

A Novel Approach for Early detection of Alzheimer's disease Based on Multi Level Fuzzy Neural Networks

Hamid Akramifard ¹, Mohammad Ali Balafar ¹, Seyed Naser Razavi ¹, Abd Rahman Ramli ²

¹. Faculty of Electrical and Computer Engineering, University of Tabriz, East Azerbaijan, Tabriz, Iran

². Faculty of Engineering, University Putra Malaysia, Selangor, Malaysia

Abstract: Timely diagnosis of Alzheimer's diseases (AD) is crucial to obtain more practical treatments. In this paper, a novel approach Based on Multi-Level Fuzzy Neural Networks (MLFNN) for early detection of AD is proposed. The focus of study was on the problem of diagnosing AD and MCI patients from healthy people using MLFNN and selecting the best feature(s) and most compatible classification algorithm. In this way, we achieve an excellent performance using only a single feature i.e. MMSE score, by fitting the optimum algorithm to the best area using optimum possible feature(s) namely one feature for a real life problem. It can be said, the proposed method is a discovery that help patients and healthy people get rid of painful and time consuming experiments. Experiments shows the effectiveness of proposed method in current research for diagnosis of AD with one of the highest performance (accuracy rates of 96.6%), ever reported in the literature.

Keywords: Alzheimer's disease; classification; early detection; Multi-Level Fuzzy Neural Networks; prognosis.

1. Introduction

Alzheimer's Diseases (AD) is one of the brain disabilities which, is common among the elderly patients. It starts with forgetting and losing memory-which causes cognitive functions become disturbed, and then lose their ability to remember recent events. Moreover, it can be one of the major causes of death in not so far future. Timely diagnosis can help to reduce its casualties. Many biomarkers are used to diagnosis of AD, such as, Magnetic Resonance Imaging (MRI) (Cuingnet et al., 2011; Davatzikos, Fan, Wu, Shen, & Resnick, 2008; Kloppel et al., 2008), Positron Emission Tomography (PET) (Foster et al., 2007; Nordberg, Rinne, Kadir, & Langstrom, 2010), Cerebrospinal Fluid (CSF) (Dubois et al., 2007). Some clinical trial and examination such as Alzheimer's Disease Assessment Scale-Cognitive (ADAS-Cog) Behaviour Section (M. W. Weiner et al., 2012), Geriatric Depression Scale (GDS) (Yesavage et al., 1982), and Functional Activities Questionnaire (FAQ) (Pfeffer, Kurosaki, Harrah, Chance, & Filos, 1982), Clinical Dementia Rating (CDR) Sum of Boxes (CDR-SOB) (Daly et al., 2000), and Mini-Mental State Examination (MMSE) (Michael W Weiner et al., 2012), were used diagnosis of Alzheimer's Disease too.

Automatic detection of AD procedures based on various biomarkers and techniques are using machine learning techniques on brain images for diagnosis of AD ((Amat-ur-Rasool, Ahmed, Hasnain, & Carter, 2021; Bivona et al., 2021; De Marchi et al., 2021; Jiang et al., 2021; McGrowder et al., 2021)). Most of these CAD (Computer Aided Diagnosis) systems have three processing steps: preprocessing, feature extraction, and classification. As another example of CAD systems (Dashtban & Li, 2019) used wide spectrum of health causes, economics and environment to improve understanding of patient behaviors by deep-learning algorithm.

There are several different techniques to achieve precise classification, such as, Principal Component Analysis (PCA), Artificial Neural Networks, Partial Least Square (PLS) and support vector machine (SVM). The purpose of this study is to compare different features for classification of patients with AD, MCI and NC subjects based on segmented MRI, personal information, and MMSE scores and select the best feature(s) and best classification algorithm to achieve better diagnosis AD patients from healthy people using Multi Level Fuzzy Neural Networks. Richer information can help

us to improve diagnosis accuracy. Finding the best features and more than that, least number of feature that can achieve accurate and precise classification is an old problem, but yet an area of considerable importance; therefore, this study has focused on issues regarding Alzheimer's Disease diagnosis.

1.1 Characteristic of subjects

It needs to be mentioned that only the first exam of each patient in ADNI database has been used in this work. 705 images and other features of patients were used for assessing the proposed approach.

Demographic data of patients and subjects is summarized in Table 1. All the data have been downloaded from the ADNI database in September 2016.

Table 1 Summary of Demographic data of patients and subjects.

	Count	Male	Female	Married	Widowed	Divorced	Never Married	Average of AGE	Average of MMSE
AD	156	76	80	127	18	8	3	74.89	23.32
NC	211	110	101	142	38	17	14	75.91	29.13
MCI	338	215	123	269	39	24	6	74.51	27.05
Total	705	401	304	538	95	49	23	75.01	26.85

Data and proposed method (preprocessing technique, feature selection) of this work are presented in Section 2. The experimental results are provided in Section 3. The discussion and conclusions are presented in the final section.

2. Proposed Method

A novel method for diagnosis of AD based on Multi-Level Fuzzy Neural Networks (MLFNN) and SVM classifiers is proposed. This method involves mainly the feature extraction and selection, data normalization, and classification of AD using MLFNN and SVM classifiers on MMSE, MRI, and personal information (i.e. age, marital status, gender, and education level of subjects). Before doing so, the proper features of interested area of those images had to be extracted. Then, classification was done using combined SVM and MLFNN classifiers as a framework.

Finding least features and a compatible algorithm that can make a precise classification on it, is the problem that was aimed and solved in this study. In this way, Alzheimer's disease diagnosis problem is solved using only a single feature i.e. MMSE scores, that is a clinical test which can free patients and healthy people from a lot of painful and time consuming experiments. To do so, lots of algorithms on several datasets were tested to find the best match to solve the problem, some of which are reported here. To achieve better results the input data were normalized using natural logarithm. Then, proposed framework, combining results of MLFNN classifier and SVM classifier, using only one feature (MMSE scores) gave us excellent performance. The proposed method was evaluated using k-fold cross validation method ($k = 10$). In the 10-fold cross validation method, all the data partition to ten equal parts, nine parts for training and one part for testing.

2.1. MRI Acquisition Parameters

In the multiple ADNI sites, multiple machines as Siemens, Philips, and GE Medical scanners are used. Standard protocol developed to evaluate 3D T1-weighted sequences for morphometric analyses (Jack et al., 2008). Structural brain MRI scans were acquired using 1.5 and 3 T MRI scanners. Most of the 1.5T MRIs were obtained from GE Medical scanners, and most of the 3T MRIs were acquired from Siemens machines.

In 1.5T protocol each subject experienced 2*1.5T T1-weighted MRI by 3D sagittal volumetric magnetization prepared rapid gradient echo (MP-RAGE) sequence. The repetition time of typical 1.5T acquisition are 2400ms, and the inversion time is 1000ms. Flipping and field of view are 8°, 24cm, respectively. Dimensionality of MRIs is $256 \times 256 \times 170$, and the voxel size is $1.25 \times 1.25 \times 1.2$ mm3.

For 3T scans, repetition time and inversion time are 2300ms and 900ms respectively. Flipping angle and field of view are 8°, 26cm. Dimensionality of MRIs is $256 \times 256 \times 170$, with voxel size of $1.0 \times 1.0 \times 1.2 \text{ mm}^3$.

For modern systems, scan time at 1.5T is 7.7min, and for 3T systems is 9.3min. This usually happens because of the difference between vulnerability artifacts, spin relaxation and chemical shift properties in 1.5T and 3T.

2.2. Preprocessing of MR Images

Spatial Parametric Mapping (SPM) software used for preprocessing (Ashburner, 2011). SPM used for realignment, smoothing and spatial normalization and feature extraction from ROI of MRIs. The preprocessing steps using VBM8¹ tools are as below:

- 1) Checking format of the images for being in a suitable condition using SPM tools.
- 2) Segmentation of the images, to identify gray matter (GM) and white matter (WM) and warp GM to the segmented image to Montreal Neurological Institute MNI space using the SPM tools.
- 3) Deformations estimation to best align the images to each other and create templates by registering the imported images with their average, iteratively using DARTEL tools of SPM.
- 4) Generation spatially normalised and smoothed GM images, normalised to MNI Space, using the estimated deformations by the DARTEL tools of SPM, generated the smoothed \modulated warped GM and WM images.

Data cleaning and selection was done in the preprocessing step. In the second step (feature extraction) the input data is converted into small vectors (Duin, 2000). The classification algorithm determines that the vectors are more similar to Mild Cognitive Impairment (MCI) or to AD patient, or to Normal Control (NC) subject.

2.3. Feature Extraction and selection

The features of MRI images were obtained based on regions of interest (ROI). The feature vector consisting of average voxel values from MRI images and also volume of MRI images of ROI using VBM-SPM were extracted. In this step volumetric changes of specific regions such as hippocampus, entorhinal cortex, temporal and parietal lobes, and ventricles, were used (Miranian & Abdollahzade, 2013). To extract volume and voxel values from ROIs, some masks using WFU PickAtlas² tools were made and used. Then, after applying the mask to gray matter images, the amounts of voxels were measured and counted. Other ROIs also used because, their effectiveness was reported in the literature (M. W. Weiner et al., 2012). For this purpose, segmented MRIs from ADNI database were used. Some masks were made for ROIs using SPM8 toolbox to get the desired features, and MMSE scores were obtained from ADNI database. MMSE, personal information and MRI data were used to make feature vector and classification AD and MCI patients, and NC subjects from each other. Also, using another tools as CAT12, that has great results in MRI segmentation and feature extraction (Farokhian et al., 2017), besides the VBM8 tools, can be considered as further work.

