

# Improved Alzheimer's Disease Classification Using Innovative Multimodal Feature Selection and Fusion Technique

S. Shobha<sup>1</sup>, Dr. B. R. Karthikeyan<sup>2</sup>

Submitted: 10/10/2023

Revised: 30/11/2023

Accepted: 11/12/2023

**Abstract:** This paper aims to enhance the accuracy of Alzheimer's Disease (AD) versus Mild Cognitive Impairment (MCI) versus Normal Controls (NC) classification through the implementation of feature reduction techniques and the fusion of multimodal features. Specifically, gray matter within specified regions of interest (ROI) in the brain is extracted from both MRI and FDG-PET images. The LASSO feature selection technique is employed to identify relevant features crucial for distinguishing AD from MCI and NC. The reduction in features results in a 92.27% accuracy, reflecting an 11% improvement compared to classification without feature selection in AD versus MCI. The classification of AD versus MCI and MCI versus NC proves challenging due to the high correlations among features. The analysis reveals that a maximum classification accuracy of 92.27% is achieved for AD versus MCI through the multi-modal combination of features using a linear Support Vector Machine (SVM) algorithm. Additionally, a 99% accuracy is attained for MCI versus NC using the linear SVM algorithm. The fusion of features across all modalities yields a 94.9% accuracy for AD versus MCI versus NC. This analysis underscores that the fusion of multimodal features consistently improves classification accuracy compared to relying on any single modality. The study utilizes the ADNI-1 database, and the corresponding subject IDs are detailed in Table 10.

**Keywords:** Alzheimer Disease (AD), Mild Cognitive Impairment (MCI), Normal Controls (NC), Feature Selection.

## 1. Introduction

Alzheimer's Disease (AD) is a prevalent neurodegenerative condition affecting the elderly, progressing over time and causing the deterioration of brain cells leading to dementia. The evaluation of AD progression involves categorizing data into three stages: Normal Controls (NC), Mild Cognitive Impairment (MCI), and Alzheimer's Disease. It is noteworthy that not all subjects diagnosed with MCI will transition to AD; only a subset will convert over time [1]. Early detection of AD is crucial for clinicians to enhance the quality of life for individuals affected by the disease.

Various imaging modalities, including Magnetic Resonance Imaging (MRI), Fluorodeoxyglucose Positron Emission Tomography (FDG-PET), genetic analysis, and Cerebrospinal Fluid (CSF) examination, are available to diagnose the progression of AD. Structural MRI, widely used for its ability to differentiate between gray and white

matter in dementia subjects [6] [35], leverages changes in gray matter volume in the hippocampus region as a biomarker for AD diagnosis. The measurement of gray matter in predefined 116 Regions of Interest (ROIs) in the hippocampus serves as a feature vector for classification.

FDG-PET plays a significant role in the early diagnosis of AD, revealing reductions in glucose metabolic rates in various brain regions [23] [27]. Genetically, dementia is associated with permanent variations in the Apolipoprotein-E (APOE) gene, particularly the  $\epsilon_2$ ,  $\epsilon_3$ , and  $\epsilon_4$  alleles. Individuals with the APOE  $\epsilon_4$  allele have a higher risk of developing AD at an early age [18] [32]. In CSF, biomarkers such as  $A\beta_{38}$ ,  $A\beta_{40}$ ,  $A\beta_{42}$ , total tau, and phosphorylated tau are utilized for AD diagnosis. Notably, low levels of  $A\beta_{42}$  are validated in AD subjects, and the combination of  $A\beta_{40}$  and  $A\beta_{42}$  improves diagnostic accuracy [25].

Researchers have demonstrated the use of the  $A\beta_{42}/A\beta_{40}$  ratio for enhanced AD identification [13] [37]. Additionally, correlations between low levels of  $A\beta_{38}$  and elevated levels of phosphorylated tau (P-tau) [15] and total tau in CSF [4] in AD subjects have been investigated. Plaques and tangles, composed of amyloid beta fragments, obstruct neuron communication in dementia subjects [5],

<sup>1</sup> Research Scholar, Department of Electronics and Communication Engineering, Sapthagiri College of Engineering Bengaluru, shobhaparirishu@gmail.com.

