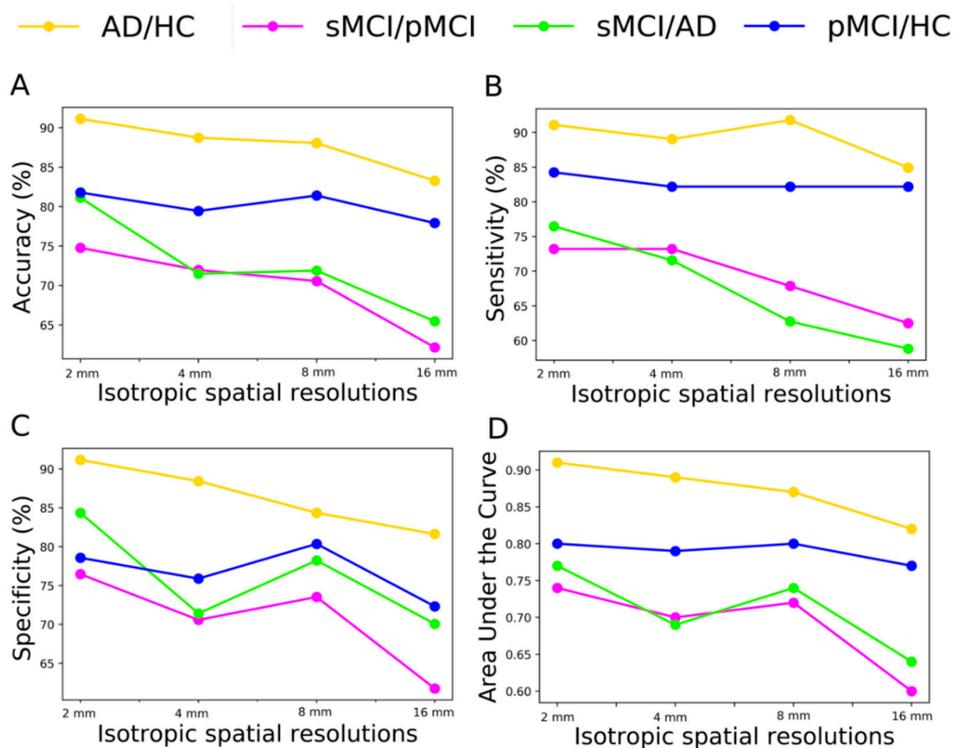


Figure 3. Classification results of the proposed CAD framework for the four pairs of clinical groups (i.e., AD/HC, pMCI/sMCI, pMCI/HC, and sMCI/AD) with respect to the different isotropic spatial resolutions. (A) accuracy (ACC) (%), (B) sensitivity (SEN) (%), (C) specificity (SPE) (%), and (D) area under the curve (AUC).

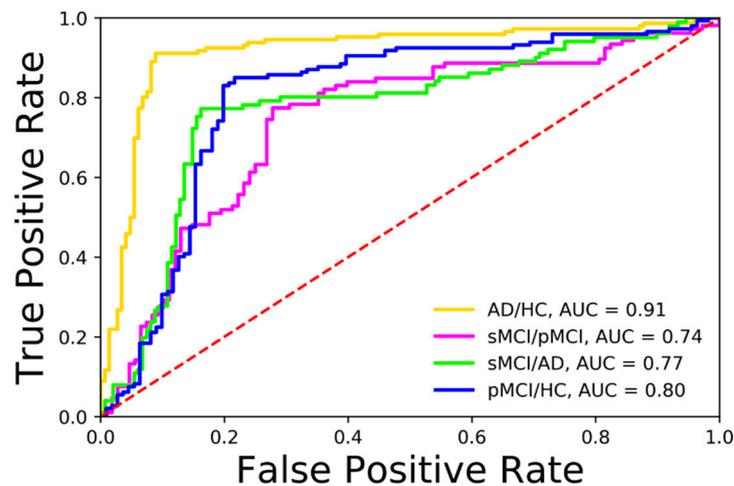


Figure 4. ROC curves of the proposed CAD framework for the four pairs of clinical groups (i.e., AD/HC, pMCI/sMCI, pMCI/HC, and sMCI/AD) with a 2-mm spatial resolution.