132 MRI features, including 117 voxel values and 15 volumes of ROIs, 1 MMSE score, and 4 personal information were used to diagnose from ADs from NCs using SVM and MLFNN.

Table 2 shows 15 Brain ROIs and those average volume values before and after applying Natural Logarithm. The average voxel value of 117 regions include these regions are used in this study.

Table 2 Volumes of Brain ROIs Before and After Applying Natural Logarithm (NL) for AD Diagnosis.

Row	ROIs Names	Average Volumes Value Before Applying NL	Average Volume Value After Applying NL
1	Brain Whole Volume	981553.82	0.9550

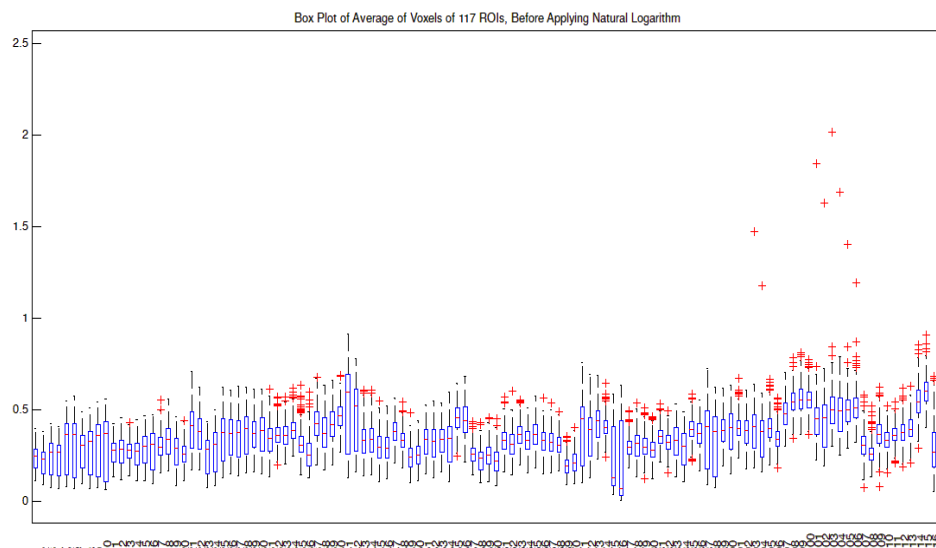
¹ <http://www.neuro.uni-jena.de/vbm/download/>

² https://www.nitrc.org/projects/wfu_pickatlas/

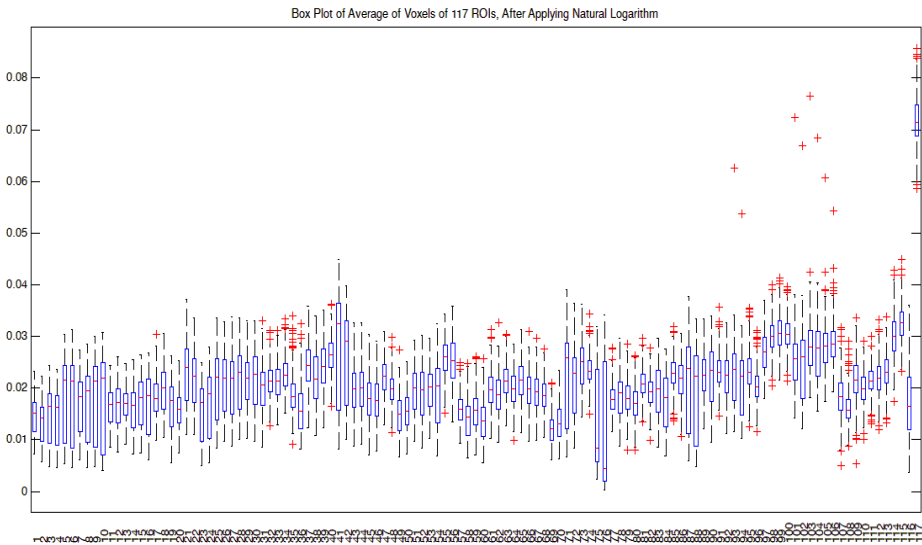
2	Estimated intracranial volume	1458425	0.9825
3	Ventricles volume	43665.516	0.7314
4	Left Hippocampus volume	3195.9745	0.5578
5	Right Hippocampus volume	3361.244	0.5612
6	Left Inferior Lateral Ventricle volume	1561.2	0.5009
7	Right Inferior Lateral Ventricle volume	1562.6468	0.5003
8	Left Middle Temporal volume	2.4296213	0.0852
9	Right Middle Temporal volume	2.4767163	0.0861
10	Left Inferior temporal volume	2.4938936	0.0865
11	Right Inferior temporal volume	2.4801191	0.0862
12	Left Fusiform volume	2.2783787	0.0821
13	Right Fusiform volume	2.2658723	0.0818
14	Left Entorhinal volume	2.9014936	0.0937
15	Right Entorhinal volume	2.9985362	0.0953

In Figure 1, BoxPlots of average of voxels values of the 117 ROIs for AD diagnosis before and after applying natural logarithm are presented. In Figure 2, Volume of 15 ROIs for these areas, before and after applying natural logarithm are presented. It is obvious in Figure 2, BoxPlots of (b) are more clear than (a) and the inter-variances of the dataset after applying natural logarithm are more dominant and tangible. These are not obvious in Figure 2 like as Figure 12. As can observe from Figures 1.b) and 2.b) values in Figure 1.b) are right between 0 and 1 but in Figure 2.b) the values are between 0 and 0.08, the cause is that we normalized all the data all together and as one block. Differences of original values are obvious in Figures 1.a) and 2.a), but after applying Natural Logarithm the values after normalization will be more dominant and tangible.

Mini-Mental State Examination (or Folstein test) is a questionnaire with 30 points that take 5 to 10 minutes. One of its advantages is requiring no specialized equipment or training for test. It has both validity and reliability for the diagnosis of Alzheimer's disease. Personal information is downloaded directly from ADNI.



a)

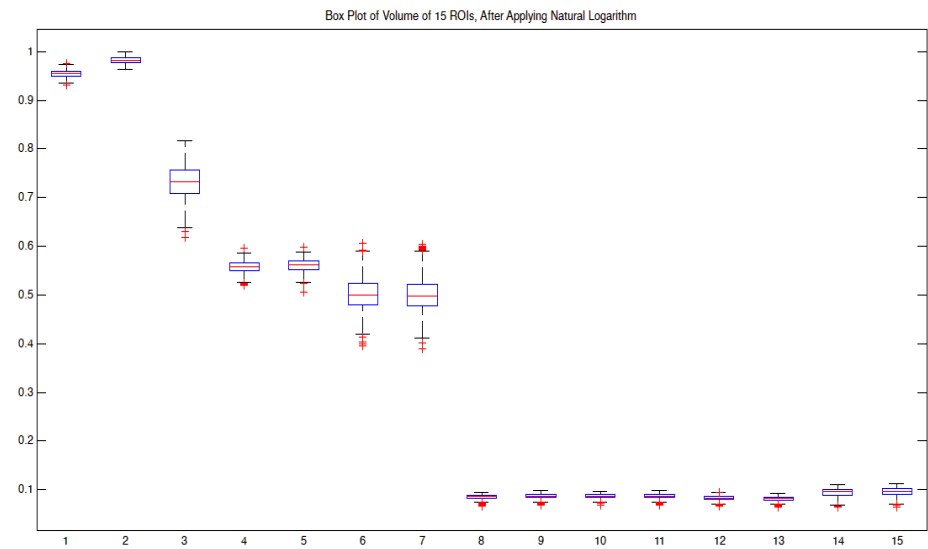


b)

Figure 1 Box Plot of Average of Voxels of 117 ROIs for all three group of patients and subjects (i.e. AD, MCI and NC). a) Before Applying Natural Logarithm, b) After Applying Natural Logarithm.



a)



b)

Figure 2 Box Plot of Volume of 15 ROIs for all three group of patients and subjects (i.e. AD, MCI and NC). a) Before Applying Natural Logarithm, b) After Applying Natural Logarithm.

