Article

VBM-Based Alzheimer's Disease Detection from the Region of Interest of T1 MRI with Supportive Gaussian Smoothing and a Bayesian Regularized Neural Network

Bijen Khagi [1], Kun Ho Lee [2,3,4], Kyu Yeong Choi [3,4], Jang Jae Lee [3,4], Goo-Rak Kwon [1,4] and Hee-Deok Yang [5,*]

1   School of Information Communication Engineering, Chosun University, 309 Pilmun-Daero, Dong-Gu, Gwangju 61452, Korea; bijenkhagi@gmail.com (B.K.); grkwon@chosun.ac.kr (G.-R.K.)
2   Department of Biomedical Science, Chosun University, Gwangju 61452, Korea; leekho@Chosun.ac.kr
3   Aging Neuroscience Research Group, Korea Brain Research Institute, Daegu 41062, Korea; khaser@Chosun.ac.kr (K.Y.C.); jjlee@Chosun.ac.kr (J.J.L.)
4   Gwangju Alzheimer's Disease and Related Dementia Cohort Research Center, Chosun University, Gwangju 61452, Korea
5   Department of Computer Engineering, Chosun University, 309 Pilmun-Daero, Dong-Gu, Gwangju 61452, Korea
*   Correspondence: heedeok_yang@chosun.ac.kr

Abstract: This paper presents an efficient computer-aided diagnosis (CAD) approach for the automatic detection of Alzheimer's disease in patients' T1 MRI scans using the voxel-based morphometry (VBM) analysis of the region of interest (ROI) in the brain. The idea is to generate a normal distribution of feature vectors from ROIs then later use for classification via Bayesian regularized neural network (BR-NN). The first dataset consists of the magnetic resonance imaging (MRI) of 74 Alzheimer's disease (AD), 42 mild cognitive impairment (MCI), and 74 control normal (CN) from the ADNI[1] dataset. The other dataset consists of the MRI of 42 Alzheimer's disease dementia (ADD), 42 normal controls (NCs), and 39 MCI due to AD (mAD) from our GARD[2] database. We aim to create a generalized network to distinguish normal individuals (CN/NC) from dementia patients AD/ADD and MCI/mAD. Our performance relies on our feature extraction process and data smoothing process. Here the key process is to generate a Statistical Parametric Mapping (SPM) *t*-map image from VBM analysis and obtain the region of interest (ROI) that shows the optimistic result after two-sample *t*-tests for a smaller value of $p < 0.001$(AD vs. CN). The result was overwhelming for the distinction between AD/ADD and CN/NC, thus validating our idea for discriminative MRI features. Further, we compared our performance with other recent state-of-the-art methods, and it is comparatively better in many cases. We have experimented with two datasets to validate the process. To validate the network generalization, BR-NN is trained from 70% of the ADNI dataset and tested on 30% of the ADNI, 100% of the GARD dataset, and vice versa. Additionally, we identified the brain anatomical ROIs that may be relatively responsible for brain atrophy during the AD diagnosis.

Keywords: Alzheimer's disease; VBM; AAL; *t*-test; feature extraction; BR-NN; computer-aided diagnosis (CAD)

1. Introduction

Alzheimer's disease is a neurodegenerative disorder causing dementia, which frequently affects elderly people. The disease mainly affects the brain and its vital parts. In this disease, the brain is affected in various regions causing a significant change in the structural and functional region of the brain [1,2]. With the breakthrough of various imaging techniques such as MRI, positron emission tomography (PET), and computed tomography (CT) scans in the medical sector, numerous efforts have been conducted to process, simulate, and interpret the results for computer-aided diagnosis (CAD) that will be

crucial for medical professionals. Similarly, various research using structural and functional MRI as the primary biomarker has been performed to efficiently develop a CAD system for Alzheimer's disease diagnosis and detection. MRI is a magnetic-field gradient-based neuroimaging biomarker technique that provides anatomic and physiological information for the diagnosis of different parts of the body including the brain. It uses a strong magnetic field and radio-wave to generate a higher-quality picture of the structure and volume of the brain, called structural MRI (sMRI). Although sMRI does not show the functional activity of the brain, it can reflect the content of brain tissue that can be useful to detect the abnormalities in the brain of AD patients compared to the healthy ones. These discriminative features lie in the content of white matter or ventricles, with a change in cortical thickness, hippocampus shape, and brain volume, which can be the probable structural features to detect AD. On group comparison, the voxel intensities of obtained clusters after the *t*-test amplifies these differences, eventually generating ROIs for feature extraction.

In neuroimaging, VBM [3] is used to investigate the anatomic morphology of the brain and its related part using the intensity value of every voxel present in the MRI scan. It is generally used with different statistical parametric mapping approaches to analyze tissue contents between two or more groups of patients under different hypothesis tests. VBM extracts features of three tissue distributions: gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF) in each voxel considered at a time. It follows the pipeline as executed by John et al. [3] and was later followed by Savio et al. [4], Busatto et al. [5], Frisoni et al. [6], and Beheshti et al. [7,8] for Alzheimer's prone structural MRI classification. Savio et al. [4] applied VBM in the Open Access Series of Imaging Studies (OASIS) dataset to detect AD patients from the NC and myotonic dystrophy of type 1 (MD1) patients from healthy controls (HCs) from the dataset of the Neurology Department of the Donostia hospital. The extracted features were voxels values and mean and standard deviation (MSD) from SPM *t*-map images. These features were classified using machine-learning tools such as Support vector machine (SVM) and have achieved the highest accuracy of 81%. Beheshti et al. [7] applied a similar technique to detect 130 AD and 130 HCs, extracting the voxel value from voxels of interest (VOI) using a probability distribution function (PDF) for feature reduction (selection), which is later classified using SVM and RBF kernels. They used the ADNI dataset and reported up to 90% accuracy from ten-fold cross-validation. Subsequently, Beheshti et al. [8] again published a paper with higher accuracy of up to 92.48% using seven feature ranking methods for the same VOI-based features. On the other hand, non-Morphometry based feature extraction technique using classical methods such as Dual-Tree Complex Wavelet Transform (DTCWT) [9], wavelet entropy [10], Singular Value Decomposition (SVD) [11], Scaled Chaotic Artificial Bee Colony (SCABC) [12], Downsized Kernel Principal Component Analysis (DKPCA) [13] including machine-learning and deep-learning techniques [14–16] like Recurrent Neural Network and Semantic Segmentation [17,18] have been used in recent times in CAD development for AD detection between NC and MCI. Also, major advances in structural and functional neuroimaging studies have resulted in improvements in the early and accurate detection of AD [19–24].

In this study, we investigated the VBM method using SPM 12 for MRI classification. We had combined the traditional VBM pipeline [3] along with our proposed feature extraction process to support the classification of the normalized feature vectors. Following the VBM execution pipeline, we obtained a *t*-map image showing the areas where distinction exists between two groups (CN and AD/MCI), used in the *t*-test. The obtained *t*-map is used to identify ROIs to generate the features for classification. To support the classifier with a normal distribution of features belonging to each group, we implemented Gaussian smoothing to generate weighted average feature vectors without any feature reduction. Subsequently, the smoothened features are fed into the neural network along with their label and trained based on Bayesian optimization. Once the network is trained, it is tested on other sets of data, and finally, the results are evaluated.

2. Materials

2.1. Image Acquisition

The T1-MR images and data used in this work were obtained from the Alzheimer's disease neuroimaging initiative (ADNI) database [25]. The selected sMRI were acquired at 1.5 T using a Siemens scanner and preprocessed using the University of California, San Diego (UCSD) ADNI pipeline in two steps for a three-dimensional (3D) Grad-warp [26,27] and B1 non-uniformity correction [28]. It is important to consider the work of Clifford et al. [29] who studied the approaches for standardization of ADNI MRI protocols and acquisition parameters for each imaging sequence, post-acquisition correction of image artifacts, and other technical non-uniformity for data variation minimization. The obtained MRI scans were in Nifti (.nii format) 3D volume, which can be processed for VBM analysis as illustrated in Figure 1.
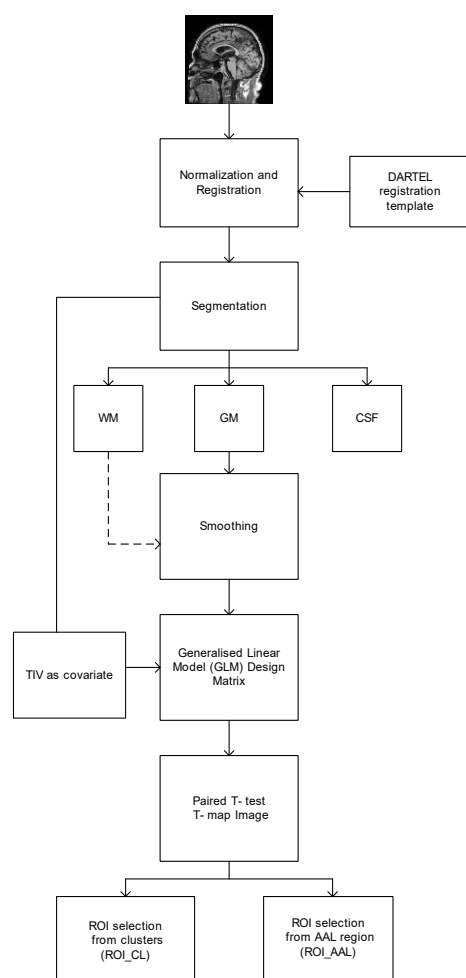
Figure 1. Workflow in VBM.

2.2. Participants

The first dataset for our study is from ADNI that consists of 74 AD, 74 CN, and 40 Mild MCI patients' scans. The training and testing sets were selectively adapted from Cuingnet et al. [2]. We have selected a limited number of AD and NC, as the available MCI subjects are comparatively less, so to balance the classwise approximation of the classifier, we have used limited subjects in both datasets for comparative study. We have selected one MRI from each patient having a uniform dimension of $256 \times 256 \times 166$ (to avoid a co-registration problem). The second dataset is from GARD that consists of 42 ADD, 42 NC, and 39 mAD. T1 weighted MRI scans are preferred in both datasets for the experiment as most of the VBM based work uses T1 scans and it has been reported that the

contrast-enhanced T1-weighted MRI was significantly superior to T2-weighted for imaging assessment without statistically significant difference [30]. The detailed demographic is shown in Table 1.