5. Discussion

Recent advances in machine learning research suggest that AD patients can be recognized at early stages of the disease via different neuroimaging modalities such as MRI [12,43,44], FDG-PET [45–47], and diffusion tensor imaging [48]. Among different neuroimaging modalities applied to AD, MRI is widely used because of its widespread availability and completely noninvasive nature [12]. Generally, AD classification studies using neuroimaging data suffer from the curse of dimensionality. Different dimensionality reduction and feature selection approaches have been proposed to overcome this drawback and improve the classification performance. For instance, the Principal Component Analysis (PCA) technique was proposed for the reduction of data extracted from SPECT and PET modalities [49]. Various AD studies additionally used the Partial Least Squares technique to transfer the data from a high-dimensional space into a lower dimensional vector [50–53]. Parallel to these approaches, a feature ranking-based strategy has been proposed for sorting the extracted features from high-dimensional space on the basis of their importance and then selecting the optimal subset of top-ranked features using Fisher criterion [54] and classification error [55].

In this study, we developed an automatic CAD framework followed by voxel-wise MRI features for AD detection and MCI conversion prediction. Although the voxel-wise metric is widely used for AD classification and brain-age estimation [31], it suffers from the curse of dimensionality (i.e., a high-dimensional feature vector compared with the number of samples). To address this limitation, we presented an MFS procedure that reduces dimensionality by eliminating redundant features (i.e., voxel intensities). The performance of the proposed CAD framework was assessed using 507 MRI scans taken at baseline from the J-ADNI data set. The experimental results indicated that the proposed MFS can identify the most informative features (i.e., voxel intensities) from high-dimensional space and that the proposed CAD system is highly competitive with state-of-the-art techniques using only MRI reported in the literature.

Lately, a comprehensive study was conducted in Reference [56] to explore the impact of different datasets (i.e., ADNI, AIBL, and OASIS), training set sizes, scanner protocols, preprocessing options, brain imaging modalities, classifiers, and feature types (i.e., voxel-wise and region-wise) on AD classification. In this study, we investigated the effect of different isotropic spatial resolutions on AD classification. Our experimental results demonstrate that spatial resolution may play an effective role in classification performance. Practically, the performance of the proposed CAD framework was improved by decreasing the spatial resolution. This may be because a lower spatial resolution (i.e., 2 mm) may help to trace atrophic regions in the AD brain. It is worth noting that a decrease in the spatial resolution increases the dimensionality and computational cost. For example, by decreasing

the spatial resolution from 8 mm to 2 mm, the dimensionality increases from 3747 to 238,955 features for each individual. Consequently, a tradeoff between the spatial resolution and computational cost is needed. For future work, we suggest to use the MFS technique for classification of AD and prediction of MCI on the basis of ROI-based metrics. In this study, we used only GM images. Another future direction will be to integrate the GM, white matter, and CSF components (i.e., data fusion) followed by MFS procedure to improve the classification accuracy for AD studies.

6. Conclusions

This study presents an automatic CAD framework for AD detection and MCI conversion prediction based on the voxel-wise metric. As part of a CAD framework, we present a MFS method to uncover the most informative voxel intensities from high-dimensional space. The performance of the proposed CAD framework was assessed in four pairs of clinical groups (i.e., AD/HC, pMCI/sMCI, pMCI/HC, and sMCI/AD) based on MRI scans taken at baseline from the J-ADNI data set with different spatial resolutions. The proposed CAD framework shows better performance with a low spatial resolution (i.e., 2 mm) than other spatial resolutions (i.e., 4 mm, 8 mm, and 16 mm). Indeed, features derived based on a fine spatial resolution improve discrimination of brain atrophy due to AD.

Author Contributions: Conceptualization: I.B. and H.M.H.; Methodology: H.M.H.; Writing-Original Draft Preparation: I.B.; Supervision: H.M.

Funding: This study was supported by an Intramural Research Grant (30-10) for Neurological and Psychiatric Disorders from the National Center of Neurology and Psychiatry (Japan).

Acknowledgments: We wish to thank all investigators and participants of Japanese-Alzheimer's Disease Neuroimaging Initiative (J-ADNI) project. Access to the original data of the J-ADNI is available on request from the NBDC Human Database (<http://humandbs.biosciencedbc.jp/en/>) hosted by the National Bioscience Database Center (NBDC) of the JST.

Conflicts of Interest: The authors declare no competing financial interests.