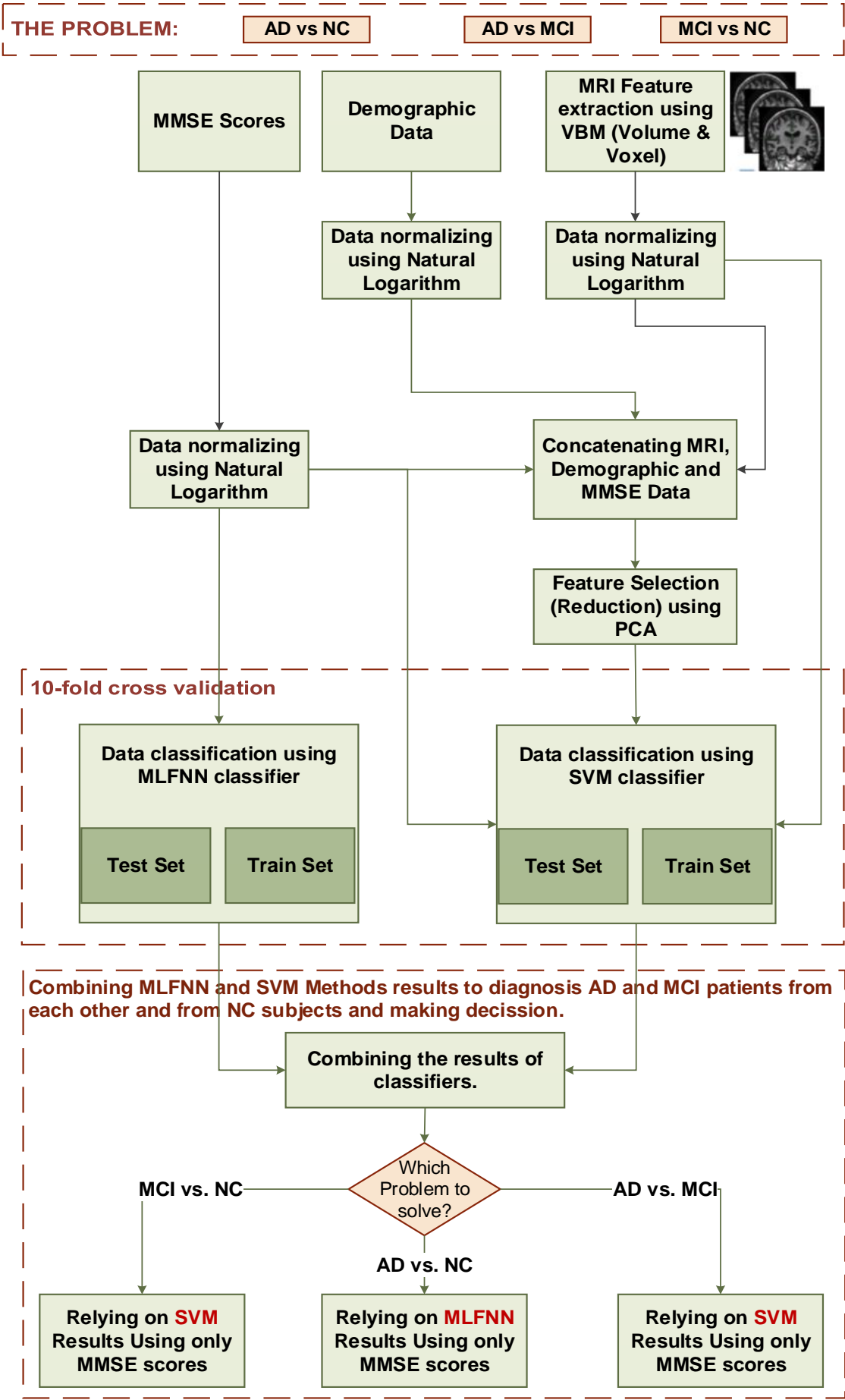


Figure 3 Block diagram of proposed framework.

Not all of the 137 extracted features can be used for Alzheimer's Disease clinical diagnosis, not only because they may be not beneficial in diagnosis, but also they can affect negatively the diagnosis performance. In this study after features extraction of the ROIs that mentioned in the literature, because of affecting the above regions of the brain in the preliminary stages of the AD and not letting go those important information, we used one of the public methods for dimension reduction that is, principal component analysis (PCA). The PCA maps feature set or data to a lesser-dimension with keeping most of the variance of them. Eigenvectors of largest eigenvalues will be principal components. The some initiatory eigenvectors often have most of the information of the original feature set. Reducing computation complexity and elimination of the negative effect of redundant features are advantages of using PCA. But when mean and covariance are not enough to define Indetermination of the principal components to keep information of data, PCA will fail.

Figure 3 displays diagram of the proposed framework. All the combinations of the data were examined in this study using both of classifiers (SVM and MLFNN). The diagram of Figure 3 represents main goal of the proposed method, that is diagnosis of AD using one feature. As can be seen in the Figure 3, based on one of the three problems we decide to solve in the top of framework, it will select the algorithm to classification in the bottom of it.

In this study we used 25 PCA component that include 99% variance of the original data, it is a lot lesser than the 137 features and more effective.

2.4. Multi Level Fuzzy Min-Max Neural Networks

A Fuzzy min-max (FMM) neural network is a type of artificial neural network that has high efficiency among other machine learning methods (Jawarkar, Holambe, & Basu, 2011). The FMM algorithm is proposed by Simpson in 1992. These kinds of networks have been used in many cases for classification and clustering purposes (Lin, Chang, & Lin, 2013), (Miranian & Abdollahzade, 2013). Training phase of FMM algorithm is online-adaptive and single-pass-through, which will be done in appropriate time. These networks are based on the hyperbox concept. Every hyperbox determines a part of one class. Hyperboxes are convex boxes. Those boxes represent a small part of pattern space. These hyperboxes expand during training phase to cover pattern space completely. There are no overlaps between hyperboxes that belong to different classes, at the end of training stage. Expansion parameter controls the maximum extension of these hyperboxes. Initialization value of this parameter is between 0 and 1.

The boxes are determined by their two opposite corners vertices. V and W points represent the beginning (min) and end (max) points of the hyperbox, respectively. Figure 4 shows a three-dimensional hyperbox. Each hyperbox belongs to one of the classes and covers a part of pattern space, defined as

$$B_i = \{X, V_j, W_j, f(X, V_j, W_j)\} \forall X \in I^n \quad (1)$$

Where V_i and W_i are min and max corners of the hyperbox, respectively. X is input vector, and n is number of input space dimensions. Each class can belong to one hyperbox at least. Hyperboxes from different classes must not overlap but same class's hyperboxes could have overlap. Figure 5 shows the final hyperboxes of an FMM network in a 2-D binary classification example.

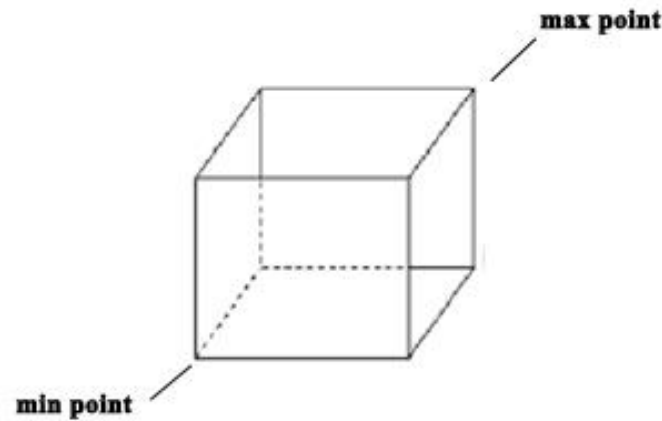


Figure 4 3-D hyperbox and its min and max points.

Various membership functions may be used for hyperboxes. For instance, Simpson's function in an earlier method of FMM is as follows:

$$hb_j(X_h) = \frac{1}{2} \sum_{i=1}^n [\max(0, 1 - \max(0, \gamma \min(1, a_{hi} - W_{ji}))) + \max(0, 1 - \max(0, \gamma \min(1, V_{ji} - a_{hi})))] \quad (2)$$

where $A_h = (a_{h1}, a_{h2}, \dots, a_{hn}) \in I^n$ is the h th sample and γ is a coefficient between 0 and 1 that adjusted how fast the membership values decrease along with increasing distance between A_h and B_j . Figure 6 shows this function and the coverage area.

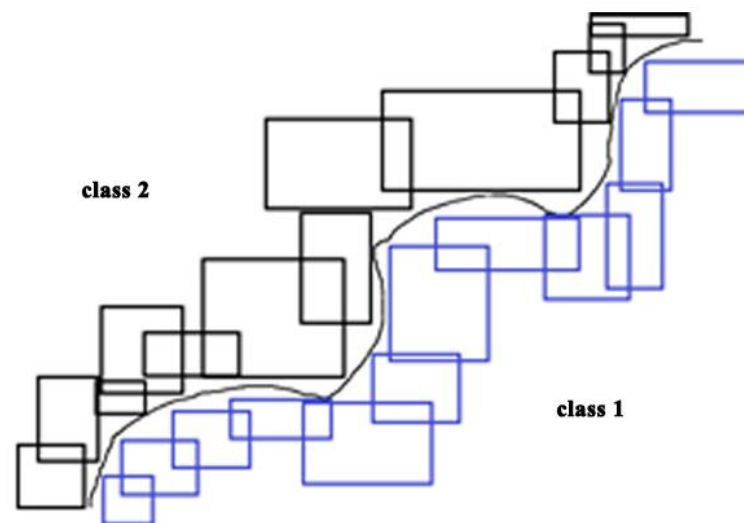


Figure 5 Classification in Fuzzy Min-Max neural network.