Table 1. Demographic details. ADNI participant complete details are included in the Supplementary File. CN and NC, in general, represent the same normal patients, it is the categorical labels used in each database, i.e., ADNI and GARD respectively so, is the AD and ADD. '#' meaning number.

Dataset	Information	Cohort Domain		
		AD/ADD	**MCI/mAD**	**CN/NC**
ADNI	# of male/female subject	35/39	21/19	36/40
	Mean age	76.18 ± 7.46	74.46 ± 6.88	75.8 9 ± 5.26
GARD	# of male/female subject	24/18	24/15	24/18
	Mean age	76.25 ± 3.33 75.03 ± 6.29	74.75 ± 3.588 72.06 ± 2.89	76.26 ± 4.57 69.66 ± 3.09

3. Methodology

We will discuss the proposed method thoroughly following each step involved. Figure 2 provides a quick illustration of the proposed method.
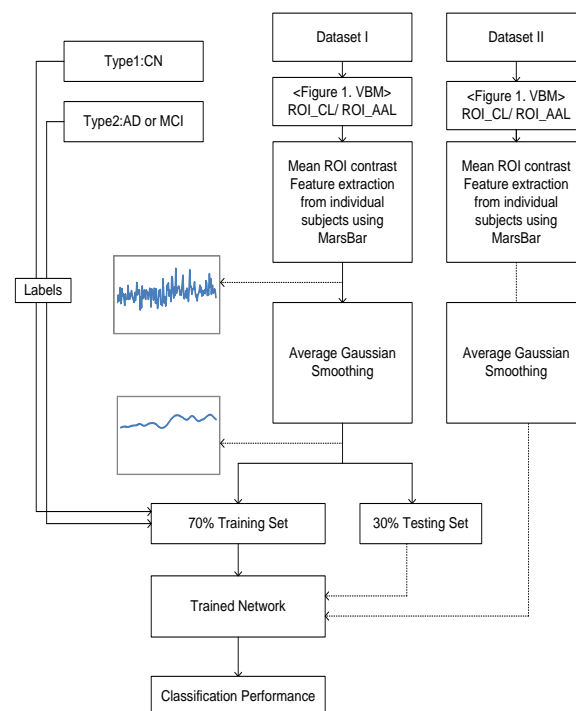
Figure 2. Proposed CAD classification method.

3.1. VBM SPM Model

All MRI scans undergo the preprocessing process as in 3.1.1 required to make all MRI scans ready for SPM modeling. We have used a GUI-based SPM toolbox for SPM model generation. 3.1.1 to 3.1.2 is implemented manually using the SPM toolbox as shown in Figure 3.
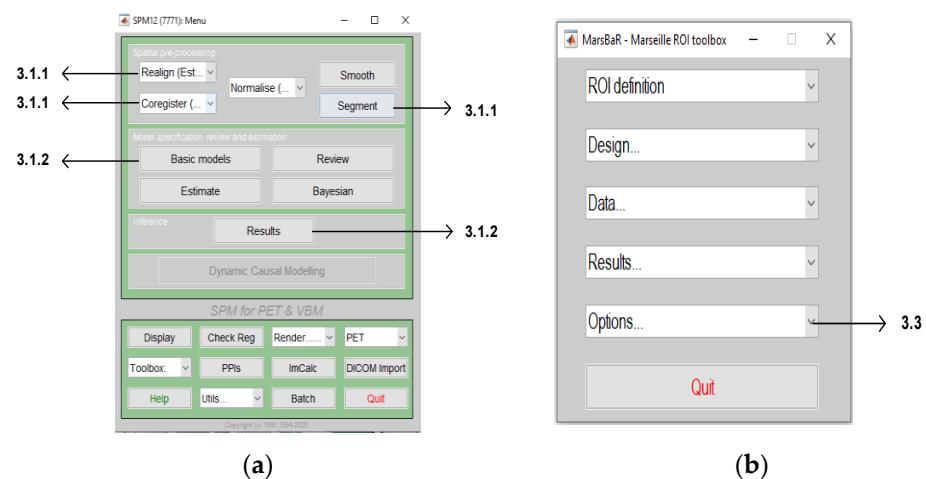
Figure 3. Screenshot from (**a**) GUI of SPM12 for 3.1.1. and 3.1.2. (**b**) GUI of MarsBaR for 3.3. Each number arrowed in the figure represents the corresponding section in this article.

3.1.1. MRI Preprocessing for Voxel-Based Model

The major process in VBM includes: (i) spatial normalization and diffeomorphic anatomical registration through exponentiated lie algebra (DARTEL) registration, (ii) modulation and segmentation, and (iii) smoothing. Spatial normalization transforms all the participants' volume to the same stereotactic space for uniformity. The SPM 12 MATLAB toolbox is used for spatial registration. It uses the DARTEL [31] registration template which was created from 555 participants of the IXI dataset [3] between 20–80 years. We used tissue probability maps (TPM) [32] for the initial spatial registration and segmentation as a reference map. The final image was smoothened using the Gaussian smoothing kernel size [8 8 8] to suppress noise and effects due to the residual difference in functional and gyral-related anatomies during inter-participant averaging. Hence, the normalized-modulated-segmented smooth image of voxel 1.5 mm and dimensions 121 × 145 × 121 was finally formed for each tissue volume, i.e., GM, WM, and CSF. We have considered the GM and WM volume as the major input for the further mapping process.

3.1.2. SPM *t*-Map

Normalized DARTEL—wrapped, modulated smooth GM and WM images were separately used to build a general linear model (GLM) to detect gray matter (or white matter) volume changes performing a voxel-wise two-sample *t*-test (one group being CN and the other being MCI or AD) in SPM12 using Equation (1):

$$Y_i = X_i\beta + \varepsilon_i \tag{1}$$

Here, Y_i is the predicted functional value for the *i*th participant and X_i (design matrix) is the vector of regressor variables for the *i*th participant. For the matrix test, X_i is set [1 0] for group 1 and [0 1] for group 2, during two-sample *t*-tests. β is a vector of parameters that varies from voxel to voxel and must be estimated after the model is built. In the two-sample *t*-tests, it represents the mean contrast of each group including covariates value as well. ε_i is a normally distributed error term with a mean of zero and a variance of unit S.D and varies across voxels. The total intracranial volume (TIV) was engaged in the design matrix as a covariate with an absolute threshold at a value of 0.2. This model was implemented using $p < 0.001$ and Gaussian random field (GRF) theory [33] for family-wise error (FWE) correction to produce a final *t*-map image. Here, the *p*-value indicates the probability of the null hypothesis being true; therefore, the lower the *p*-value, the better is the model. Once SPM *t*-map image was generated, we could concentrate ROI selection on those areas with GM volumes change detected by a voxel-based analysis using contrast as a feature as in Figures 3 and 4, for further feature extractions. We have considered the selection and

generation of the ROI as the most crucial and result-oriented process. If the selected ROI is not chosen properly, an unfavorable result may be obtained.
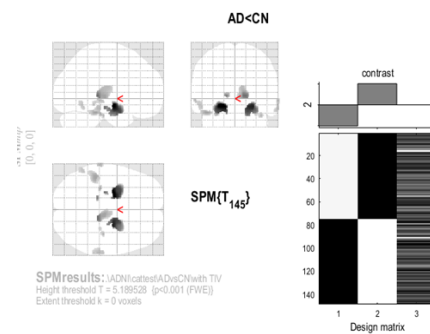

Figure 4. ADNI SPM *t*-map glass view for [CN AD] = [1 −1].

3.2. ROI Selection

It is done after SPM modeling i.e., Sections 3.1.2, 3.2.1 and 3.2.2 describe our proposed method for ROI selection. It is done manually as described in the respective sections. Corresponding results using screenshots are shown in Figures 3 and 4, respectively. Section 3.3 is performed via MarsBaR (also SPM toolbox extension) and manual selection. Each section explains in detail how the process is done and what parameters were selected for the final implementation.

3.2.1. From Each Cluster Generated into Separate ROI

These ROIs are identified based on the *t*-map generated from VBM analysis. The regions which are significantly different and bundled into the same region are considered as a cluster, as shown in Figures 4–6. The generated clusters vary in shape and size, based on the significance of the *p*-values and the error correction used; *p*-value was computed based on a repetitive experiment in the range 0.1 to 0.0001, the one with the highest voxel number is reported i.e., *p*< 0.001. Thus, we have selected for *p* < 0.001 where 14 clusters were selected out of 30 clusters from the *t*-map of GM volumes, and one cluster from the *t*-map of the WM volumes both from the AD and CN two-sample *t*-tests. Similarly, three clusters were generated from the t-map of the GM volumes of the CN and MCI two-sample *t*-tests. The selected 18 clusters were used as ROI directly to obtain the mean weighted voxel value from each participating MRI; we will term this ROI as ROI_CL hereinafter. The number of voxels in each ROI_CL differs abruptly so, if we accumulate all the voxels from each ROI, the number of voxels is going to be substantially large; hence one of the logical solutions would be taking mean weighted voxel value as stated in Equation (2) to represent a single value for each selected ROI for each MRI, demonstrated as shown in Appendix A, Table A1.
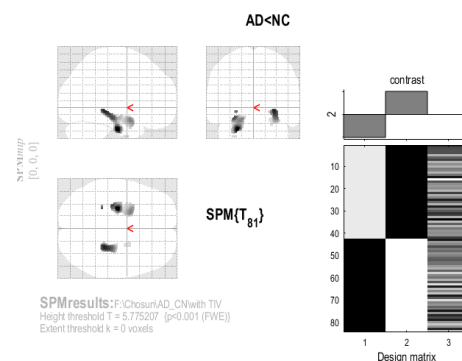

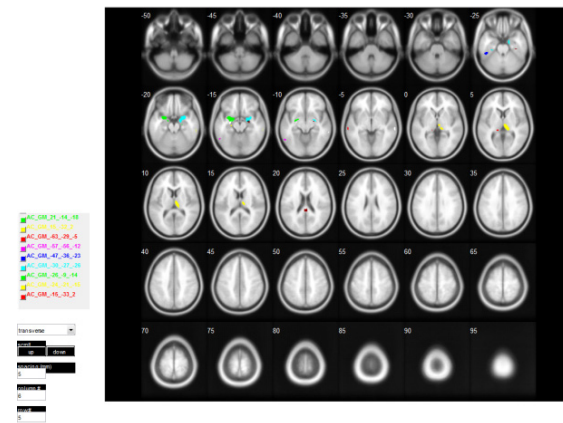**Figure 5.** GARD SPM *t*-map glass view for [NC AD] = [1 −1].