References

1. Alzheimer's Association|Alzheimer's Disease and Dementia. 2015. Available online: <http://www.alz.org/> (accessed on 5 April 2015).
2. Association, A. Alzheimer's Association Report 2015 Alzheimer's disease facts and figures. *Alzheimer's Dement* **2015**, *11*, 332–384.
3. Papakostas, G.A.; Savio, A.; Graña, M.; Kaburlasos, V.G. A lattice computing approach to Alzheimer's disease computer assisted diagnosis based on MRI data. *Neurocomputing* **2015**, *150*, 37–42. [[CrossRef](#)]
4. Gao, N.; Tao, L.X.; Huang, J.; Zhang, F.; Li, X.; O'Sullivan, F.; Chen, S.P.; Tian, S.J.; Mahara, G.; Luo, Y.X.; et al. Contourlet-based hippocampal magnetic resonance imaging texture features for multivariate classification and prediction of Alzheimer's disease. *Metab. Brain Dis.* **2018**, *33*, 1899–1909. [[CrossRef](#)] [[PubMed](#)]
5. Hanyu, H.; Sato, T.; Hirao, K.; Kanetaka, H.; Iwamoto, T.; Koizumi, K. The progression of cognitive deterioration and regional cerebral blood flow patterns in Alzheimer's disease: A longitudinal SPECT study. *J. Neurol. Sci.* **2010**, *290*, 96–101. [[CrossRef](#)] [[PubMed](#)]
6. Gray, K.R.; Wolz, R.; Heckemann, R.A.; Aljabar, P.; Hammers, A.; Rueckert, D. Multi-region analysis of longitudinal FDG-PET for the classification of Alzheimer's disease. *Neuroimage* **2012**, *60*, 221–229. [[CrossRef](#)] [[PubMed](#)]
7. Chen, Y.J.; Deutsch, G.; Satya, R.; Liu, H.-G.G.; Mount, J.M.Z. A semi-quantitative method for correlating brain disease groups with normal controls using SPECT: Alzheimer's disease versus vascular dementia. *Comput. Med. Imaging Graph.* **2013**, *37*, 40–47. [[CrossRef](#)]
8. Górriz, J.M.M.; Segovia, F.; Ramírez, J.; Lassel, A.; Salas-Gonzalez, D. GMM based SPECT image classification for the diagnosis of Alzheimer's disease. *Appl. Soft Comput.* **2011**, *11*, 2313–2325. [[CrossRef](#)]
9. Ding, X.; Bucholc, M.; Wang, H.; Glass, D.H.; Wang, H.; Clarke, D.H.; Bjourson, A.J.; Le Roy, C.D.; O'Kane, M.; Prasad, G.; et al. A hybrid computational approach for efficient Alzheimer's disease classification based on heterogeneous data. *Sci. Rep.* **2018**, *8*, 9774. [[CrossRef](#)]