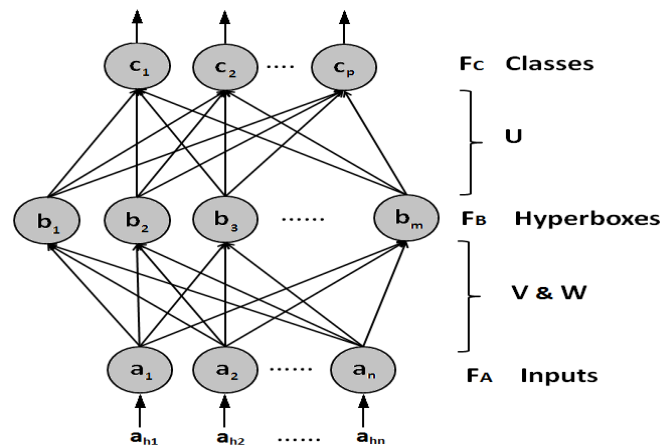


Figure 6 Structure of the classic Fuzzy Min-Max neural network.

As shown in Figure 6, FMM neural network is composed of three layers. The first layer is input layer, the second layer represents hyperboxes, and each node in the third layer represents one class. Also, each hyperbox is a node in middle layer, and the membership function of this hyperbox is the transfer function of that node. Figure 7 depicts details of a hyperbox. Each node of F_A layer is connected to all nodes in the F_B layer and each of these links has two weights (V_{ij} and W_{ij}), which are, respectively, the min and max points of the B_j hyperbox, and i is the index of the nodes in the first (Input) layer. Each node of the F_B (middle) layer is also connected to all nodes in the F_C (Output) layer. Weights of those links are

$$u_{ij} = \begin{cases} 1, & \text{if } b_j \in C_i \\ 0, & \text{if } b_j \notin C_i \end{cases} \quad (3)$$

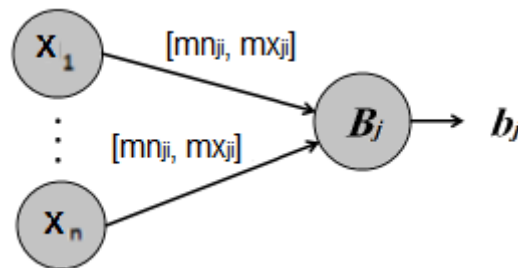


Figure 7 Details of one node in FMM neural network.

In learning phase all hyperboxes will be formed and adjusted. Every new sample will be checked to find out if there is a box that belong to the same class of the sample. If such a hyperbox appears, no further processing is needed and it is next sample's turn. If algorithm cannot find one, following three steps will be executed.

- 1) *Expansion*: When a sample is sent to algorithm, a hyperbox of same class must be found and it also needs to be capable of expansion to cover the sample. Hyperbox size does not exceed the θ parameter. If there is none, a new hyperbox is created with min and max points equal to coordinate of the sample.
- 2) *Overlap Test*: In this step, the overlapping area of the new hyperbox is checked with all hyperboxes of other classes. When each dimension is recognized in one case of (5), two hyperboxes overlap. To elimination of overlaps, the dimension (Δ) with least overlap is selected for contraction

$$\begin{aligned} \text{Case 1: } & v_{ji} < v_{ki} < w_{ji} < w_{ki} \\ \text{Case 2: } & v_{ki} < v_{ji} < w_{ki} < w_{ji} \end{aligned} \quad (4)$$

Case 3: $v_{ji} < v_{ki} < w_{ki} < w_{ji}$

Case 4: $v_{ki} < v_{ji} < w_{ji} < w_{ki}$

3) Contraction: When there was no overlap, there is no need of this step; otherwise, due to the type of the overlap in the Δ direction, one case of following will be executed.

Case 1: $v_{ji} < v_{ki} < w_{ji} < w_{ki}$

$$v_{k\Delta}^{new} = w_{k\Delta}^{new} = \frac{v_{k\Delta}^{old} + v_{j\Delta}^{old}}{2} \quad \text{OR} \quad w_{j\Delta}^{new} = v_{k\Delta}^{old} \quad (5)$$

Case 2: $v_{ki} < v_{ji} < w_{ki} < w_{ji}$

$$v_{j\Delta}^{new} = w_{k\Delta}^{new} = \frac{v_{j\Delta}^{old} + v_{k\Delta}^{old}}{2} \quad \text{OR} \quad v_{j\Delta}^{new} = w_{k\Delta}^{old} \quad (6)$$

Case 3: $v_{ji} < v_{ki} < w_{ki} < w_{ji}$

If $w_{k\Delta} - v_{j\Delta} < w_{k\Delta} - v_{j\Delta}$ Then

$$v_{j\Delta}^{new} = w_{k\Delta}^{old}$$

Else

$$w_{j\Delta}^{new} = v_{k\Delta}^{old} \quad (7)$$

Case 4: $v_{ki} < v_{ji} < w_{ji} < w_{ki}$

If $w_{k\Delta} - v_{j\Delta} < w_{j\Delta} - v_{k\Delta}$ Then

$$w_{k\Delta}^{new} = v_{j\Delta}^{old}$$

Else

$$v_{k\Delta}^{new} = w_{j\Delta}^{old} \quad (8)$$

Here, Δ is the selected dimension. These three steps are executed on every learning sample to create and adjust the required hyperboxes.

MLFNN algorithm uses a multilevel tree structure. The network function is not like a decision tree but like a homogenous classifier. This method, despite of the classic FMM methods, does not manage the overlap problem in the contraction step. Hyperboxes of each level are smaller and more detailed than hyperboxes of previous levels. The proposed method, creation of hyperboxes in the levels, and the classification are illustrated in Figure 8.

The first level classifier, classifies the non-boundary region of the data, and nodes (classifiers) of second level take care of the boundary region. After all, the node that has the best output among all nodes is selected as the network's output.

In MLFNN, like in other FMM methods, all hyperboxes are created and adjusted during training phase and are used in the test phase. Other FMM methods handle overlap problem as soon as it is found, but MLF handles the overlaps when creation and adjustment of all hyperboxes are done, that results in reduction of memory and time complexity.

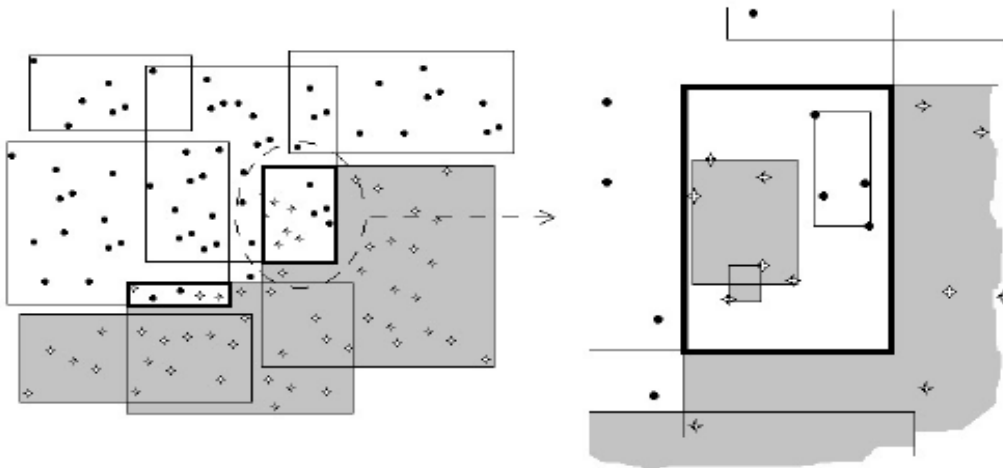


Figure 8 Different levels of classification in MLF.

3. Experiments and Results

Experimental results using MRI, MMSE, and Personal Information data have been conducted for classifying NC, AD and MCI patients. Evaluations were performed under different features subsets including the features of all the data, only the MRI data, only the personal information, and only the MMSE scores. A total number of 137 features were selected to form the final feature subset involving 132 MRI features (including 117 voxel values and 15 volumes of ROIs), 1 MMSE score, and 4 personal Information. After normalization and feature reduction, the Multi-Level Fuzzy Neural Network and SVM classifiers were trained using these feature vectors, to make a framework. The average accuracy of 10-fold cross validation was used to represent the performance of the learning models, and the obtained results were compared with other state-of-the-art methods. Some popular performance measures such as Accuracy (ACC), sensitivity (SEN), positive predictive value (PPV), specificity (SPE), negative predictive value (NPV), and area under curve (AUC) are employed to demonstrate the real performance of the constructed model. To this end, MATLAB 2014 software, on a PC having 2 GHz Core i7 CPU, and 12GB of RAM was employed. The elapsed time for recognizing binary groups are as below table:

Table 3 Time complexity of MLFNN classifier vs. SVM classifier.