Figure 6. ROI_CL selected from the cluster of AD < CN (*p* < 0.001).

3.2.2. From AAL Template Based on the Anatomic Region

Each participant's brain differs in terms of the structure and size of the anatomic regions. This can be more problematic when anatomical ROIs are defined from a single participant and applied to the remaining data, as a significant anatomical inconsistency exists between the participants. Hence, we selected anatomic regions with a higher number of voxels in each cluster. After obtaining those regions, we generated a mask from the automated anatomic labeling (AAL) [34] atlas for those anatomic regions and co-registered with our *t*-map image. The AAL atlas is a digital human brain atlas with 116 labelled volumes indicating macroscopic brain structures. Hence, these masks can be used subsequently for feature extraction. The most significant regions of the brain based on the *t*-map image using AAL atlas labels are tabulated in Tables 2 and 3 for the ADNI and the GARD test conditions, respectively. From the tables, 14 anatomic regions were selected, and the standard ROIs were generated from the AAL atlas, as shown in Figure 7. We will term this ROI as ROI_AAL hereinafter. We applied the same weight-based mean calculation to obtain the mean features to represent each ROI value from all MRI, as in ROI_CL, a few demonstrated as in shown in the Appendix A, Table A2.
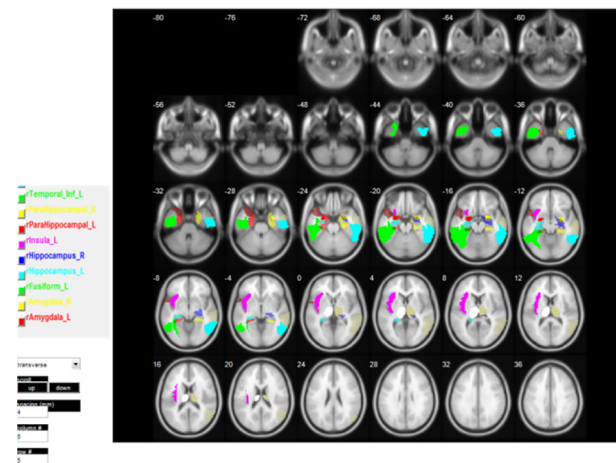

Figure 7. ROI_AAL selected from the cluster of AD < CN (*p* < 0.001).

Table 2. Result of ROI analysis for ADNI test condition.

ADNI Test Condition	AAL Region	Number of Voxels	Mean T	SD of T	Cluster Number
CN > AD (FWE $p < 0.001$)	Insula_L	508	3.5646	0.333	1
	Amygdala_L	256	7.602	0.6846	1
	Olfactory_L	107	4.3285	0.9015	1
	Amygdala_R	329	6.5708	0.9011	2
	ParaHippocampal_R	138	6.9469	0.5042	2
	Hippocampus_R	60	7.6221	0.3329	2
	Thalamus_R	1020	5.2492	0.6611	3
	Thalamus_L	981	4.4001	0.4153	3
	Hippocampus_L	13	6.2737	0.3692	4
	Temporal_Inf_L	1148	4.6223	0.6231	5
	Fusiform_L	490	3.9798	0.684	5
	Temporal_Inf_R	440	4.3544	0.6455	6
	Temporal_Mid_R	238	4.7228	0.6999	6
CN > MCI (FWE $p < 0.001$)	Hippocampus_L	53	5.5327	0.1259	1
	Amygdala_L	20	5.4556	0.1059	1
	Hippocampus_R	25	5.4315	0.0893	2
	Amygdala_R	24	5.5113	0.1407	2

Table 3. Result of ROI analysis for GARD test condition.

GARD Test Condition	AAL Region	Number of Voxels	Mean T	SD of T	Cluster Number
NC > AD (FWE $p < 0.001$)	Hippocampus_L	247	6.192	0.3964	1
	ParaHippocampal_L	212	6.4757	0.6251	1
	Amygdala_L	208	6.0849	0.2441	1
	Fusiform_L	103	6.4647	0.5286	1
	Temporal_Pole_Sup_L	90	6.2633	0.2723	1
	ParaHippocampal_R	179	6.4194	0.4744	2
	Hippocampus_R	145	6.7702	0.5669	2
	Amygdala_R	29	5.8039	0.0179	4
NC > MCI (no FWE $p < 0.001$)	Hippocampus_L	82	3.5717	0.2273	1
	Hippocampus_R	53	3.5466	0.2281	2
	Thalamus_R	19	3.4174	0.1511	2

3.3. Feature Extraction from Both Types of ROIs

Once the mask is generated from both ROI_CL and ROI_AAL, they are processed further for ROI analysis where each mask fuses with every image of each patient to generate only the important voxels. This voxel is higher in dimensions and is difficult to use as a feature. Consequently, the feature dimensionality is reduced with their intensity values averaged to obtain a mean intensity value. For each participant's MRI, voxels that are within the ROI are averaged to obtain a single representative value. The larger the ROI area, the higher is this value as we have used the weighted mean as in Equation (2), which becomes biased in the highly activated non-binary ROI. Although the ROI selection is a general and typical process, it can be a pivotal factor for obtaining good results. An efficient

SPM extension tool, MARSeille Boîte À Région d'Intérêt (MarsBaR) [35] was used for this feature extraction process:

$$x' = \frac{\sum_{i=1}^{n} x_i \times w_i}{\sum_{i=1}^{n} w_i} \tag{2}$$

Here, x_i represents the voxel intensity value for n voxels lying inside the ROI and w_i is the weight of the voxel. Here, w_i will have high values representing high confidence for voxels within the ROI, and w_i values are nearly zero representing low confidence in outlier voxels.

For VBM-based ROI, the selection was done manually based on cluster size, ROI_CL, for which we selected the cluster with the highest number of voxels, e.g., AD vs. CN voxels result from the ADNI dataset *t*-test yielded 30 clusters, of which we selected 14 clusters with significant voxel numbers, while rest small-sized clusters were neglected (please refer to the Supplementary Data file ADNI_AD_CN_p0.001_FWE_BrainLabels). Similarly, the remaining four clusters were selected from the AD vs. CN WM *t*-test and CN vs. MCI *t*-test (please refer to the Supplementary Data file ADNI_CN_MCI_p0.001_FWE_BrainLabels). For ROI_AAL the selected clusters were matched with the AAL template, to obtain significant anatomic regions, then those standard ROIs were used to create the mask for further voxel detection from each ROI, for feature detection. On initial matching 23 ROI from AD vs. CN (16 GM and seven WM), nine ROIs from AD vs. MCI, and 11 ROIs from CN vs. MCI were detected. We selected the one with the highest number of voxels from each cluster and the overlapped regions. In total there were 18 ROIs from the cluster and 14 ROIs from the AAL atlas, as shown in Figures 6 and 7, which were tested separately to obtain the best classification result.
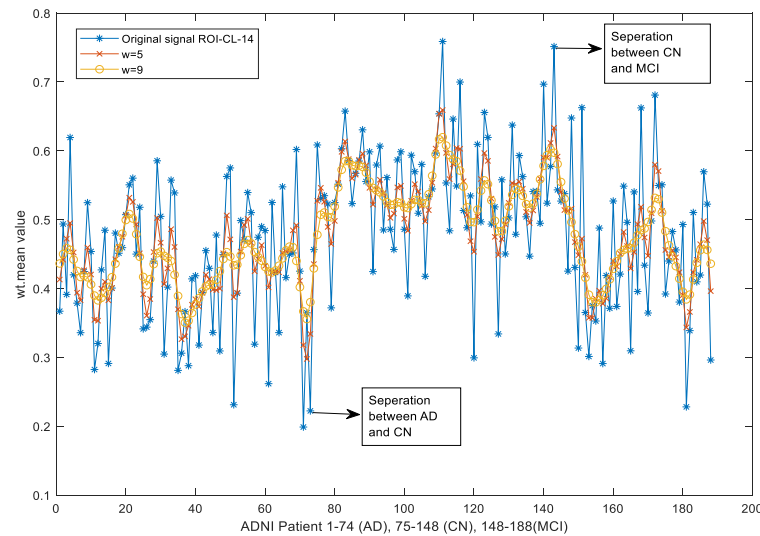
4. Gaussian Smoothing

The extracted features from each participant are the weighted mean intensity value of each ROI, which must be more homogeneous to their class and heterogeneous to others. However, upon careful observation as in Figure 8, the original data (blue) contains some highly unusual spikes at some point and can be considered abnormal. These spikes (x = 70, 110, 120, 144) were considered abnormal for the graph due to its extreme shifting from mean positioning (however this doesn't mean they are a less useful feature, each feature values depends on its cluster size, only in regard to the distribution we consider it as abnormal, which were averaged using Gaussian smoothing (GS) for accurate classification using BR-NN). In general, the Gaussian distribution for one-dimensional vector operates under the following equation:

$$G(x_{i,w}) = \frac{1}{\sqrt{2\pi}\sigma_{i,w}} e^{\frac{-x_i^2}{2\sigma_{i,w}^2}} \tag{3}$$

where x_i represents the input value for the *i*th feature of each image for each ROI, and $\sigma_{i,w}$ represents the standard deviation for each group of ROI, w is the size of the window over which the Gaussian weighted moving average is calculated. This window slides over the specified length, and provides the average Gaussian value over that window, as shown in Figure 8. With a Gaussian weighted moving window, data is transformed to its normal distribution after the completion of its window operation. Please note the window is an averaging moving function hence $\sigma_{i,w}$ also keeps on changing however we can approximate the final distribution to be a normal distribution for higher window size. In our case, we have chosen $w = 9$ from our experiment. In Figure 8, the right part separated from $x = 1$ to 74 represents the mean value for the AD patients and $x = 75$ to $x = 148$ represents the mean value of CN patients from one cluster of ROI_CL. Here, GS transforms the data from unknown haphazard distribution to the normal distribution, so that it can be classified with less computational burden using the Bayesian regularized neural network. It shifted the characteristics of each vector toward the mean value of its target class and bring homogeneity in its cohort. This can be considered as a data cleaning

step similar to the normalization process. Furthermore, it can work as a replacement for the feature selection step where the features are transformed for better classification results.