10. Wang, Y.; Klunk, W.E.; Debnath, M.L.; Huang, G.F.; Holt, D.P.; Shao, L.; Mathis, C.A. Development of a PET/SPECT agent for amyloid imaging in Alzheimer's disease. *J. Mol. Neurosci.* **2004**, *24*, 55–62. [[CrossRef](#)]
11. Csernansky, J.G.; Hamstra, J.; Wang, L.; McKeel, D.; Price, J.L.; Gado, M.; Morris, J.C. Correlations between antemortem hippocampal volume and postmortem neuropathology in AD subjects. *Alzheimer Dis. Assoc. Disord.* **2004**, *18*, 190–195.
12. Bron, E.E.; Smits, M.; Van Der Flier, W.M.; Vrenken, H.; Barkhof, F.; Scheltens, P.; Papma, J.M.; Steketee, R.M.; Orellana, C.M.; Meijboom, R.; et al. Standardized evaluation of algorithms for computer-aided diagnosis of dementia based on structural MRI: The CADDementia challenge. *Neuroimage* **2015**, *111*, 562–579. [[CrossRef](#)] [[PubMed](#)]
13. Moradi, E.; Pepe, A.; Gaser, C.; Huttunen, H.; Tohka, J. Alzheimer's Disease Neuroimaging Initiative. Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects. *Neuroimage* **2015**, *104*, 398–412. [[CrossRef](#)] [[PubMed](#)]
14. Ezzati, A.; Zammit, A.R.; Habeck, C.; Hall, C.B.; Lipton, R.B. Detecting biological heterogeneity patterns in ADNI amnesic mild cognitive impairment based on volumetric MRI. *Brain Imaging Behav.* **2019**, 1–13. [[CrossRef](#)] [[PubMed](#)]
15. Cho, Y.; Seong, J.; Jeong, Y.; Yong, S. Individual subject classification for Alzheimer's disease based on incremental learning using a spatial frequency representation of cortical thickness data. *Neuroimage* **2012**, *59*, 2217–2230. [[CrossRef](#)]
16. Zhang, D.; Shen, D. Multi-modal multi-task learning for joint prediction of multiple regression and classification variables in Alzheimer's disease. *Neuroimage* **2012**, *59*, 895–907. [[CrossRef](#)] [[PubMed](#)]
17. Suk, H.I.; Lee, S.W.; Shen, D. Initiative, Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis. *Neuroimage* **2014**, *101*, 569–582. [[CrossRef](#)] [[PubMed](#)]
18. Beheshti, I.; Maikusa, N.; Daneshmand, M.; Matsuda, H.; Demirel, H.; Anbarjafari, G. Classification of Alzheimer's disease and prediction of mild cognitive impairment conversion using histogram-based analysis of patient-specific anatomical brain connectivity networks. *J. Alzheimer's Dis.* **2017**, *60*, 295–304. [[CrossRef](#)]
19. Filipovych, R.; Davatzikos, C. Semi-supervised pattern classification of medical images: Application to mild cognitive impairment (MCI). *Neuroimage* **2011**, *55*, 1109–1119. [[CrossRef](#)]
20. Westman, E.; Muehlboeck, J.S.; Simmons, A. Combining MRI and CSF measures for classification of Alzheimer's disease and prediction of mild cognitive impairment conversion. *Neuroimage* **2012**, *62*, 229–238. [[CrossRef](#)]
21. Davatzikos, C.; Bhatt, P.; Shaw, L.M.; Batmanghelich, K.N.; Trojanowski, J.Q. Prediction of MCI to AD conversion, via MRI, CSF biomarkers, and pattern classification. *Neurobiol. Aging* **2011**, *32*, 2322.e19–2322.e27. [[CrossRef](#)]
22. Gray, K.R.; Aljabar, P.; Heckemann, R.A.; Hammers, A.; Rueckert, D. Random forest-based similarity measures for multi-modal classification of Alzheimer's disease. *Neuroimage* **2013**, *65*, 167–175. [[CrossRef](#)] [[PubMed](#)]
23. Young, J.; Modat, M.; Cardoso, M.J.; Mendelson, A.; Cash, D.; Ourselin, S. Alzheimer's Disease Neuroimaging Initiative. Accurate multimodal probabilistic prediction of conversion to Alzheimer's disease in patients with mild cognitive impairment. *NeuroImage Clin.* **2013**, *2*, 735–745. [[CrossRef](#)]
24. Zhou, Q.; Goryawala, M.; Cabrerizo, M.; Wang, J.; Barker, W.; Loewenstein, D.A.; Duara, R.; Adjouadi, M. An Optimal Decisional Space for the Classification of Alzheimer's Disease and Mild Cognitive Impairment. *IEEE Trans. Biomed. Eng.* **2014**, *61*, 2245–2253. [[CrossRef](#)] [[PubMed](#)]
25. Hu, K.; Wang, Y.; Chen, K.; Hou, L.; Zhang, X. Multi-scale features extraction from baseline structure MRI for MCI patient classification and AD early diagnosis. *Neurocomputing* **2016**, *175*, 132–145. [[CrossRef](#)]
26. Willette, A.A.; Calhoun, V.D.; Egan, J.M.; Kapogiannis, D. Initiative, Prognostic classification of mild cognitive impairment and Alzheimer's disease: MRI independent component analysis. *Psychiatry Res. Neuroimaging* **2014**, *224*, 81–88. [[CrossRef](#)] [[PubMed](#)]
27. Youssofzadeh, V.; McGuinness, B.; Maguire, L.P.; Wong-Lin, K. Corrigendum: Multi-Kernel Learning with Dartel Improves Combined MRI-PET Classification of Alzheimer's Disease in AIBL Data: Group and Individual Analyses. *Front. Hum. Neurosci.* **2017**, *11*, 380. [[CrossRef](#)] [[PubMed](#)]
28. Folstein, M.F.; Folstein, S.E.; McHugh, P.R. 'Mini-mental state'. A practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr. Res.* **1975**, *12*, 189–198. [[CrossRef](#)]