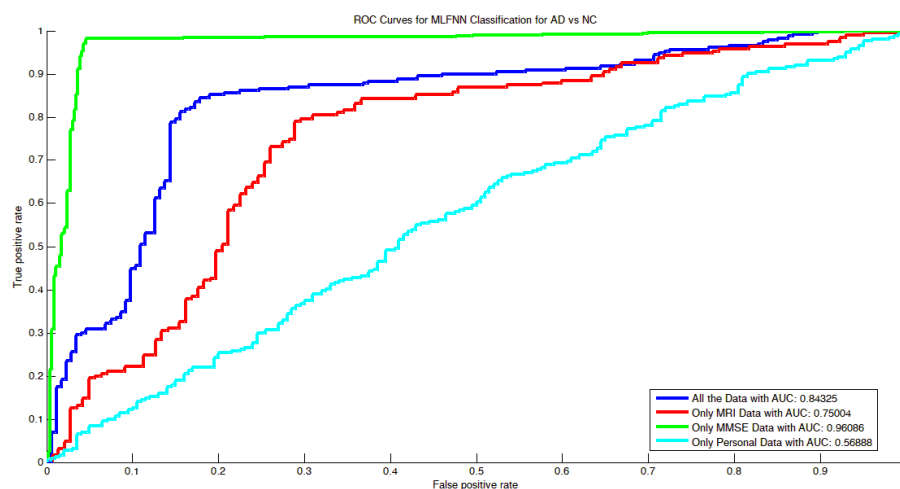
Data	Classes	MLFNN train and test time (seconds)	SVM train and test time (seconds)
All Data Using MLFNN	NC-AD	16.44226	5.500653
	AD-MCI	22.49492	8.895344
	MCI-NC	25.74769	11.04229
Only MRI Data Using MLFNN	NC-AD	15.99595	5.6213
	AD-MCI	22.07155	10.60535
	MCI-NC	24.83326	9.899532
Only MMSE Data Using MLFNN	NC-AD	2.176935	4.143991
	AD-MCI	7.561059	7.604824
	MCI-NC	12.23362	8.662185
Only Personal Information Data using MLFNN	NC-AD	7.546106	5.859331
	AD-MCI	13.21055	8.251767
	MCI-NC	15.00944	9.350641

In this study SVM with RBF Kernel was utilized to compare to and combine with MLFNN results. The SVM box constraint is set to 50 and RBF sigma was set to 100 (scaling factor with default 1). Table 3 shows that the time complexity of MLFNN for one feature is better than SVM with 137 features. Moreover, resources usage such as CPU while running MLFNN is so lower than SVM, but from the table can observe that when number of the sample increased MLFNN running time go to increase progressively. Table 4 shows the classifier and features parameter used in this study.

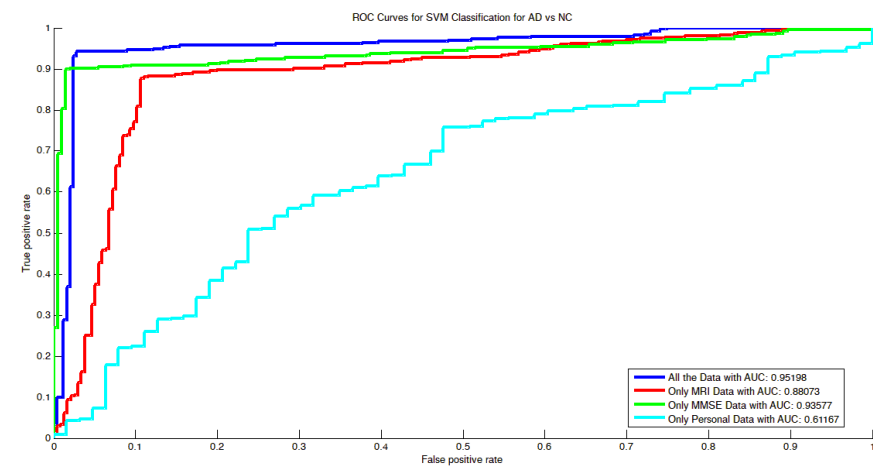
Table 4 The parameters of classifier and features set.

Number of MRI features	132
Number of Personal Info	4
Number of MMSE Scores	1
Total Number of features	137
Number of featured used after PCA	25
Variance of selected PCA elements	99%
SVM Kernel	RBF
RBF Sigma	100
SVM Box Constraint	50

According to Table 5 and Figure 9 - Figure 12, the AUC value of MLFNN using only MMSE scores is the highest compared to other feature subsets used here. On the other hand, using this setup the model could achieve high accuracy of 96.6%. Figure 12 shows that box-plots of method using MLFNN with MMSE scores were more compressed which depicts the more stability in the constructed classification model. Plus, the accuracy of NC versus MCI was higher when using SVM with MMSE scores. The MLFNN algorithm showed lower performance for classifying NC vs. MCI, yet it has the highest NPV in this area. As shown in Figure 12, the demographic information in all the three binary classifications has slightly more than 50% accuracy. The SVM using all the features achieved best results in NC vs. MCI classification in the most of the performance measures. MLFNN using MMSE scores for AD vs. NC classification achieved best performance in the most of measures, which is a great result, and using SVM over all the data for AD vs. MCI and NC vs. MCI showed accuracy was 79.08% and 70.43% which are appropriate results compared to other methods in this realm and other state of the art methods. To achieve final appropriate performance, we combined results of MLFNN and SVM classifiers using only MMSE scores, as can be seen in Figure 3.

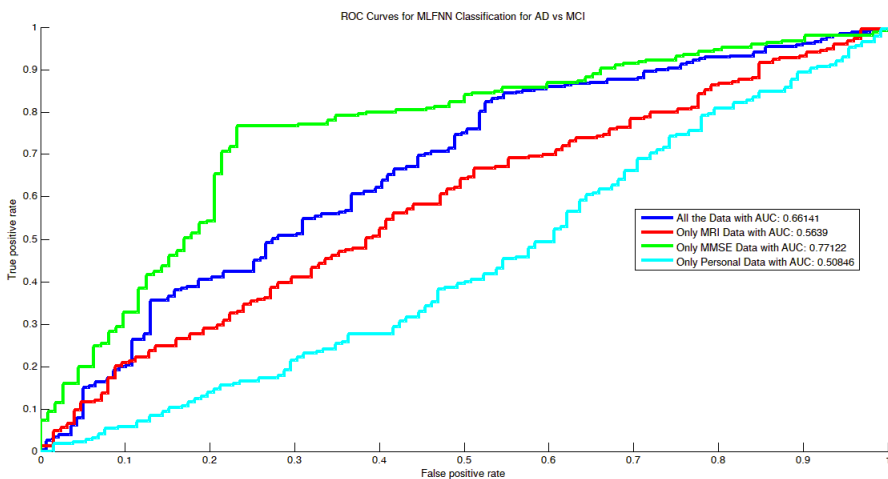


a)

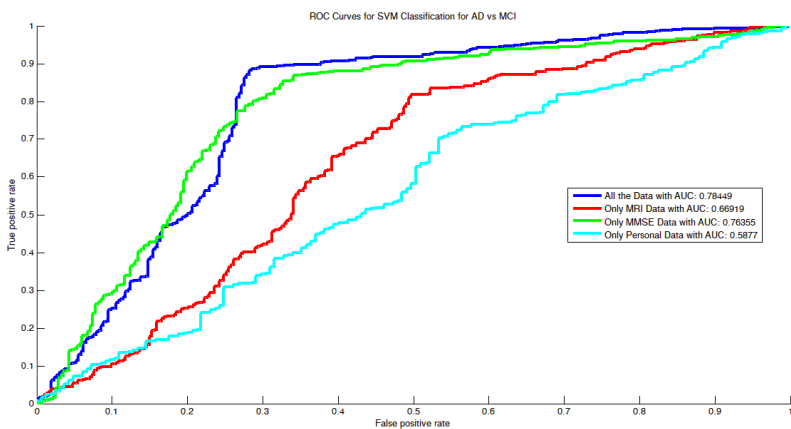


b)

Figure 9 ROC curves for classification of AD versus NC. a) Using MLFNN, b) Using SVM.