Figure 8. Data processing result before and after Gaussian-weighted smoothing average filter, window size, shown for 14th out of 18 features from ROI_CL. Blue colored curve is the raw data whereas orange and red curve represents the smoothened data for $w = 9$ and $w = 5$ respectively.

5. Bayesian Regularized Feed-Forward Neural Network (BR–FNN)

We used the Bayesian regularization [36] proposed by MacKay in 1992 for training the NN. This regularization technique minimizes the linear combination of squared errors and weights so that, at the end of the training the resulting network has good generalization qualities. Later in 1997 Foresee and Hagan [37] proposed optimization using the Levenberg–Marquardt (LM) algorithm in the BR-FNN to reduce computational overhead. It reduces the linear arrangement of squared errors and weights and finally produces a good, generalized network after iterative training. The Bayesian regularization takes place within the Levenberg-Marquardt algorithm. This can be summarized using the following equations.

For training the sets of the form $\{x_1, t_1\}, \{x_2, t_2\} \ldots \{x_n, t_n\}$, where $\{x_i, t_i\}$ represents the input value and the corresponding target for the *i*th term. The sum of squared error during training is represented by:

$$E_D = \sum_i^n (t_i - y_i)^2 \tag{4}$$

Here, t_i are the target values and y_i is the predicted response by the neural network during the process of training for *n* number of training inputs.

The objective function for Bayesian optimization and regulation exhibits a standard form as:

$$F(w) = \beta E_D + \alpha E_W \tag{5}$$

Here, E_W is the sum of squares of the network weights and the E_D error calculated from (4). Optimization parameters α and β are calculated as in Equations (6) and (7), respectively.

$$\alpha = \frac{\gamma}{2E_W(W)} \tag{6}$$

$$\beta = \frac{n - \gamma}{2E_D(W)} \tag{7}$$

Initially, $\alpha = 0$ and $\beta = 1$ such that $F(w) = E_D$, where γ is the effective number of parameters, calculated using the Gauss-Newton approximation to the Hessian matrix available in the LM training algorithm until final convergence is met [37,38]. We first randomly initialized a 2-layer feed-forward NN with the inputs connected to the first hidden layers and the second layer is the output layers with only two output values for binary classification. Here, the Nguyen-Widrow [39] method is used to initialize the weights and bias values for each hidden layer so that the active regions of the layer's neurons are distributed almost evenly over the input space. Hyperbolic tangent sigmoid [40] is used as an activation function in hidden layers to generate squashed outputs between $[-1 \; 1]$ for input vectors whereas in the output layer it is a linear purelin function. Random partition of training and testing (7:3) was performed manually on the whole set of features so that only the partitioned training set participates in training BR-NN, while the test set remains untouched. Later once the BR-NN model is trained, the test set is fed into the NN to find out the testing accuracy.

To obtain the optimal number of hidden layers, the network was trained for a different number of hidden layers starting from 1 to 20, initialized each time randomly. The classification performance parameter was tested iteratively ten times for each number of hidden layers such that the average performance is calculated each time; the resulting graph is shown in Figure 9. Here, training will yield us, 10 independent models, for each hidden layer i.e., when tested for up to 20 hidden layers in total $10 \times 20 = 200$ models were developed. Out of which every 10 models with the same number of hidden layers were passed iteratively for untouched test features, the average of ten iterations is then reported. Later the best performance was observed for 17 hidden layers. Table A3 in the appendix shows the result of each iteration. Subsequently, this network was used to verify the dataset from the GARD dataset with its mean ROI feature as the input.

The same process was applied to train the BR-NN from the GARD dataset features and was tested under the same condition for the ADNI dataset features. Both ROI_CL and ROI_AAL were used separately as features.
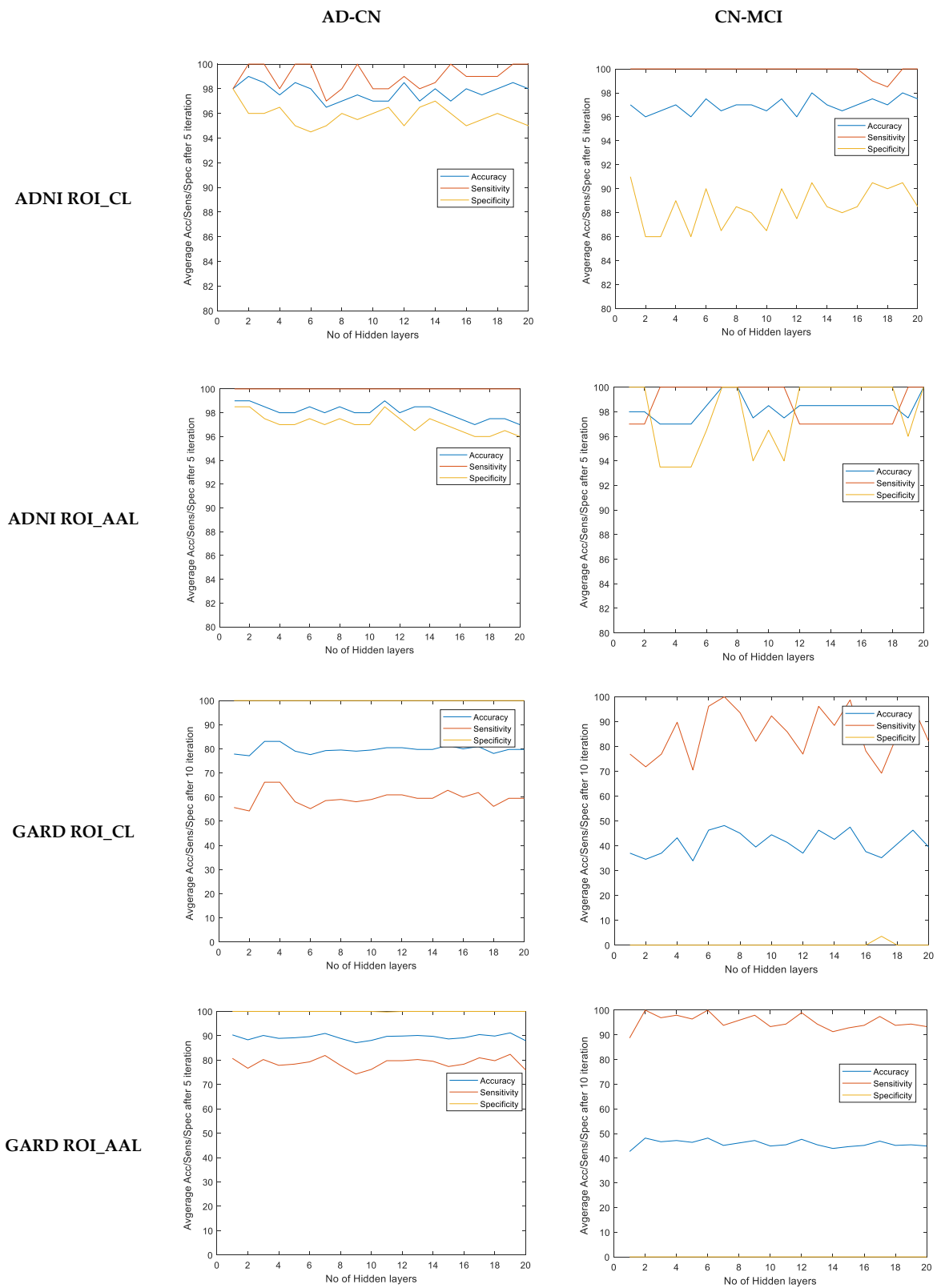
Figure 9. Testing results from the proposed method by training the ADNI dataset; here, the plotted curve is the average results by performing ten iterations for every value of n, i.e., the number of hidden layers.

6. Performance and Discussion

6.1. ROI Selection and Its Importance

Table 4 presents the clinical ROIs found on *t*-test group comparison between participating class domains; it lists the important brain regions identified using AAL matching. There were majorly six clusters to distinguish the voxel intensities, with Left Temporal Inferior Lobe, both Right and left Thalamus, left and right Amygdala, left Fusiform, left Insula, and right Hippocampus being the prominent ones for CN versus AD, at $p < 0.001$, whereas left and right Hippocampus, Temporal mid lobe, right Parahippocampal region were highlighted as dense cluster for CN versus MCI. Details of all the clusters and numbers of voxel detected are presented in Supplementary files. The ROIs identified using ADNI datasets are clearer and larger whereas GARD datasets are smaller and dimmer, (please see Figures 4 and 5) which may be due to the difference of MRI acquisition protocol [41], still the regions like the hippocampus, Amygdala, Temporal lobe, Thalamus are commonly deteriorated during AD/ADD and MCI/mAD (please see Tables 2 and 3, the T-value measures the size of the difference relative to the variation in the group data. A higher T-value supports our null hypothesis of group difference condition. SD of T represents the standard deviation measure of T-value and mean-T, the mean value of all T-values for voxels in its cluster). As it is believed that the loss of memory is mostly related to the limbic system including its subcortical structures like Hippocampus, Amygdala, more prominent in the left hemisphere [42]. However various cortical structures including frontal, Parietal, Temporal lobes are also affected during neurodegeneration [43,44]. Hence, it is wise to state that overall brain structures associated with memory are affected during the AD memory dysfunction phase.

Table 4. Voxel values for each AAL anatomic region from each cohort *t*-test of ADNI dataset for ROI_AAL selection. Bold represents the area from ROI_CL selected for ROI_AAL.