<sup>2</sup> Associate Professor, Department of Electronics and Communication Engineering, M S Ramaiah University of Applied Sciences, Bengaluru, India. Karthikeyan.ec.et@msruas.ac.in

and a correlation between P-tau and tangle count has been reported [11].

While individual modalities provide limited information about dementia, the fusion of multi-modal features extracts diverse aspects of AD, offering complementary information. Integrating features from MRI, PET, CSF, APOE, and clinical features enhances the early detection of dementia [39]. Numerous researchers have developed computerized techniques for AD diagnosis, frequently utilizing MRI [8] [9] [14] [17] [26] and PET [10]. MRI provides detailed structural anatomy, while PET offers insight into metabolic brain function [24]. Consequently, the fusion of multimodal features holds promise for improving AD classification accuracy [12][16] [21].

This paper delves into the classification analysis of Alzheimer's Disease (AD) versus Mild Cognitive Impairment (MCI) versus Normal Controls (NC) subjects through the fusion of features from multiple modalities employing machine learning algorithms. Features from the Region of Interest (ROI) in both MRI and PET scans were extracted and utilized for the classification task. To streamline the process, the LASSO cross-validation technique was applied to identify relevant features and reduce their number. The investigation aimed to pinpoint optimal combinations of modalities for enhancing the accuracy of AD diagnosis. The reduced features from MRI and PET scans were amalgamated with data from the APOE gene, Cerebrospinal Fluid (CSF), and clinical information in various combinations to assess their impact on classification accuracy. The analysis underscores the consistent improvement in classification accuracy achieved through the fusion of multi-modality features, surpassing the efficacy of any single modality. Notably, the results highlight that while the APOE genotype, identified as a biomarker for AD, does not individually offer significant classification accuracy, its fusion with other modalities yields notable improvements in accuracy.

## 2.Dataset and Feature Extraction

The analysis presented in this paper utilized the ADNI LONI Image Data archives of the Alzheimer's Disease Neuroimaging Initiative (ADNI). The demographic details of the subjects selected for the analysis are provided in Table 1. To facilitate pre-processing and feature extraction from MRI and PET images, the SPM12 and CAT12 software toolboxes [3] were employed.

**Table 1.** Summary of Demographic characteristics of subjects

Group/ Gender	No. of Subjects	Age (Mean ± Std)	MMSE score (Mean ± Std)	GDS (Mean ± Std)	CDR (Mean ± Std)
AD Male/ Female	78	75.87 ± 7.6	23.50± 2.07	1.53 ± 1.32	1.00 ± 1.06
MCI Male/ Female	100	75.85 ± 6.81	28.92 ± 1.08	1.00 ± 1.27	0 ± 0
NC Male/ Female	96	75.81 ± 4.8	27.31 ± 2.27	1.77 ± 1.59	0.51 ± 0.09

The values are denoted as mean ± standard deviation.

MMSE=Mini Mental State Examination,

CDR=Clinical Dementia Rating, GDS=Global Deterioration Scale

To extract features from Magnetic Resonance Imaging (MRI) and Fluorodeoxyglucose Positron Emission Tomography (FDG-PET) images, a Region of Interest (ROI) Analysis [2] was executed utilizing an atlas specified by [22]. The FDG-PET images underwent pre-processing in accordance with standard protocols, involving a Dynamic 3D scan of six 5-minute frames conducted 30-60 minutes post-injection, as outlined in detail at (<http://adni.loni.usc.edu/methods/pet-analysis-method/pet-analysis/>). Following this pre-processing, the FDG-PET scans were co-registered with the corresponding pre-processed MRI scans (<http://adni.loni.usc.edu/methods/mri-tool/mri-analysis/>) using a three-dimensional method. The genetic component linked to Alzheimer's Disease is governed by the Apolipoprotein E (APOE) gene, characterized by three alleles: e2, e3, and e4. Each subject's genetic feature is defined by one of six genetic combinations: (e2, e2), (e2, e3), (e2, e4), (e3, e3), (e3, e4), and (e4, e4). The specific APOE measures utilized in the current study are detailed in Table 2.