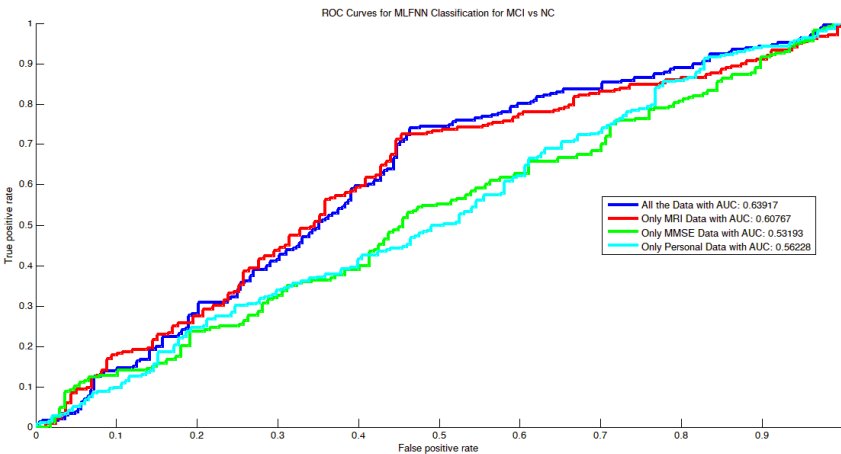


a)

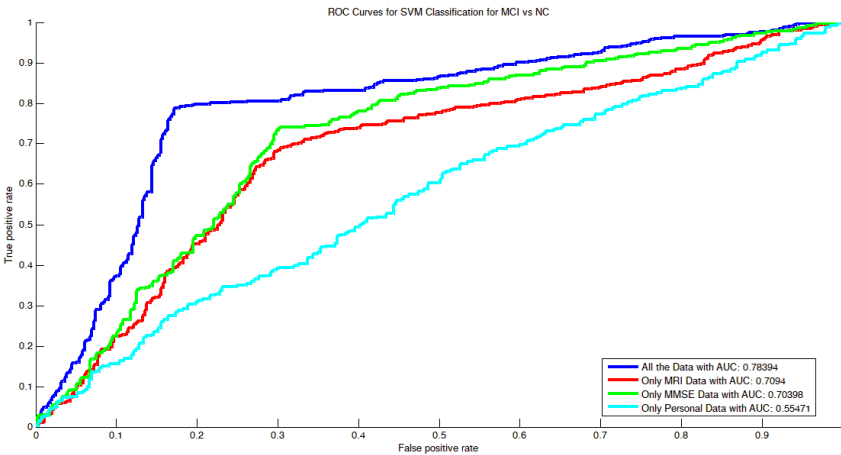


b)

Figure 10 ROC curves for classification of MCI versus NC. a) Using MLFNN, b) Using SVM.



a)



b)

Figure 11 ROC curves for AD versus MCI classification. a) Using MLFNN, b) Using SVM.
Table 5 Performance metrics of SVM and MLFNN over extracted data and its subsets. Classification accuracy (ACC), sensitivity (SEN), specificity (SPE), positive predictive value (PPV), negative predictive value (NPV), and area under curve (AUC).

Data	Classes	ACC %	SEN %	SPE %	PPV %	NPV %	AUC
All Data Using MLFNN	AD-NC	84.94	82.11	87.04	82.56	86.74	0.8432
	AD-MCI	72.97	46.99	85.13	59.82	77.46	0.6614
	MCI-NC	63.82	82.71	33.56	66.66	54.81	0.6392
All Data Except Marital Data Using MLFNN	AD-NC	84.87	83.61	85.82	81.54	87.61	0.8506
	AD-MCI	73.89	49.28	85.36	61.14	78.34	0.6750
	MCI-NC	64.21	85.72	29.77	66.17	56.67	0.5772
Only MRI Data Using MLFNN	AD-NC	75.17	76.2	74.41	69.04	80.75	0.75
	AD-MCI	65.5	33.42	80.47	44.37	72.18	0.5639
	MCI-NC	62.1	81.76	30.7	65.36	51.55	0.6077
Only MMSE Data Using MLFNN	AD-NC	96.6	94.83	97.92	97.14	96.24	0.9609
	AD-MCI	76.08	78.03	75.18	59.55	88.16	0.7712

Only Personal Information Data using MLFNN	MCI-NC	54.12	43.2	60.97	40.8	63.22	0.5319
	AD-NC	57.93	50.71	63.31	50.54	63.49	0.5689
	AD-MCI	57.36	33.58	68.42	33.03	68.94	0.5085
	MCI-NC	56.73	64.92	43.66	64.81	43.81	0.5623
All Data Using SVM	AD-NC	94.49	97.58	92.51	89.27	98.36	0.952
	AD-MCI	81.08	68.76	87.66	74.84	84.01	0.7845
	MCI-NC	78.24	93.34	65.36	69.7	91.99	0.7839
	AD-NC	94.46	97.32	92.61	89.51	98.16	0.9505
All Data Except Marital Data Using SVM	AD-MCI	80.71	68.21	87.39	74.23	83.75	0.7758
	MCI-NC	78.58	93.59	65.74	70.03	92.30	0.7976
	AD-NC	87.89	88.81	87.3	82.1	92.24	0.8807
	AD-MCI	69.31	51.41	82.63	68.77	69.56	0.6692
Only MRI Data Using SVM	MCI-NC	70.74	86.53	58.25	62.12	84.53	0.7094
	AD-NC	91.97	99.22	87.98	81.95	99.51	0.9358
	AD-MCI	79.08	64.65	87.91	76.43	80.34	0.7636
	MCI-NC	70.43	93.15	57.1	56.01	93.43	0.704
Only MMSE Data Using SVM	AD-NC	60.89	62.11	60.7	21.49	90.19	0.6117
	AD-MCI	64.93	44.56	73.07	39.77	76.75	0.5877
	MCI-NC	55.86	68.31	44.53	52.85	60.68	0.5547

Table 6 Comparison of the proposed method with other state of the art methods for diagnosis of AD.

	AD vs NC				AD vs MCI				MCI vs NC			
	Acc%	Sen%	Spec%	AUC	Acc%	Sen%	Spec%	AUC	Acc%	Sen%	Spec%	AUC
(Zhang, Wang, Zhou, Yuan, & Shen, 2011)	93.20	93.00	93.30	0.98	-	-	-	-	76.40	81.80	66.00	0.81
(Dai, He, Vogelstein, & Hou, 2013)	90.81	92.59	90.33	0.94	85.92	82.46	87.59	0.87	81.92	78.51	88.34	0.81
(Suk, Lee, Shen, & Initiative, 2014)	93.05	90.86	94.57	0.95	88.98	82.11	90.65	0.90	83.67	96.79	57.28	0.82
(Liu et al., 2016)	94.65	95.03	91.76	0.95	88.63	91.55	86.25	0.91	84.79	88.91	80.34	0.83
(Mishra, Beheshti, Khanna, & Initiative, 2018)	89.15	85.06	92.53	0.93	-	-	-	-	-	-	-	-
(Khedher, Ramírez, Górriz, Brahim, & Segovia, 2015)	88.96	92.35	86.24	0.93	84.59	88.75	83.07	0.89	82.41	84.12	80.48	0.81
(Lian, Liu, Zhang, & Shen, 2019)	90.00	82.00	97.00	0.95	-	-	-	-	-	-	-	-
(Ben Ahmed, Benois-Pineau, Allard, Ben Amar, &	87.00	75.50	100	0.85	72.23	75.00	70.00	0.76	78.22	70.73	83.34	0.77

Catheline, 2014)												
(Zhou, He, Xu, Xiong, & Cai, 2018)	93.75	87.5	100	-	-	-	-	-	-	-	-	-
Proposed Method	96.60	94.83	97.92	0.96	79.08	64.65	87.91	0.77	70.43	93.15	57.1	0.61

4. Discussion

For the early detection of AD, we developed a novel machine learning method with MMSE scores. The information provided in Figures 9-11 and Table 4 shows that using MLFNN classifier and MMSE scores together makes better results for distinguishing AD patients from NC subjects, and the Box-Plots are more compressed, clearly, as shown in Figure 12, that means more reliability and stability in classification. As mentioned at the beginning of the paper, the first exam of every person was adopted from ADNI. The proposed method yielded higher performance in distinguishing AD, that can assist with detecting Alzheimer’s Disease in its initial stages.

In this study, it was found that with only a single indicator (MMSE), a drastically high accuracy detection can be built for distinguishing between the AD, MCI patients, and NC subjects. This can be suitable for patients and patrons who are exhausted of coming and going to MRI imaging centers and of the pain of injecting materials for other imaging methods.

As can be seen in Figure 12.d, personal information were experimented to see how that might contribute to the results. However, it shows a seemingly poor discriminant potential. It actually can distinguish between all three binary groups marginally better than a random classifier. Despite that, it can help us to build up a better feature set for classification. From Table 5, it can be realized that employing personal information for example marital status have a small positive impact on AD diagnostics; but employing them can help to achieve a slightly better performance.

From the table 5, MLFNN classifier using all the features obtains accuracy of 84.94%, and using all the features without marital data obtains 84.87%. SVM classifier obtains accuracy of 94.49% using all the features, and 94.46% using all the features without marital data. The above mentioned shows a little friction between classification with and without marital data. But, we always have to use all the information we have.

Table 6 represents comparison of the proposed framework to other state of the art methods. As shown in the table, the proposed framework outperformed other methods in most performance measures. Most of the listed methods in Table 6 have used ADNI database and we used all the images and data in ADNI (at the time of start), consists of 705 participants. Whereas some of the reported methods used a portion of the ADNI samples, we have rights to compare our proposed method to them.

In the Table 5, for instance, the AUC of the MLFNN is significantly higher than the other methods. The proposed method also yielded a higher performance in most metrics of AD diagnosis evaluation.