	Region from ROI_CL	# of Voxels	Mean T	# of Cluster	Selected Region for ROI_AAL (Alphabetically)
	Insula_L	508	3.5646	1	**Amygdala_L**
	Amygdala_L	256	7.602	1	**Amygdala_R**
	Amygdala_R	329	6.5708	2	**Fusiform_L**
	ParaHippocampal_R	138	6.9469	2	**Hippocampus_L**
	Hippocampus_R	60	7.6221	2	**Hippocampus_R**
	Olfactory_R	46	3.8332	2	**Insula_L**
GM based Region	Temporal_Pole_Sup_R	30	5.5589	2	**ParaHippocampal_L**
(AD vs. CN)	**Thalamus_R**	1020	5.2492	3	**ParaHippocampal_R**
	Thalamus_L	981	4.4001	3	**Temporal_Inf_L**
	Hippocampus_L	13	6.2737	4	**Temporal_Inf_R**
	Temporal_Inf_L	1148	4.6223	6	**Temporal_Mid_R**
	Fusiform_L	490	3.9798	6	**Temporal_Pole_Sup_L**
	Temporal_Mid_L	405	4.2446	6	**Thalamus_L**
	Temporal_Inf_R	440	4.3544	7	**Thalamus_R**
	Temporal_Mid_R	238	4.7228	7	

Table 4. *Cont.*

	Region from ROI_CL	# of Voxels	Mean T	# of Cluster	Selected Region for ROI _AAL (Alphabetically)
WM based Region (AD vs. CN)	**Thalamus_R**	124	4.0965	1	
	Hippocampus_R	79	4.7669	1	
	Thalamus_L	75	3.8945	1	
	Pallidum_R	27	4.1347	1	
	Caudate_R	26	3.6705	4	
	Putamen_R	24	3.7313	4	
GM based Region (AD vs. MCI)	Cingulum_Mid_R	64	2.7408	1	
	Insula_L	75	2.7419	2	
	Olfactory_R	61	2.6409	4	
	Temporal_Sup_R	42	2.5427	5	
	Lingual_R	79	2.5406	5	
	Insula_L	114	2.4783	6	
	Hippocampus_L	66	2.5187	10	
GM based Region (CN vs. MCI)	Temporal_Sup_R	1475	4.6274	1	
	Temporal_Mid_R	1310	4.5555	1	
	Hippocampus_R	873	5.9036	1	
	ParaHippocampal_R	730	4.6921	1	
	Temporal_Mid_L	1337	4.6585	1	
	Precuneus_L	888	4.8104	2	
	Fusiform_L	852	4.498	2	
	Hippocampus_L	667	5.6445	2	
	Temporal_Inf_L	505	4.6409	2	

6.2. Experimental Result

The classification performance of our proposed method was tested based on three parameters: accuracy, specificity, and sensitivity. Figure 10 displays the comparative classification performance results using the ADNI dataset trained model on the left part and the same parameter with the GARD trained dataset on the right part. Accuracy for ADNI trained dataset drops from 97% to 80% in the GARD test set in the case of ROI_CL based features, whereas the accuracy drop is not so severe in ROI_AAL based features. Similarly, accuracy from the GARD-trained dataset drops from 96% to 91% in the ADNI test set using ROI_CL features and from 97% to 88% in ROI_AAL features. Here all the results are from ADNI-trained BR-NN models; the first four results are for 30% test sets of ADNI, whereas the last four are for 100% of the GARD dataset used as the test set. To test the validation of the proposed method, some ablation experiments were also performed to conclude the need for smoothing and the importance of all features. Besides, this experiment compares the efficiency via time consumption of each experiment. The results of all performed experiments are shown in Table 5.
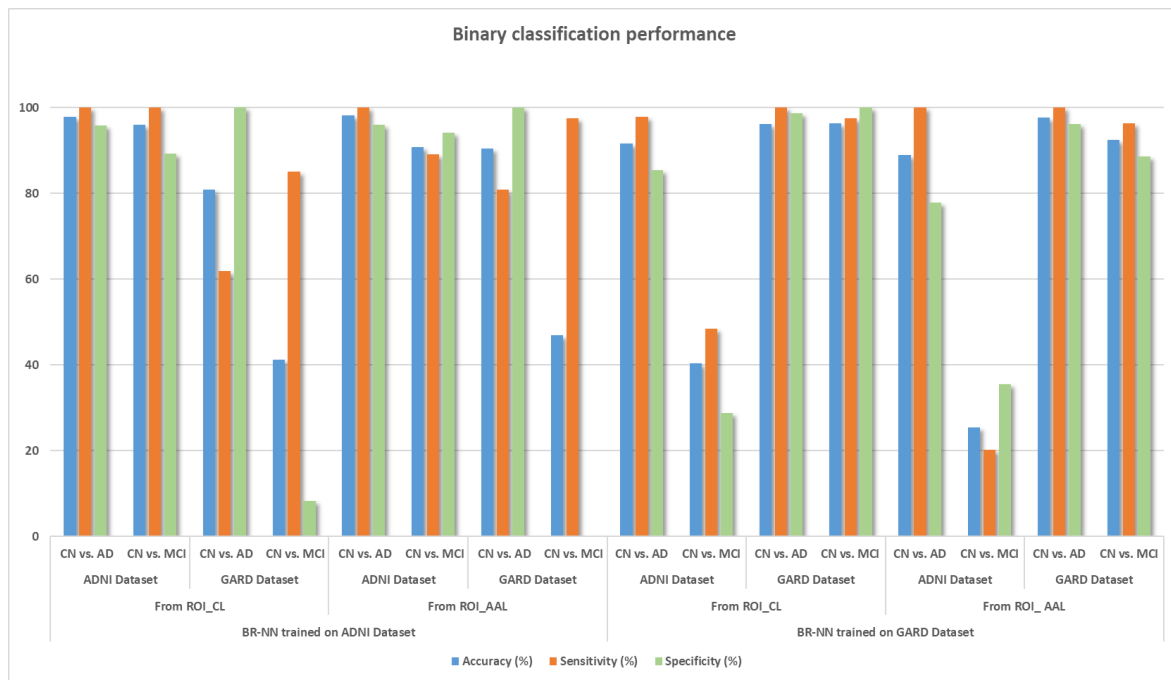
Figure 10. Classification test result from BR-NN trained on ADNI or GARD dataset. For ideal classification, the accuracy, sensitivity, and specificity are all 100%. Training on both dataset features would overfit the data and might not produce a convincing result for an independent test, hence the dataset was trained independently.

Table 5. Ablation experiments results. Please note the difference in the performance result with and without GS.

ADNI 70% Dataset Trained Experiment		Training Result: Proposed (ROI_CL+GS +BR-NN)	ROI_CL+BR-NN	Testing Result: ROI_CL+BR-NN)	50% Top Features ROI_CL+GS +BR-NN	Training Result Proposed ROI_AAL+GS +BR-NN	ROI_AAL +BR-NN	Testing Result: ROI_AA +BR-NN)	50% Top Features ROI_AAL+GS +BR-NN
ADNI Test (AD vs. CN)	ACC	97.78 ± 1.7	70.22 + 7.20	88.67 + 0.7	99.56 + 0.49	98.22 ± 0.99	62.89 + 4.45	78.22 + 2.78	98.51 + 1.48
	SEN	100 ± 0.00	66.4 + 7.11	90.87 + 1.37	100 + 0.00	100 ± 0.00	56 + 10.83	75 + 3.21	98.96 + 1.14
	SPE	95.83 ± 2.22	75 + 13.33	86.36 + 0.5	95.65 + 1.69	96 ± 2.24	71.5 + 8.18	77.39 + 4.49	100
ADNI Test (NC vs. MCI)	ACC	96 ± 1.56	77.14 + 10.26	72.86 + 4.31	98.57 + 2.02	90.86 ± 2.63	73.43 + 10.69	80.57 + 4.63	98 + 1.38
	SEN	100 ± 0.00	78.8 + 12.23	75.45 + 3.83	99.6 + 1.26	89.13 ± 2.29	82.4 + 8.68	82.31 + 6.83	98 + 0.52
	SPE	89.2 ± 4.21	73 + 8.23	68.46 + 6.74	96 + 3.16	94.17 ± 7.91	51 + 11.14	75.56 + 4.68	93 + 4.83
Total Training Time (s) AD vs. CN		319.0688	366.7428	319.0688	256.5436	287.4786	341.7982	287.4786	251.0212
GARD Test (ADD vs. NC)	ACC	80.95 ± 4.45	49.75 + 7.01	71.67 + 1.09	66.19 + 2.46	90.48 ± 2.45	54.64 + 4.64	75.6 + 0.84	90 + 2.82
	SEN	61.9 ± 8.91	33.85 + 9.35	57.38 + 2.62	64.76 + 1.00	80.93 ± 4.89	46.43 + 6.85	72.86 + 1.66	80.24 + 6.05
	SPE	100 ± 0.00	64.52 + 13.11	85.95 + 2.85	67.62 + 5.41	100 ± 0	62.86 + 8.49	78.33 + 1.76	99.76 + 0.75
GARD Test (NC vs. mAD)	ACC	41.23 ± 9.5	40.74 + 4.16	59.75 + 1.19	31.98 + 7.53	46.91 ± 2.14	41.11 + 3.99	42.35 + 3.03	39.38 + 6.20
	SEN	85.13 ± 8.6	66.15 + 15.66	31.54 + 2.11	7.18 + 7.13	97.44 ± 4.44	55.9 + 9.80	73.59 + 10.04	81.79 + 12.87
	SPE	8.3 ± 6.3	17.14 + 12.13	85.95 + 2.85	55 + 16.51	0.00 ± 0.0	27.38 + 8.49	13.33 + 5.04	0 + 0.00
Total Training Time (s) NC vs. MCI		304.7792	912.7463	304.7792	199.8685	245.374	324.8097	245.374	379.8156

6.3. Discussion

Our result suggests the proposed method shows good classification results for AD vs. CN (ADD vs. NC), and the network can be trained from either database, as the result is good in both cases. However, in the case of CN vs. MCI (NC vs. mAD), the result is satisfactory on the trainee dataset, but poor on the non-trainee dataset. This shows that the distinction between CN vs. MCI is unfavorable with the proposed method; further, on the fine analysis of feature, we found that the characteristics of the CN and MCI feature after Gaussian smoothing become indistinguishable in contrary to CN vs. AD, which may be the reason why the classifier cannot distinguish between them. In general AD

and CN are comparatively more discriminative than CN and MCI (this is our primary assumption), in the course of smoothing, the CN features tend more towards the CN mean value and so does AD features, so because of this, the MCI group (which lies between AD and CN) when smoothen loosens its MCI features and tends to follow either AD or CN features, depending on its feature weights. This is also highlighted by the result of the ROI_CL/ROI_AAL +BR-NN result column of Table 5, where the performance of all test conditions i.e., AD vs. CN and CN vs. MCI abruptly drops if Gaussian smoothing is not mediated before classification. Similarly, the test with half of the features results in poor accuracy for the non-trainee dataset i.e., GARD classification.