**Table 2.** Each diagnostic group possessing each of the Six APOE allele pairs.

Genetic biomarkers	AD	MCI	NC
(e3, e3)	22	48	54
(e3, e4)	37	36	21
(e4, e4)	12	11	2
(e2, e3)	3	2	16
(e2, e4)	4	3	1
(e2, e2)			2
<b>Total Subjects</b>	<b>78</b>	<b>100</b>	<b>96</b>

The CSF biomarkers such as A $\beta$ 42, A $\beta$ 40, A $\beta$ 38, A $\beta$ , t-tau and p-tau were acquired using the notable fully automated Roche Elecsys and cobas e 601 immunoassay analyzer system from ADNI. The CSF biomarkers were not available for all the subjects, the missing data were compensated using mean value imputation technique. The measure of CSF biomarkers used in the classification of AD is depicted in Table 3.

**Table 3.** CSF measures and biomarkers used in classification of AD

CSF Biomarkers	AD	MCI	NC
A $\beta$ <sub>38</sub> , pg/ml	1718.20± 499.99	1747.11± 584.04	1754.81± 566.55
A $\beta$ <sub>40</sub> , pg/ml	7398± 2098.97	7319.08± 2306.10	7573.32± 2260.77
A $\beta$ <sub>42</sub> , pg/ml	651.17 ± 258.75	868.63± 456.40	1096.05± 404.72
A $\beta$	624.58± 267.26	943.39± 616.30	1174.17± 507.89
t-tau	363.06±	264.27±	227.23±

	141.00	90.67	77.59
p-tau	37.47897± 17.30	25.74± 9.87	21.165± 7.96

pg/ml=Picograms/milliliter

### 3. Feature Selection using LASSO Technique

The Least Absolute Shrinkage and Selection Operator (LASSO) feature selection selects the relevant features among the available features. The LASSO performs L1 regularization and was developed by Tibshirani [30], it identifies the most relevant features by performing cross validation of features against its class labels. This intern reduces the number of features, result in improving accuracy of the classifier. The authors have explored the LASSO based features selection on volumetric features and cortical thickness features to accomplish the improvement of classification Accuracy [28]. In the proposed analysis the relevant features of MRI and FDG-PET identified using LASSO technique are provided in Table 4.

**Table 4.** List of ROI biomarkers selected from MRI and PET images using LASSO Model

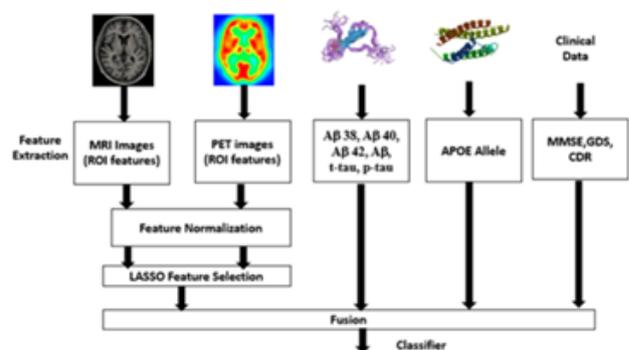
Sl. no	ROI biomarkers	MRI	PET	Sl. no	Volumetric Features	MRI	PET
1	Precentral_L			59	Parietal_Sup_L	√	
2	Precentral_R	√		60	Parietal_Sup_R	√	
3	Frontal_Sup_L			61	Parietal_Inf_L		
4	Frontal_Sup_R	√		62	Parietal_Inf_R		
5	Frontal_Sup_Orb_L	√		63	SupraMarginal_L		
6	Frontal_Sup_Orb_R			64	SupraMarginal_R		
7	Frontal_Mid_L			65	Angular_L		
8	Frontal_Mid_R			66	Angular_R		
9	Frontal_Mid_Orb_L			67	Precuneus_L		
10	Frontal_Mid_Orb_R			68	Precuneus_R		
11	Frontal_Inf_Oper_L			69	Paracentral_Lobule_L	√	√
12	Frontal_Inf_Oper_R			70	Paracentral_Lobule_R		
13	Frontal_Inf_Tri_L		√	71	Caudate_L	√	
14	Frontal_Inf_Tri_R	√		72	Caudate_R		√
15	Frontal_Inf_Orb_L			73	Putamen_L	√	
16	Frontal_Inf_Orb_R		√	74	Putamen_R	√	√
17	Rolandic_Oper_L			75	Pallidum_L	√	√
18	Rolandic_Oper_R		√	76	Pallidum_R		√
19	Supp_Motor_Area_L			77	Thalamus_L	√	
20	Supp_Motor_Area_R		√	78	Thalamus_R	√	
21	Olfactory_L			79	Heschl_L		
22	Olfactory_R		√	80	Heschl_R		
23	Frontal_Sup_Medial_L			81	Temporal_Sup_L		
24	Frontal_Sup_Medial_R			82	Temporal_Sup_R	√	
25	Frontal_Mid_Orb_L			83	Temporal_Pole_Sup_L	√	