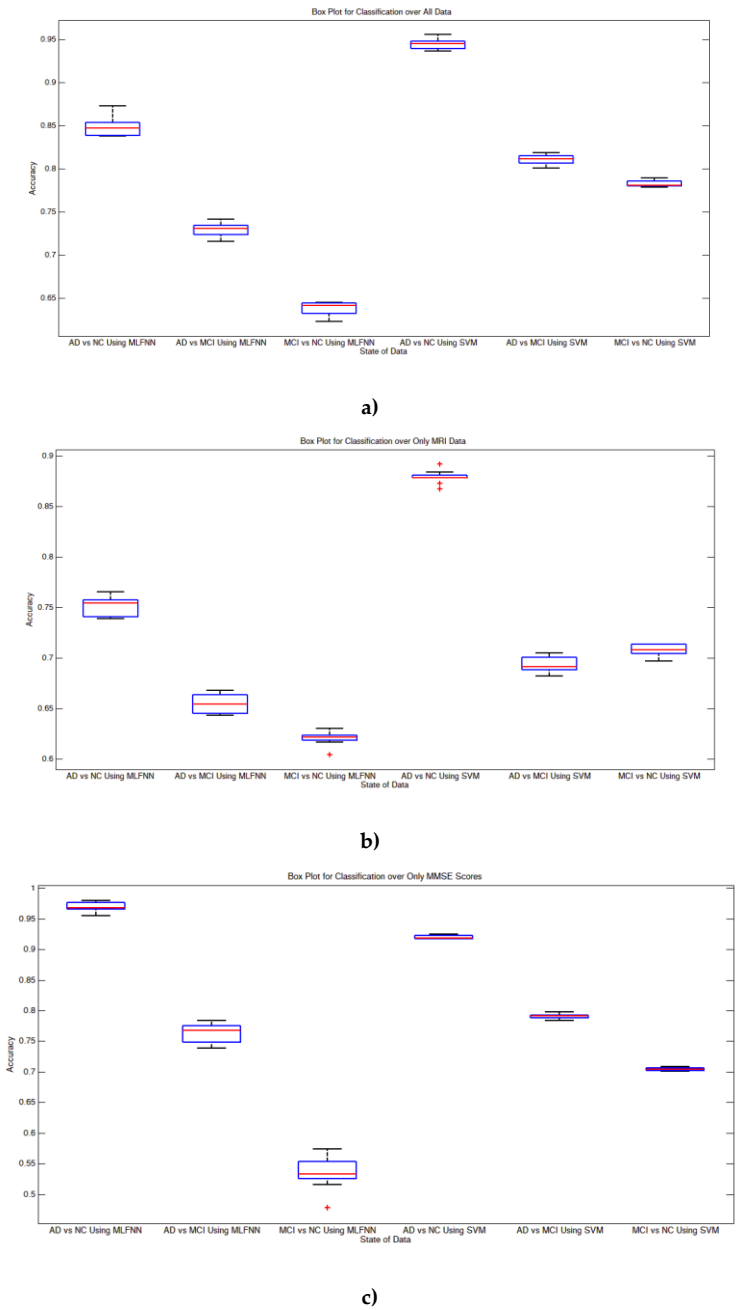
Comparing to typical supervised learning approaches, MLFNN represented a significant increase in AUC scores. This single biomarker provided 0.9609 for AUC score. Since the average value of AUC is 0.9609, the proposed model has near perfect diagnostic capability on testing (unseen) samples to differentiate the diseased and non-diseased patients. As can be seen in Table 6, the proposed framework obtained highest accuracy (96.6%) in classifying AD patients from NC subjects, and due to the balanced dataset, the accuracy is a suitable performance measure.

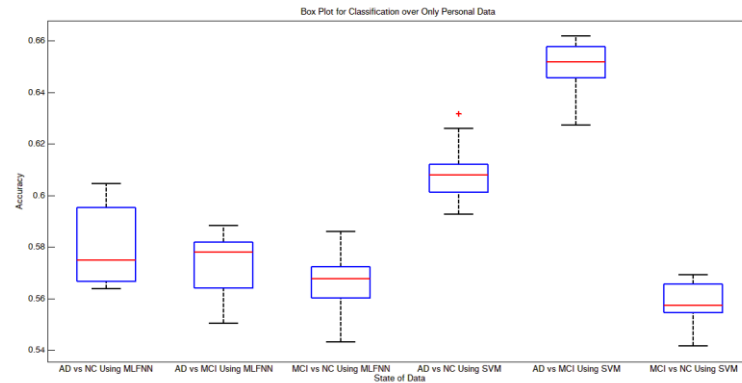
It should be noted, imbalanced dataset is a dataset that the number of its samples is not the same for all the classes. In some fields, same as fraud detection, where the majority of the cases will be in the “Not-Fraud” class and a very small minority will be in the “Fraud” class, imbalanced datasets are quite common problem. In these cases, it is recommended calculating the sensitivity of the classifier in addition to accuracy (the accuracy achieved for the balanced test dataset is not suitable for estimating the accuracy of imbalanced dataset).

In addition, sensitivity (recall rate) of proposed method (94.83) is the second highest sensitivity rate as can be seen in Table 6, after (Liu et al., 2016) the specificity rate of this method which is

significantly lower than our proposed method. The proposed method stood at the second place with specificity of 97.9% against perfect detection rate of (Ben Ahmed et al., 2014), similar to sensitivity. However,(Ben Ahmed et al., 2014) method had a high false alarm rate while the proposed method succeeded to achieve a trade-off between the two! It is worth mentioning that the lower the specificity, the higher the sensitivity (94.83 vs. 87 for example in case of (Ben Ahmed et al., 2014) method). Nevertheless, combined methods can be used to achieve better performance.

According to the Table 5, when we use all the data together as feature set, SVM outperforms MLFNN, nevertheless, adopting only one feature (MMSE score) for MLFNN makes it to outperform SVM in 6 measurements. Further, the SVM classifier using all the data (137 features) in 2 of measurements (sensitivity and NPV) is better than MLFNN.





d)

Figure 12 BoxPlots for recognition of AD, NC, and MCI subjects: (a) Using All the data. (b) Using only MRI data. (c) Using only MMSE scores. (d) Using subjects' Personal data.

Adopting MMSE scores and MLFNN algorithm achieved a lower performance in classifying MCI vs. NC or MCI vs. AD. A better performance could be reached using SVM and MRI features, though. As a result, even though the MLFNN classifier performs rather poor in detecting direct combination (concatenation) of multimodal data, when we adopted a great singular feature, its performance outperformed the other methods significantly. Even though, MLFNN had comparatively lower performance in distinguishing MCI patients, integrating the proposed method simply with SVM can yield great performance. Actually, when the problem is distinguishing AD patients from NC subjects we can rely on the MLFNN and when the problem is finding MCI patients we use SVM. It is worth noting, this combination is a wise approach. For this, Figure 3 demonstrate this framework.

It should be noted, as can be seen in the Figure 12.d and Table 5 the recognition performance of the personal information is a bit higher than random, but adding this little information to the model can impacts the performance of the whole model, and a bit of information can help to improve performance of whole system. In this regards all the information used in this study can be useful, that is age, gender, marital status, and education can have effect in dementia and AD diagnosis (Sundstrom, Westerlund, & Kotyrló, 2016).

Furthermore, one important justification of this method's outstanding achievement comes from the fact that the distribution of testing samples for the diseased and non-diseased subjects has no overlap (Folstein, Folstein, & McHugh, 1975). From the Table 6 the MLFNN algorithm for AD vs NC classification in all of the performance measure is in first or second place, and in other two classification problems have comparable results. From that can realize that the proposed framework is almost dominant one compare to the other methods.

5. Conclusion

An effective framework for early detection of Alzheimer Disease was proposed. Three separated group of data was used in this study, i.e. extracted features from MR Images, Personal information and MMSE scores of the subjects. For evaluation purpose the ADNI database was utilized. In the realm of Alzheimer's disease diagnosis, we achieved an excellent performance using only a single feature i.e. MMSE indicator by developing and adopting Multi Level Fuzzy neural networks that discriminated AD patients from non-patient people near perfectly and SVM in the other two classification problems, that is AD vs MCI and MCI vs NC. In other words, the current study fit the optimum algorithm to the best area using optimum possible feature(s), namely one feature. The discovery of this study can free patients and healthy people from painful and time-consuming experiments, and classification algorithm from time and memory complexity. However demographic information had not great performance in diagnosis AD patients from healthy subjects, and we use them to tune performance of the model. We used Natural Logarithm for data normalization purpose.

The performance was tested using 10-fold cross validation. To validate the effectiveness of our proposed method, several tests were done on ADNI database and they were compared to other AD diagnosis methods. Ultimately, experiments over 705 subjects of ADNI database showed the effectiveness of proposed method in current research for diagnosis of AD with one of the highest performance (accuracy rates of 96.6%), reported in the literature. The final model outperformed other methods in distinguishing AD from NC, however the performance of SVM was better for classifying other types of patients. This suggest that the MLFNN is highly applicable for classifying AD vs NC. Furthermore, an integrative approach can be adopted to achieve the best of results i.e. the performance of MLFNN and SVM together as a framework. However, efficiency of proposed method was proved by comparing to the state of the art works. For further research, longitudinal data could be studied for predicting the MCI to AD conversion, considering that only the baseline data were utilized in this study. Combining ADAS-Cog, FAQ and features of this study for AD diagnostics can be a future work. This study employed VBM8 tools which we got promising result, however there are other tools to explore (e.g. CAT12) can be used to improve diagnosis results.