Cuingnet et al. [2] have shown in their experiments that the voxel-based direct method (similar to ROI_CL) and the atlas-based method (similar to ROI_AAL) for AD vs. CN independent test in ADNI outperformed cortical thickness and hippocampus volume-based classification. Also, methods using the whole brain reached specificity over 90% than those based on the hippocampus. Besides in prodromal AD detection (CN vs. MCIc), the sensitivity was substantially lower as in our case. In comparison to the cortical thickness and hippocampus analysis, VBM-based methods are simple and direct, SPM software can be used to generate tissue probability maps without extra parcellation (with parcellation error during registration) as required by the Freesurfer software (https://surfer.nmr.mgh.harvard.edu/ (accessed on 21 February 2021), which accounts for heavy time consumption and hardware operation. Similarly, contrary to deep neural networks like CNN, which are beneficial for standalone classification purposes and claiming the state-of-the-art performance in image classification. However, considering very few CNN architectures being designed for 3D MRI scans [45] and the significances of ROIs analysis as discussed in Sections 3.2 and 6.1 we were interested in VBM. In VBM, we follow an established preprocessing pipeline to extract and visualize the group cluster difference for generating ROIs of significant regions, hence is more specific in voxel-based group statistics analysis, whereas CNN is a gradient-loss supervised algorithm for image-attributes based feature extraction process primarily focused on classification rather than ROI detection.

6.4. Comparison with State-of-the-Art Performance

Table 6 shows a brief comparison with other state-of-the-art performances. Zhang et al. [46] used a multimodal imaging technique combining MRI, CSF, and PET data for classification and reported up to 86.2% accuracy. Westman et al. [47] combined MRI and CSF measurements to classify AD with NC using Freesurfer-extracted cortical thickness and subcortical volumes. Aguilar et al. [48] used the same features for classification based on the clinical dementia rating (CDR) score. Beheshti et al. [7] used the VBM extracted VOI features of the voxel and PDF for feature selection, which was classified using the SVM classifier. Subsequently, Beheshti et al. [8] reported a higher accuracy of 92.48% using feature-ranking techniques to select the discriminative top feature instead of a feature selection technique. The most used dataset is ADNI, whereas few researchers have used OASIS [49] to report the accuracy. Most of the stated accuracy is from cross-validation (CV). Conversely, we have reported average accuracy results from the test set. We preferred a completely untouched test set so that the neural network model is trained only using training features (Figure 2). Besides, it will also report a generalized cum unbiased performance parameter. However, for verification we also tested 5-fold CV on the whole feature set of ADNI and GARD datasets, to be accuracy equivalent to 98% and 97% respectively, however, the most important accuracy is in other untouched datasets, which cannot be calculated using k-fold. Besides, in contrast to single-step feature extraction from voxels of *t*-map, we have implemented multistep AAL based masking (ROI_AAL) and cluster-based selection of ROI (ROI_CL), then used the selected voxels as features (results in Supplementary Data). This mask creation steps from AAL and cluster helps to identify the atrophic anatomic regions of the brain, signifying the probable area of interest for AD/ADD detection. All these methods are still progressively being updated and modified

for better classification performance. Our reported accuracy of 98.22% for CN vs. AD is the highest hitherto using VBM features on the ADNI database. Further, the accuracy of 90.48% in a completely different dataset, i.e., GARD indicated that it has generalized well.

Table 6. Comparison of the proposed method with other methods. * Number of hidden layers = 17, trained on 70% of ADNI dataset, Test set: 30% ADNI, 100% GARD. Abbreviations: Orthogonal Partial Least Squares (OPLS), Multivariate analysis (MVA), SVM Recursive Feature Elimination (SVM-RFE), Feedforward Neural Network (FNN), Multilayer Perceptron (MLP).

Author	Biomarker	Dataset	Classifier	AD/HC	Validation Method	ACC (%)	SEN (%)	SPE (%)
Zhang et al. [46]	MRI	ADNI	SVM	51/52	Ten-fold	86.2	86	86.3
Westman et al. [47]	MRI	ADNI	OPLS MVA	96/111	Ten-fold	87	83.3	90.193
Aguilar et al. [48]	MRI	ADNI	SVM, OPLS	116/110	Ten-fold	84.9	80.2	90.88
Zhou et al. [50]	MRI	ADNI	SVM	59/127	Two-fold	78.2	68.5	81.7
Papakostaset al. [51]	MRI	OASIS	KNN	49/49	Ten-fold	85	78	92
Khedher et al. [52]	MRI	ADNI	SVM	188/229	Ten-fold	88.49	85.11	91.27
Iman Behesti et al. [7]	MRI	ADNI	SVM	130/130	Ten-fold	89.65	87.73	91.57
Iman Behesti et al. [8]	MRI	ADNI	SVM	130/130	Ten-fold	92.48	91.07	93.89
Yi Ding et al. [53]	MRI	ADNI	SVM-RFE	54/58	Five-fold	92.86	87.04	98.28
Savio et al. [4]	MRI	OASIS	NN	49/49	Ten-fold	86	80	92
Jha et al. [9]	MRI	OASIS	FNN	28/98	Five-fold	90.06 ± 0.01	92.00 ± 0.04	87.78 ± 0.04
Wang et al. [10]	MRI	OASIS	MLP	28/98	Five-fold	92.40 ± 0.83	92.14 ± 4.39	92.47 ± 1.23
Proposed Method	MRI	ADNI (ROI_CL)	BR-NN	74/74	Average of ten iterations *	97.78 ± 1.7	100.00 ± 0.00	95.83 ± 2.22
	MRI	GARD (ROI_CL)	BR-NN	42/42	Average of ten iterations *	80.95 ± 4.45	61.9 ± 8.91	100.00 ± 0.00
	MRI	ADNI (ROI_AAL)	BR-NN	74/74	Average of ten iterations *	98.22 ± 0.99	100	96 ± 2.24
	MRI	GARD (ROI_AAL)	BR-NN	42/42	Average of ten iterations *	90.48 ± 2.45	71.43 ± 6.30	100

7. Conclusions

We present herein an efficient approach to extract mean weighted features from the *t*-map obtained after VBM processing and then apply GS on the raw data. This will help to distribute the features into a Gaussian distribution such that the classifier does not require many operations to identify the distinguishing feature. Using an excessive number of training materials could over-train the network and make it prone to over-fitting, so few training materials were used for optimal training of the classifier model, which is also one supportive idea for generalization. The use of the BR-NN helps to generalize the network. Hence, we presented an optimized approach for the classification of an AD/ADD from the CN/NC based on VBM ROI extracted-smoothed features. Our uniqueness lies in the selection of ROIs and the extraction of ROI features. Unlike using all the voxel of ROI generated from the *t*-test of VBM, we have used only the mean weighted value from the entire ROIs voxels thus reducing feature and applying feature selection at once. Besides, using AAL maps generated standard co-registered ROIs, priorities based on *t*-map generated clusters of each anatomic region for mean weighted feature extraction are firstly implemented by us as per our knowledge. To improve the result, we have suggested Gaussian smoothing on a small window, which may also be the new attempt on post-processing. Final classification based on Bayesian regularized NN finds this distribution computationally easier to classify improving the overall performance accuracy of the system. Additionally, we have experimented with two datasets in reference to each other to validate the process. To summarize, our proposed method to optimize and generalize

the AD/ADD detection using T1-MRI yields significant improvement over the previously used VBM based method, at the same time, our findings of brain ROIs can help clinicians to focus on those specific regions for further future discovery. Hence, we hope our attempt will have some meaningful contribution in the future.

(accessed on 12 May 2021)). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California. Correspondence should be addressed to Goo-Rak Kwon: grkwon@chosun.ac.kr.

Appendix A

Table A1. ADNI-based ROI_CL feature sample data for first three MRIs from each cohort, read as AC_GM cluster at [−46.5 −36.0 −22.5] AC meaning AD vs. CN else CM meaning CN vs. MCI, GM stated the obtained ROI is from gray matter else WM meaning White matter, coordinates indicate its center location.