26	Frontal_Mid_Orb_R			84	Temporal_Pole_Sup_R	√	
27	Rectus_L			85	Temporal_Mid_L		
28	Rectus_R			86	Temporal_Mid_R		
29	Insula_L			87	Temporal_Pole_Mid_L	√	√
30	Insula_R			88	Temporal_Pole_Mid_R	√	√
31	Cingulum_Ant_L			89	Temporal_Inf_L		
32	Cingulum_Ant_R			90	Temporal_Inf_R		
33	Cingulum_Mid_L			91	Cerebelum_Crus1_L		
34	Cingulum_Mid_R			92	Cerebelum_Crus1_R	√	
35	Cingulum_Post_L			93	Cerebelum_Crus2_L		
36	Cingulum_Post_R	√		94	Cerebelum_Crus2_R	√	√
37	Hippocampus_L	√		95	Cerebelum_3_L		
38	Hippocampus_R	√		96	Cerebelum_3_R	√	
39	ParaHippocampal_L		√	97	Cerebelum_4_5_L		
40	ParaHippocampal_R	√		98	Cerebelum_4_5_R	√	
41	Amygdala_L			99	Cerebelum_6_L	√	
42	Amygdala_R	√		100	Cerebelum_6_R	√	√
43	Calcarine			101	Cerebelum_7b_L		
44	Calcarine		√	102	Cerebelum_7b_R		
45	Cuneus			103	Cerebelum_8_L	√	√
46	Cuneus		√	104	Cerebelum_8_R		
47	Lingual			105	Cerebelum_9_L	√	
48	Lingual_R			106	Cerebelum_9_R	√	
49	Occipital_Sup_L			107	Cerebelum_10_L	√	
50	Occipital_Sup_R	√		108	Cerebelum_10_R	√	
51	Occipital_Mid_L			109	Vermis_1_2		
52	Occipital_Mid_R		√	110	Vermis_3		
53	Occipital_Inf_L			111	Vermis_4_5		
54	Occipital_Inf_R		√	112	Vermis_6		
55	Fusiform_L			113	Vermis_7	√	
56	Fusiform_R			114	Vermis_8	√	
57	Postcentral_L	√		115	Vermis_9	√	
58	Postcentral_R	√		116	Vermis_10	√	√

#### 4. Classification Performance Analysis

The comprehensive workflow integrated into the analysis is illustrated in Figure 1. This method encompasses four pivotal steps: Image pre-processing and Feature Extraction, Feature Selection employing the LASSO algorithm, Fusion, and Classification. All subjects included in the current analysis had complete MRI, PET, and APOE genotype data available. In instances where Cerebrospinal Fluid (CSF) data were lacking for specific subjects, the missing values were replaced with the mean value within each subgroup (AD, MCI, and NC group) [7]. The extracted features underwent normalization through the min-max normalization technique. Subsequently, the LASSO feature selection technique was applied to identify relevant features for subsequent classification utilizing a Support Vector Machine (SVM) classifier. The classification accuracy was computed by averaging over 100 iterations using 5-fold cross-validation. To evaluate

the performance of the classification model, Accuracy (ACC) and F1 score were computed in each iteration. The classification results derived from the proposed method are systematically tabulated in Table 6.