Conflict of Interests: There are no conflicts of interest associated with this publication and there *has not* been financial support for this research work that could have influenced its outcome.

Acknowledgements: The data used in this paper were obtained from the Alzheimer's Diseases Neuroimaging Initiative (ADNI) database. The ADNI was launched in 2003 by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA). The primary goal of ADNI has been to test whether magnetic resonance imaging, other biomarkers, and clinical and neuropsychological estimations can be combined to measure the progression of early stage of Alzheimer's disease. The purpose of determination of sensitive and specific biomarkers of early AD progression is to develop new treatments and monitor their effectiveness, and reducing the time and cost of clinical experiments.³

References

- Amat-ur-Rasool, H., Ahmed, M., Hasnain, S., & Carter, W. G. (2021). Anti-Cholinesterase Combination Drug Therapy as a Potential Treatment for Alzheimer's Disease. *Brain Sciences*, 11(2), 184.
- Ben Ahmed, O., Benois-Pineau, J., Allard, M., Ben Amar, C., & Catheline, G. (2014). Classification of Alzheimer's disease subjects from MRI using hippocampal visual features. *Multimedia Tools and Applications*, 74(4), 1249-1266. doi: 10.1007/s11042-014-2123-y
- Bivona, G., Lo Sasso, B., Gambino, C. M., Giglio, R. V., Scazzone, C., Agnello, L., & Ciaccio, M. (2021). The Role of Vitamin D as a Biomarker in Alzheimer's Disease. *Brain Sciences*, 11(3), 334.
- Cuingnet, R., Gerardin, E., Tessieras, J., Auzias, G., Lehericy, S., Habert, M. O., . . . Colliot, O. (2011). Automatic classification of patients with Alzheimer's disease from structural MRI: a comparison of ten methods using the ADNI database. *Neuroimage*, 56, 766–781.
- Dai, D., He, H., Vogelstein, J. T., & Hou, Z. (2013). Accurate prediction of AD patients using cortical thickness networks. *Machine vision and applications*, 24(7), 1445-1457.
- Daly, E., Zaitchik, D., Copeland, M., Schmahmann, J., Gunther, J., & Albert, M. (2000). Predicting conversion to Alzheimer disease using standardized clinical information. *Arch Neurol*, 57(5), 675-680. doi: 10.1001/archneur.57.5.675
- Dashtban, M., & Li, W. (2019). Deep learning for predicting non-attendance in hospital outpatient appointments.
- Davatzikos, C., Fan, Y., Wu, X., Shen, D., & Resnick, S. M. (2008). Detection of Prodromal Alzheimer's Disease via Pattern Classification of MRI. *Neurobiology of Aging*, 29(4), 514-523. doi: 10.1016/j.neurobiolaging.2006.11.010

³This section has been added because of ADNI request.

- De Marchi, F., Contaldi, E., Magistrelli, L., Cantello, R., Comi, C., & Mazzini, L. (2021). Telehealth in Neurodegenerative Diseases: Opportunities and Challenges for Patients and Physicians. *Brain Sciences*, 11(2), 237.
- Dubois, B., Feldman, H. H., Jacova, C., DeKosky, S. T., Barberger-Gateau, P., Cummings, J., . . . Scheltens, P. (2007). Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS–ADRDA criteria. *Lancet Neurol*, 6, 734–746.
- Duin, R. P. W. (2000). *Classifiers in almost empty spaces*. Paper presented at the Proceedings 15th International Conference on Pattern Recognition.
- Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res*, 12(3), 189-198.
- Foster, N. L., Heidebrink, J. L., Clark, C. M., Jagust, W. J., Arnold, S. E., Barbas, N. R., . . . Minoshima, S. (2007). FDG-PET improves accuracy in distinguishing frontotemporal dementia and Alzheimer's disease. *Brain*, 130, 2616–2635.
- Jack, C. R., Jr., Bernstein, M. A., Fox, N. C., Thompson, P., Alexander, G., Harvey, D., . . . Weiner, M. W. (2008). The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. *J Magn Reson Imaging*, 27(4), 685-691. doi: 10.1002/jmri.21049
- Jawarkar, N. P., Holambe, R. S., & Basu, T. K. (2011, 3-5 June 2011). *Use of fuzzy min-max neural network for speaker identification*. Paper presented at the 2011 International Conference on Recent Trends in Information Technology (ICRTIT).
- Jiang, J., Benhamou, E., Waters, S., Johnson, J. C. S., Volkmer, A., Weil, R. S., . . . Hardy, C. J. D. (2021). Processing of Degraded Speech in Brain Disorders. *Brain Sciences*, 11(3), 394.
- Khedher, L., Ramírez, J., Górriz, J. M., Brahim, A., & Segovia, F. (2015). Early diagnosis of Alzheimer's disease based on partial least squares, principal component analysis and support vector machine using segmented MRI images. *Neurocomputing*, 151, Part 1, 139-150. doi: <http://dx.doi.org/10.1016/j.neucom.2014.09.072>
- Kloppel, S., Stonnington, C. M., Chu, C., Draganski, B., Scahill, R. I., Rohrer, J. D., . . . Frackowiak, R. S. (2008). Automatic classification of MR scans in Alzheimer's disease. *Brain*, 131(Pt 3), 681-689. doi: 10.1093/brain/awm319
- Lian, C., Liu, M., Zhang, J., & Shen, D. (2019). Hierarchical Fully Convolutional Network for Joint Atrophy Localization and Alzheimer's Disease Diagnosis using Structural MRI. *IEEE Trans Pattern Anal Mach Intell*, 1-1. doi: 10.1109/TPAMI.2018.2889096
- Lin, Y. Y., Chang, J. Y., & Lin, C. T. (2013). Identification and prediction of dynamic systems using an interactively recurrent self-evolving fuzzy neural network. *IEEE Trans Neural Netw Learn Syst*, 24(2), 310-321. doi: 10.1109/tnnls.2012.2231436
- Liu, J., Li, M., Lan, W., Wu, F.-X., Pan, Y., & Wang, J. (2016). Classification of alzheimer's disease using whole brain hierarchical network. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 15(2), 624-632.
- McGrowder, D. A., Miller, F., Vaz, K., Nwokocha, C., Wilson-Clarke, C., Anderson-Cross, M., . . . Alexander-Lindo, R. (2021). Cerebrospinal Fluid Biomarkers of Alzheimer's Disease: Current Evidence and Future Perspectives. *Brain Sciences*, 11(2), 215.
- Miranian, A., & Abdollahzade, M. (2013). Developing a local least-squares support vector machines-based neuro-fuzzy model for nonlinear and chaotic time series prediction. *IEEE Trans Neural Netw Learn Syst*, 24(2), 207-218. doi: 10.1109/tnnls.2012.2227148
- Mishra, S., Beheshti, I., Khanna, P., & Initiative, f. t. A. s. D. N. (2018). A statistical region selection and randomized volumetric features selection framework for early detection of Alzheimer's disease. *International Journal of Imaging Systems and Technology*, 28(4), 302-314. doi: 10.1002/ima.22290
- Nordberg, A., Rinne, J. O., Kadir, A., & Langstrom, B. (2010). The use of PET in Alzheimer disease. *Nat. Rev. Neurol.*, 6, 78–87.
- Pfeffer, R. I., Kurosaki, T. T., Harrah, C. H., Jr., Chance, J. M., & Filos, S. (1982). Measurement of functional activities in older adults in the community. *J Gerontol*, 37(3), 323-329.

- Suk, H.-I., Lee, S.-W., Shen, D., & Initiative, A. s. D. N. (2014). Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis. *Neuroimage*, *101*, 569-582.
- Sundstrom, A., Westerlund, O., & Kotyrló, E. (2016). Marital status and risk of dementia: a nationwide population-based prospective study from Sweden. *BMJ Open*, *6*(1), e008565. doi: 10.1136/bmjopen-2015-008565
- Weiner, M. W., Veitch, D. P., Aisen, P. S., Beckett, L. A., Cairns, N. J., Green, R. C., . . . Liu, E. (2012). The Alzheimer's Disease Neuroimaging Initiative: A review of papers published since its inception. *Alzheimer's & Dementia*, *8*(1), S1-S68.
- Weiner, M. W., Veitch, D. P., Aisen, P. S., Beckett, L. A., Cairns, N. J., Green, R. C., . . . Trojanowski, J. Q. (2012). The Alzheimer's Disease Neuroimaging Initiative: a review of papers published since its inception. *Alzheimers Dement*, *8*(1 Suppl), S1-68. doi: 10.1016/j.jalz.2011.09.172
- Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M., & Leirer, V. O. (1982). Development and validation of a geriatric depression screening scale: a preliminary report. *J Psychiatr Res*, *17*(1), 37-49.
- Zhang, D., Wang, Y., Zhou, L., Yuan, H., & Shen, D. (2011). Multimodal classification of Alzheimer's disease and mild cognitive impairment. *Neuroimage*, *55*, 856-867.
- Zhou, K., He, W., Xu, Y., Xiong, G., & Cai, J. (2018). Feature Selection and Transfer Learning for Alzheimer's Disease Clinical Diagnosis. *Applied Sciences*, *8*(8), 1372.