29. Morris, J.C. The Clinical Dementia Rating (CDR): Current version and scoring rules. *Neurology* **1993**, *43*, 2412–2414. [[CrossRef](#)]
30. Fujishima, M.; Kawaguchi, A.; Maikusa, N.; Kuwano, R.; Iwatsubo, T.; Matsuda, H. Sample Size Estimation for Alzheimer's Disease Trials from Japanese ADNI Serial Magnetic Resonance Imaging. *J. Alzheimer's Dis.* **2017**, *56*, 75–88. [[CrossRef](#)]
31. Franke, K.; Ziegler, G.; Klöppel, S.; Gaser, C. Estimating the age of healthy subjects from T1-weighted MRI scans using kernel methods: Exploring the influence of various parameters. *Neuroimage* **2010**, *50*, 883–892. [[CrossRef](#)]
32. Tohka, J.; Moradi, E.; Huttunen, H. Comparison of Feature Selection Techniques in Machine Learning for Anatomical Brain MRI in Dementia. *Neuroinformatics* **2016**, *14*, 279–296. [[CrossRef](#)]
33. Dimitrovski, I.; Kocev, D.; Kitanovski, I.; Loskovska, S.; Džeroski, S. Improved medical image modality classification using a combination of visual and textual features. *Comput. Med. Imaging Graph.* **2015**, *39*, 14–26. [[CrossRef](#)]
34. Al-Kadi, O.S. A multiresolution clinical decision support system based on fractal model design for classification of histological brain tumours. *Comput. Med. Imaging Graph.* **2015**, *41*, 67–79. [[CrossRef](#)]
35. Li, M.; Qin, Y.; Gao, F.; Zhu, W.; He, X. Discriminative analysis of multivariate features from structural MRI and diffusion tensor images. *Magn. Reson. Imaging* **2014**, *32*, 1043–1051. [[CrossRef](#)]
36. Xue, T.; Bai, L.; Chen, S.; Zhong, C.; Feng, Y.; Wang, H.; Liu, Z.; You, Y.; Cui, F.; Ren, Y.; et al. Neural specificity of acupuncture stimulation from support vector machine classification analysis. *Magn. Reson. Imaging* **2011**, *29*, 943–950. [[CrossRef](#)]
37. Song, X.; Chen, N.K. A SVM-based quantitative fMRI method for resting-state functional network detection. *Magn. Reson. Imaging* **2014**, *32*, 819–831. [[CrossRef](#)]
38. Hinrichs, C.; Singh, V.; Xu, G.; Johnson, S.C. Alzheimers Disease Neuroimaging Initiative. Predictive markers for AD in a multi-modality framework: An analysis of MCI progression in the ADNI population. *Neuroimage* **2011**, *55*, 574–589. [[CrossRef](#)]
39. Cheng, B.; Liu, M.; Zhang, D.; Shen, D. Alzheimer's Disease Neuroimaging Initiative. Robust multi-label transfer feature learning for early diagnosis of Alzheimer's disease. *Brain Imaging Behav.* **2019**, *13*, 138–153. [[CrossRef](#)]
40. Wang, H.; Shen, Y.; Wang, S.; Xiao, T.; Deng, L.; Wang, X.; Zhao, X. Ensemble of 3D densely connected convolutional network for diagnosis of mild cognitive impairment and Alzheimer's disease. *Neurocomputing* **2019**, *333*, 145–156. [[CrossRef](#)]
41. Beheshti, I.; Demirel, H.; Matsuda, H. Classification of Alzheimer's disease and prediction of mild cognitive impairment-to-Alzheimer's conversion from structural magnetic resource imaging using feature ranking and a genetic algorithm. *Comput. Biol. Med.* **2017**, *83*, 109–119. [[CrossRef](#)]
42. Basaia, S.; Agosta, F.; Wagner, L.; Canu, E.; Magnani, G.; Santangelo, R.; Filippi, M. Alzheimer's Disease Neuroimaging Initiative. Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks. *NeuroImage Clin.* **2019**, *21*, 101645. [[CrossRef](#)]
43. Guo, H.; Liu, L.; Chen, J.; Xu, Y.; Jie, X. Alzheimer classification using a minimum spanning tree of high-order functional network on fMRI dataset. *Front. Neurosci.* **2017**, *11*, 639. [[CrossRef](#)]
44. Challis, E.; Hurley, P.; Serra, L.; Bozzali, M.; Oliver, S.; Cercignani, M. Gaussian process classification of Alzheimer's disease and mild cognitive impairment from resting-state fMRI. *Neuroimage* **2015**, *112*, 232–243. [[CrossRef](#)]
45. Zhang, D.; Wang, Y.; Zhou, L.; Yuan, H.; Shen, D. Alzheimer's Disease Neuroimaging Initiative. Multimodal classification of Alzheimer's disease and mild cognitive impairment. *Neuroimage* **2011**, *55*, 856–867. [[CrossRef](#)]
46. Higdon, R.; Foster, N.L.; Koeppe, R.A.; DeCarli, C.S.; Jagust, W.J.; Clark, C.M.; Barbas, N.R.; Arnold, S.E.; Turner, R.S.; Heidebrink, J.L.; et al. A comparison of classification methods for differentiating fronto-temporal dementia from Alzheimer's disease using FDG-PET imaging. *Stat. Med.* **2004**, *23*, 315–326. [[CrossRef](#)]
47. Gray, K.R.; Wolz, R.; Keihaninejad, S.; Heckemann, R.A. Regional Analysis of FDG-PET for use in The Classification of Alzheimer's Disease. In Proceedings of the 2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, Chicago, IL, USA, 30 March 2011; pp. 1082–1085.
48. Graña, M.; Termenon, M.; Savio, A.; Gonzalez-Pinto, A.; Echeveste, J.; Pérez, J.M.; Besga, A. Computer aided diagnosis system for Alzheimer disease using brain diffusion tensor imaging features selected by Pearson's correlation. *Neurosci. Lett.* **2011**, *502*, 225–229. [[CrossRef](#)]