**Fig. 1.** The workflow of AD vs MCI vs NC identification using LASSO feature selection

**Table 6.** Classification results obtained using Linear SVM with Feature selection

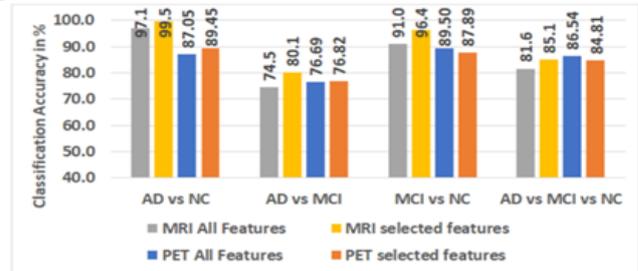
Sl. no	Modality	AD Vs NC		AD Vs MCI		MCI vs NC		AD Vs MCI Vs NC
		ACC	F1	ACC	F1	ACC	F1	ACC
1	MRI	99.49±0.37	99.43±0.41	80.13±1.75	75.8±2.18	96.35±1.99	96.54±1.83	85.09 ± 1.22
2	PET	89.45±4.38	89.63±3.95	76.82±1.98	78.6±1.46	87.89±0.74	87.86±0.73	84.81 ±0.81
3	MRI + PET	98.49±0.6	98.3±0.68	78.47±1.99	79.0±1.77	98.57±0.87	98.61±0.83	88.08 ±1.16
4	MRI + CSF	98.35±0.45	98.17±0.49	84.87±1.36	83.2±1.64	96.58±1.09	96.66±1.05	87.82 ±0.98
5	MRI + Genetic	99.28±0.44	99.19±0.5	79.26±1.72	74.7±2.28	96.68±1.07	96.81±1.01	84.34 ±1.14
6	MRI + Clinical	100±0	100±0	89.11±1.41	87.1±1.71	99.49±0	99.5±0	92.57 ±0.85
7	PET + CSF	95.94±0.6	95.55±0.65	81.71±2.07	82.0±1.71	84.95±0.8	84.93±0.77	83.81 ±0.79
8	PET + Genetic	90.65±4.3	90.62±3.94	76.17±1.9	78.0±1.45	80.59±1.53	80.57±1.55	78.59 ±1.47
9	PET + Clinical	99.99±0.08	99.99±0.09	86.11±2.98	86.2±2.56	99.49±0	99.5±0	97.26 ±0.5
10	MRI + PET + CSF	99.32±0.23	99.24±0.26	86.94±1.42	86.5±1.41	97.51±0.76	97.55±0.74	90.9 ± 0.9
11	MRI + PET + Genetic	98.22±0.63	97.99±0.72	79.41±2.05	79.9±1.79	98.06±0.63	98.11±0.6	87.75 ± 1.17
12	MRI + PET + Clinical	99.95±0.15	99.95±0.17	88.93±1.9	87.9±1.99	99.49±00	99.5±0.0	94.8 ± 4 ± 0.85
13	MRI + CSF + Clinical	99.79±0.28	99.76±0.31	88.78±1.23	87.3±1.44	99.49±0	99.5±0	92.22 ± 0.93
14	MRI + Genetic + Clinical	100±0	100±0	88.18±1.33	85.9±1.65	99.49±0	99.5±0	92.14 ± 0.84
15	PET + CSF + Genetic	95.48±0.57	95.06±0.62	82.49±2.15	82.6±1.79	85.15±0.82	85.15±0.8	83.78 ±0.71
16	PET + CSF + Clinical	98.81±0.41	98.68±0.45	92.87±0.8	92.2±0.82	99.49±0	99.5±0	95.99 ±0.43
17	MRI + PET + CSF + Genetic	99.11±0.3	99.01±0.34	87.98±1.31	87.5±1.32	96.79±0.82	96.86±0.79	90.68 ±0.89
18	MRI + PET + CSF + Clinical	100±0	100±0	92.15±1.2	91.3±1.32	99.53±0.14	99.54±0.14	95 ± 0.56
19	MRI +	99.98±0.11	99.97±0.13	89.37±1.74	88.4±1.83	99.4±0	99.5±	94.71