Location/	AC_GM cluster at [−15.0 −33.0 1.5]	AC_GM cluster at [−24.0 −21.0 −15.0]	AC_GM cluster at [−25.5 −9.0 −13.5]	AC_GM cluster at [−30.0 −27.0 −25.5]	AC_GM cluster at [−46.5 −36.0 −22.5]	AC_GM cluster at [−57.0 −55.5 −12.0]	AC_GM cluster at [−63.0 −28.5 −4.5]	AC_GM cluster at [15.0 −31.5 1.5]	AC_GM cluster at [21.0 −13.5 −18.0]	AC_GM cluster at [25.5 −9.0 −13.5]	AC_GM cluster at [30.0 −27.0 −25.5]	AC_GM cluster at [37.5 −22.5 −27.0]	AC_GM cluster at [64.5 −24.0 −6.0]	AC_GM cluster at [64.5 −31.5 −16.5]	AC_WM cluster at [1.5 −36.0 21.0]	CM_GM cluster at [−27.0 −10.5 −13.5]	CM_GM cluster at [13.5 −31.5 1.5]	CM_GM cluster at [27.0 −9.0 −15.0]	Target class/Label
MR1	0.3093	0.2374	0.3281	0.2701	0.3236	0.2851	0.3153	0.4931	0.3490	0.3527	0.3216	0.3242	0.3461	0.3991	0.1228	0.2657	0.3673	0.3670	**AD**
MR2	0.3925	0.5445	0.4526	0.5315	0.3892	0.4626	0.3869	0.4416	0.5676	0.4565	0.5105	0.4238	0.3969	0.3980	0.1114	0.5032	0.3712	0.4937	**AD**
MR3	0.3253	0.3617	0.3490	0.3129	0.4553	0.4456	0.4495	0.3161	0.4353	0.3725	0.3614	0.3375	0.3813	0.3794	0.1176	0.3147	0.3229	0.3914	**AD**
MR75	0.5074	0.5122	0.5295	0.5587	0.4627	0.5046	0.4328	0.5831	0.6103	0.5532	0.5588	0.4860	0.4192	0.4568	0.1192	0.5566	0.4118	0.6085	**CN**
MR76	0.4533	0.4398	0.4810	0.4092	0.5092	0.4700	0.4434	0.5087	0.5012	0.5178	0.4451	0.4844	0.4669	0.5073	0.1567	0.4826	0.3516	0.5306	**CN**
MR77	0.5358	0.4974	0.4879	0.5012	0.4016	0.3805	0.3538	0.5088	0.6036	0.5197	0.4281	0.3673	0.3385	0.3621	0.1126	0.4999	0.4863	0.5348	**CN**
MR149	0.2897	0.4256	0.4016	0.4596	0.4078	0.3795	0.3595	0.4271	0.4182	0.4126	0.4333	0.3582	0.2984	0.3184	0.0808	0.3981	0.2601	0.4306	**MCI**
MR150	0.2912	0.3234	0.4049	0.4397	0.3846	0.3320	0.2157	0.4364	0.4046	0.3805	0.4556	0.5041	0.1985	0.3619	0.0941	0.3287	0.3490	0.3135	**MCI**
MR151	0.3904	0.6184	0.4838	0.4348	0.3609	0.3279	0.3592	0.5221	0.7012	0.6146	0.5170	0.4805	0.3935	0.3767	0.1603	0.5817	0.4235	0.6626	**MCI**

Table A2. ADNI-based ROI_AAL feature sample data for first three MRIs from each cohort, read as rAmygdala_L_f_img, r meaning the region of Amygdala, _L indicating Left region, else R indicating Right region, img for image. Please refer [34] for the coordinate location of each ROIs.

Location	rAmygdala_L_f_img	rAmygdala_R_f_img	rFusiform_L_f_img	rHippocampus_L_f_img	rHippocampus_R_f_img	rInsula_L_f_img	rParaHippocampal_L_f_img	rParaHippocampal_R_f_img	rTemporal_Inf_L_f_img	rTemporal_Inf_R_f_img	rTemporal_Mid_R_f_img	rTemporal_Pole_Sup_L_f_img	rThalamus_L_f_img	rThalamus_R_f_img	Target class/Label
MR1	0.3063	0.3209	0.2122	0.1769	0.2674	0.3203	0.2141	0.2677	0.2654	0.2794	0.2672	0.2041	0.3475	0.3533	**AD**
MR2	0.4729	0.3926	0.3446	0.3796	0.3473	0.4205	0.4001	0.3818	0.3550	0.3068	0.3186	0.2946	0.3656	0.3195	**AD**
MR3	0.3469	0.3264	0.2977	0.2506	0.2616	0.3823	0.2678	0.3069	0.3429	0.3112	0.3060	0.2353	0.2274	0.2357	**AD**
MR75	0.5721	0.4882	0.3812	0.3908	0.4046	0.4206	0.4166	0.4157	0.3927	0.3580	0.3530	0.3380	0.4459	0.4360	**CN**
MR76	0.5252	0.4617	0.3808	0.3626	0.3669	0.3988	0.3864	0.3987	0.4061	0.3820	0.3699	0.3246	0.3908	0.3738	**CN**
MR77	0.5287	0.4413	0.3150	0.3775	0.3726	0.4020	0.4223	0.4084	0.2933	0.2943	0.2772	0.3368	0.3458	0.3355	**CN**
MR149	0.4055	0.3584	0.2730	0.3055	0.2877	0.3131	0.3126	0.3173	0.2794	0.2640	0.2800	0.2409	0.3247	0.3313	**MCI**
MR150	0.4358	0.3292	0.2948	0.2669	0.2827	0.3610	0.4250	0.4012	0.3014	0.2882	0.2119	0.4619	0.2773	0.2914	**MCI**
MR151	0.5278	0.5028	0.3238	0.4351	0.4315	0.4136	0.4195	0.4683	0.3103	0.3312	0.3073	0.2985	0.4123	0.4060	**MCI**

Table A3. Results of validation and testing of performance on BR-NN trained on 70% ADNI extracted features. Validation was performed in remaining 30% ADNI and tested in 100% GARD features. Iteration result is mentioned as [accuracy sensitivity specificity] highest being [1 1 1] for ideal case.

Test Condition	Features	BR-NN Hidden Layer=17	Iteration 1	Iteration 2	Iteration 3	Iteration 4	Iteration 5	Iteration 6	Iteration 7	Iteration 8	Iteration 9	Iteration 10
AD vs. CN	ROI_CL	Validation (ADNI 30%)	[0.97 1 0.95]	[1 1 1]	[1 1 1]	[0.97 1 0.96]	[1 1 1]	[1 1 1]	[1 1 1]	[1 1 1]	[0.98 1 0.94]	[0.98 1 0.93]
		Testing (GARD 100%)	[0.80 0.61 1]	[0.82 0.64 1]	[0.77 0.54 1]	[0.83 0.66 1]	[0.78 0.57 1]	[0.91 0.83 1]	[0.91 0.83 1]	[0.78 0.57 1]	[0.83 0.66 1]	[0.80 0.61 1]
	ROI_AAL	Validation (ADNI 30%)	[1 1 1]	[1 1 1]	[1 1 1]	[1 1 1]	[0.97 1 0.95]	[1 1 1]	[1 1 1]	[1 1 1]	[0.95 1 0.90]	[0.97 1 0.95]
		Testing (GARD 100%)	[0.84 0.69 1]	[0.88 0.76 1]	[0.90 0.80 1]	[0.85 0.71 1]	[0.89 0.78 1]	[0.97 0.95 1]	[0.94 0.88 1]	[0.96 0.92 1]	[0.88 0.76 1]	[0.85 0.71 1]
CN vs. MCI	ROI_CL	Validation (ADNI 30%)	[1 1 1]	[0.97 1 0.91]	[1 1 1]	[1 1 1]	[1 1 1]	[0.91 1 0.84]	[1 1 1]	[1 1 1]	[1 1 1]	[0.94 1 0.91]
		Testing (GARD 100%)	[0.39 0.82 0]	[0.48 1 0]	[0.38 0.79 0]	[0.44 0.92 0]	[0.48 1 0]	[0.38 0.79 0]	[0.39 0.82 0]	[0.46 0.97 0]	[0.48 1 0]	[0.41 0.87 0]
	ROI_AAL	Validation (ADNI 30%)	[1 1 1]	[0.94 1 0.875]	[1 1 1]	[1 1 1]	[0.94 1 0.85]	[1 1 1]	[1 1 1]	[1 1 1]	[0.97 1 0.92]	[0.97 1 0.92]
		Testing (GARD 100%)	[0.43 0.89 0]	[0.44 0.92 0]	[0.43 0.89 0]	[0.41 0.87 0]	[0.48 1 0]	[0.40 0.84 0]	[0.40 0.84 0]	[0.46 0.97 0]	[0.35 0.74 0]	[0.29 0.61 0]

References

1. Guo, H.; Zhang, F.; Chen, J.; Xu, Y.; Xiang, J. Machine Learning Classification Combining Multiple Features of a Hyper-Network of fMRI Data in Alzheimer's Disease. *Front. Neurosci. Orig. Res.* **2017**, *11*, 615. [CrossRef]

2. Cuingnet, R.; Gerardin, E.; Tessieras, J.; Auzias, G.; Lehéricy, S.; Habert, M.O.; Chupin, M.; Benali, H.; Colliot, O.; Alzheimer's Disease Neuroimaging Initiative. Automatic classification of patients with Alzheimer's disease from structural MRI: A comparison of ten methods using the ADNI database. *NeuroImage* **2011**, *56*, 766–781. [CrossRef] [PubMed]

3. Ashburner, J.; Friston, K.J. Voxel-Based Morphometry—The Methods. *NeuroImage* **2000**, *11*, 805–821. [CrossRef]

4. Savio, A.; Sebastian, M.G.; Chyzyk, D.; Hernandez, C.; Graña, M.; Sistiaga, A.; de Munain, A.L.; Villanua, J. Neurocognitive disorder detection based on feature vectors extracted from VBM analysis of structural MRI. *Comput. Biol. Med.* **2011**, *41*, 600–610. [CrossRef] [PubMed]

5. Busatto, G.E.; Garrido, G.E.J.; Almeida, O.P.; Castro, C.C.; Camargo, C.H.; Cid, C.G.; Buchpiguel, C.A.; Furuie, S.; Bottino, C.M. A voxel-based morphometry study of temporal lobe gray matter reductions in Alzheimer's disease. *Neurobiol. Aging* **2003**, *24*, 221–231. [CrossRef]

6. Frisoni, G.B.; Testa, C.; Zorzan, A.; Sabattoli, F.; Beltramello, A.; Soininen, H.; Laakso, M.P. Detection of grey matter loss in mild Alzheimer's disease with voxel based Morphometry Journal of Neurology. *Neurosurg. Psychiatry* **2002**, *73*, 657–664. [CrossRef]

7. Beheshti, I.; Demirel, H. Probability distribution function-based classification of structural MRI for the detection of Alzheimer's disease. *Comput. Biol. Med.* **2015**, *64*, 208–216. [CrossRef] [PubMed]

8. 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]