49. López, M.; Ramírez, J.; Górriz, J.M.; Álvarez, I.; Salas-Gonzalez, D.; Segovia, F.; Puntonet, C.G. Computer Aided Diagnosis of Alzheimer's Disease Using Principal Component Analysis and Bayesian Classifiers. In Proceedings of the Sixth International Symposium on Neural Networks (ISNN 2009), Wuhan, China, 26–29 May 2009; pp. 213–221.
50. Segovia, F.; Górriz, J.M.; Ramírez, J.; Salas-González, D.; Álvarez, I. Early diagnosis of Alzheimer's disease based on Partial Least Squares and Support Vector Machine. *Expert Syst. Appl.* **2013**, *40*, 677–683. [[CrossRef](#)]
51. Chaves, R.; Ramírez, J.; Górriz, J.M.; Puntonet, C.G. Association rule-based feature selection method for Alzheimer's disease diagnosis. *Expert Syst. Appl.* **2012**, *39*, 11766–11774. [[CrossRef](#)]
52. Ramírez, J.; Górriz, J.M.; Segovia, F.; Chaves, R.; Salas-Gonzalez, D.; López, M.; Álvarez, I.; Padilla, P. Computer aided diagnosis system for the Alzheimer's disease based on partial least squares and random forest SPECT image classification. *Neurosci. Lett.* **2010**, *472*, 99–103. [[CrossRef](#)]
53. Khedher, L.; Ramírez, J.; Górriz, J.M.; Brahim, A.; Segovia, F. Early diagnosis of Alzheimer's disease based on partial least squares, principal component analysis and support vector machine using segmented MRI images. *Neurocomputing* **2015**, *151*, 139–150. [[CrossRef](#)]
54. Beheshti, I.; Demirel, H. Feature-ranking-based Alzheimer's disease classification from structural MRI. *Magn. Reson. Imaging* **2016**, *34*, 252–263. [[CrossRef](#)]
55. Beheshti, I.; Demirel, H.; Farokhian, F.; Yang, C.; Matsuda, H. Structural MRI-based detection of Alzheimer's disease using feature ranking and classification error. *Comput. Methods Programs Biomed.* **2016**, *137*, 177–193. [[CrossRef](#)]
56. Samper-Gonzalez, J.; Burgos, N.; Bottani, S.; Fontanella, S.; Lu, P.; Marcoux, A.; Routier, A.; Guillon, J.; Bacci, M.; Wen, J.; et al. Reproducible evaluation of classification methods in Alzheimer's disease: Framework and application to MRI and PET data. *Neuroimage* **2018**, *183*, 504–521. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).