9. Jha, D.; Kim, J.I.; Kwon, G.R. Diagnosis of Alzheimer's disease using dual-tree complex wavelet transform, PCA, and feed-forward neural network. *J. Healthc. Eng.* **2017**, *2017*, 9060124. [CrossRef]

10. Wang, S.H.; Zhang, Y.; Li, Y.; Jia, W.; Yang, W.; Zhang, Y. Single slice based detection for Alzheimer's disease via wavelet entropy and multilayer perceptron trained by biogeography-based optimization. *Multimed. Tools Appl.* **2016**, *77*, 10393–10417. [CrossRef]

11. Zhang, Y.; Wang, S.; Dong, Z. Classification of Alzheimer disease based on structural magnetic resonance imaging by kernel uspport vector machine decision tree. *Prog. Electromagn. Res.* **2014**, *144*, 171–184. [CrossRef]

12. Zhang, Y.; Wu, L.; Wang, S. Magnetic resonance brain image classification by an improved artificial bee colony algorithm. *Prog. Electromagn. Res.* **2011**, *116*, 65–79. [CrossRef]

13. Neffati, S.; Ben Abdellafou, K.; Jaffel, I.; Taouali, O.; Bouzrara, K. An improved machine learning technique based on downsized KPCA for Alzheimer's disease classification. *Int. J. Imaging Syst. Technol.* **2018**, *29*, 121–131. [CrossRef]

14. Suk, H.; Lee, S.W.; Shen, D. Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis. *NeuroImage* **2014**, *101*, 569–582. [CrossRef]

15. Amoroso, N.; Diacono, D.; Fanizzi, A.; La Rocca, M.; Monaco, A.; Lombardi, A.; Guaragnella, C.; Bellotti, R.; Tangaro, S. Deep learning reveals Alzheimer's disease onset in MCI subjects: Results from an international challenge. *J. Neurosci. Methods* **2018**, *302*, 3–9. [CrossRef]

16. Khagi, B.; Kwon, G.R.; Lama, R. Comparative analysis of Alzheimer's disease classification by CDR level using CNN, feature selection, and machine-learning techniques. *Int. J. Imaging Syst. Technol.* **2019**, *29*, 297–310. [CrossRef]

17. Shuai, B.; Zuo, Z.; Wang, B.; Wang, G. Dag-recurrent neural networks for scene labeling. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 3620–3629.

18. Yu, L.; Fan, G. DrsNet: Dual-resolution semantic segmentation with rare class-oriented superpixel prior. *Multimed. Tools Appl.* **2020**, *80*, 1687–1706. [CrossRef]

19. Chincarini, A.; Bosco, P.; Gemme, G.; Morbelli, S.; Arnaldi, D.; Sensi, F.; Solano, I.; Amorosco, N.; Tangaro, S.; Longo, R.; et al. Alzheimer's disease markers from structural MRI and FDG-PET brain images. *Eur. Phys. J. Plus* **2012**, *127*, 135. [CrossRef]

20. Shaikh, T.A.; Ali, R. Automated atrophy assessment for Alzheimer's disease diagnosis from brain MRI images. *Magn. Reson. Imaging* **2019**, *62*, 167–173. [CrossRef]

21. 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]

22. 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]

23. Fernández, A.; Hornero, R.; Mayo, A.; Poza, J.; Gil-Gregorio, P.; Ortiz, T. MEG spectral profile in Alzheimer's disease and mild cognitive impairment. *Clin. Neurophysiol.* **2006**, *117*, 306–314. [CrossRef]

24. Wang, K.; Liang, M.; Wang, L.; Tian, L.; Zhang, X.; Li, K.; Jiang, T. Altered functional connectivity in early Alzheimer's disease: A resting-state fMRI study. *Hum. Brain Mapp.* **2007**, *28*, 967–978. [CrossRef]

25. Available online: https://ida.loni.usc.edu/home/projectPage.jsp?project=ADNI&page=HOME&subPage=OVERVIEW_PR (accessed on 10 February 2021).

26. Hajnal, J.V.; Hill, D.L.G.; Hawkes, D.J. *Medical Image Registration*; CRC Press: New York, NY, USA, 2001.

27. Jovicich, J.; Czanner, S.; Greve, D.; Haley, E.; van der Kouwe, A.; Gollub, R.; Kennedy, D.; Schmitt, F.; Brown, G.; Macfall, J.; et al. Reliability in multi-site structural MRI studies: Effects of gradient non-linearity correction on phantom and human data. *Neuroimage* **2006**, *30*, 436–443. [CrossRef]

28. Narayana, P.A.; Brew, W.W.; Kulkarni, M.V.; Sievenpiper, C.L. Compensation for surface coil sensitivity variation in magnetic resonance imaging. *Magn. Reson. Imaging* **1998**, *6*, 271–274. [CrossRef]

29. Jack, C.R., Jr.; Bernstein, M.A.; Fox, N.C.; Thompson, P.; Alexander, G.; Harvey, D.; Borowski, B.; Britson, P.J.L.; Whitwell, J.; Ward, C.; et al. The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. *J. Magn. Reson. Imaging Off. J. Int. Soc. Magn. Reson. Med.* **2008**, *27*, 685–691. [CrossRef]

30. Bladowska, J.; Biel, A.; Zimny, A.; Lubkowska, K.; Bednarek-Tupikowska, G.; Sozański, T.; Zaleska-Dorobisz, U.; Sąsiadek, M. Are T2-weighted images more useful than T1-weighted contrast-enhanced images in assessment of postoperative sella and parasellar region? *Med. Sci. Monit. Int. Med. J. Exp. Clin. Res.* **2011**, *17*, MT83–MT90. [CrossRef] [PubMed]

31. Ashburner, J. A fast diffeomorphic image registration algorithm. *NeuroImage* **2007**, *38*, 95–113. [CrossRef]

32. Mazziotta, J.; Toga, A.; Evans, A.; Fox, P.; Lancaster, J.; Zilles, K.; Woods, R.; Paus, T.; Simpson, G.; Pike, B.; et al. A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). *Philos. Trans. R. Soc. Lond. Ser. B* **2001**, *356*, 1293–1322. [CrossRef] [PubMed]

33. Matthew, B.; Penny, W.; Kiebel, S. Introduction to Random Field Theory. *Hum. Brain Funct.* **2003**, *2*, 867–879.

34. Tzourio-Mazoyer, N.; Landeau, B.; Papathanassiou, D.; Crivello, F.; Étard, O.; Delcroix, N.; Mazoyer, B.; Joliot, M. Automated Anatomical Labeling of Activations in SPM Using a Macroscopic Anatomical Parcellation of the MNI MRI Single-Subject Brain. *NeuroImage* **2002**, *15*, 273–289. [CrossRef]

35. Brett, M.; Anton, J.L.; Valabregue, R.; Poline, J.B. Region of interest analysis using an SPM toolbox [abstract]. In Proceedings of the 8th International Conference on Functional Mapping of the Human Brain, Sendai, Japan, 2–6 June 2002.

36. MacKay, D.J. Bayesian interpolation. *Neural. Comput.* **1992**, *4*, 415–447. [CrossRef]

37. Foresee, F.D.; Hagan, M.T. Gauss-Newton approximation to Bayesian learning. In Proceedings of the 1997 International Joint Conference on Neural Networks, Houston, TX, USA , 12 June 1997; IEEE: Piscataway, NJ, USA, 1997; Volume 3, pp. 1930–1935.

38. Hagan, M.T.; Demuth, H.B.; Beale, M.H.; De Jesús, O. *Neural Network Design*; Pws Pub: Boston, MA, USA, 1996.

39. Nguyen, D.; Widrow, B. Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights. In Proceedings of the 1990 IJCNN International Joint Conference on Neural Networks, San Diego, CA, USA, 17–21 June 1990; pp. 21–26.

40. Vogl, T.P.; Mangis, J.K.; Rigler, A.K.; Zink, W.T.; Alkon, D.L. Accelerating the convergence of the back-propagation method. *Biol. Cybern.* **1988**, *59*, 257–263. [CrossRef]

41. Baert, A.L.; Günther, R.W.; von Schulthess, G.K. *Interventional Magnetic Resonance Imaging*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2012.

42. Amoroso, N.; La Rocca, M.; Bruno, S.; Maggipinto, T.; Monaco, A.; Bellotti, R.; Tangaro, S. Multiplex networks for early diagnosis of Alzheimer's disease. *Front. Aging Neurosci.* **2018**, *10*, 365. [CrossRef] [PubMed]

43. Available online: https://courses.lumenlearning.com/wsu-sandbox/chapter/parts-of-the-brain-involved-with-memory/ (accessed on 1 September 2020).

44. Rugg, M.D.; Yonelinas, A.P. Human recognition memory: A cognitive neuroscience perspective. *Trends Cogn. Sci.* **2003**, *7*, 313–319. [CrossRef]

45. Khagi, B.; Kwon, G.R. *3D CNN Design for the Classification of Alzheimer's Disease Using Brain MRI and PET*; IEEE Access: Piscataway, NJ, USA, 2020.

46. 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] [PubMed]

47. 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] [PubMed]

48. Aguilar, C.; Westman, E.; Muehlboeck, J.S.; Mecocci, P.; Vellas, B.; Tsolaki, M.; Kloszewska, I.; Soininen, H.; Lovestone, S.; Spenger, C.; et al. Different multivariate techniques for automated classification of MRI data in Alzheimer's disease and mild cognitive impairment. *Psychiatry Res. Neuroimaging* **2013**, *212*, 89–98. [CrossRef] [PubMed]

49. Available online: https://www.oasis-brains.org/#data (accessed on 10 February 2020).

50. 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]

51. 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]

52. Khedher, L.; Ramírez, J.; Górriz, J.M.; Brahim, A.; Segovia, F.; Alzheimer's Disease Neuroimaging Initiative. 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]

53. Ding, Y.; Zhang, C.; Lan, T.; Qin, Z.; Zhang, X.; Wang, W. Classification of Alzheimer's disease based on the combination of morphometric feature and texture feature, in Bioinformatics and Biomedicine (BIBM). In Proceedings of the 2015 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Washington, DC, USA, 9–12 November 2015; pp. 409–412.