	PET+ Genetic + Clinical						0	$\pm 1.01$
20	MRI + CSF + Genetic+ Clinical	$99.74 \pm 0.3$	$99.71 \pm 0.33$	$88.6 \pm 1.31$	$87.1 \pm 1.56$	$99.49 \pm 0$	$99.5 \pm 0$	$92.55 \pm 0.89$
21	PET + CSF+ Genetic+ Clinical	$98.65 \pm 0.44$	$98.5 \pm 0.49$	$92.74 \pm 0.86$	$92.1 \pm 0.88$	$99.49 \pm 0$	$99.5 \pm 0$	$96.0 \pm 0.39$
22	MRI + PET + CSF + Genetic + Clinical	$99.99 \pm 0.08$	$99.99 \pm 0.09$	$92.27 \pm 1.29$	$91.6 \pm 1.38$	$99.49 \pm 0$	$99.5 \pm 0$	$94.99 \pm 0.8$

## 5 Results and Discussion

In the dataset encompassing individuals with Alzheimer's Disease (AD) and Normal Controls (NC), insights gleaned from relevant literature [19] [20] [29] [31] [33] [36] [38] highlight the achievement of a notable classification

accuracy exceeding 90%. This success is primarily attributed to the lower correlation of features between AD and NC. Conversely, in the case of Mild Cognitive Impairment (MCI) versus NC, the discernible features are comparatively fewer, resulting in reported accuracies predominantly below the 90% threshold [19] [20] [29] [31] [33] [36] [38]. Importantly, there is a noticeable gap in reported literature concerning the classification of AD versus MCI. The application of the LASSO feature reduction technique to MRI and PET features yielded a substantial reduction of 65% and 82%, respectively. Specifically, the original set of 116 features from MRI was streamlined to 41 features, while for PET, the reduction was from 116 features to a more concise 21 features. These reductions were derived from the list of Region of Interest (ROI) features. A comparative analysis of the Support Vector Machine (SVM) classifier performance, utilizing the original set of ROI features versus the reduced sets from MRI and PET, is visually presented in Figure 2. The classification performance of the original set of features and the reduced set of features for MRI and PET was compared against the fusion of MRI and PET features with modalities including CSF, Genetic, and Clinical data. The results indicate a notable improvement of 4 to 5% in classification accuracy with the feature reduction of MRI. However, in the case of PET, only a marginal 2% improvement is observed for AD versus NC, which is a relatively modest enhancement for classification. Moreover, the feature reduction is less effective for MCI versus NC and AD versus MCI versus NC.



**Fig. 2.** Comparison of classification accuracy with all 116 ROI features and Selected (41 features of MRI and 21 features of PET) ROI features

This paper underscores the importance of feature selection with LASSO to enhance classification accuracy, reporting the accuracy using the linear SVM technique with 5-fold cross-validation. With the proposed technique, using MRI as a single modality achieves an accuracy of 99.49% for AD versus NC and 96.35% for MCI versus NC. In comparison, literature utilizing Hierarchical classifier learning [29] reports an accuracy of 92.38% for AD versus NC and 84.24% for MCI versus NC, which is notably lower than the proposed technique. Similar comparisons were made for PET and other modalities against the classification accuracy reported in the literature. Detailed comparisons of accuracy for single and multi-modality fusion of features are provided in Tables 7, 8, and 9 for AD versus NC, MCI versus NC, and AD versus MCI versus NC, respectively. The results from Table 6 indicate that, for both single and multi-modality fusion of features, the proposed LASSO feature selection demonstrates improved classification accuracy compared to other methods reported in the literature. The proposed LASSO-based feature selection technique is also compared for multiclass classification of AD versus MCI versus NC against methods reported in the literature, showing a significant improvement in classification accuracy, as illustrated in Table 9.

The analysis also sought to compare the impact of the

proposed LASSO-based feature selection technique for both single and multi-modality fusion of features. The feature selection was applied to MRI and PET modalities and compared with the fusion of CSF, Genetic, and Clinical features for the classification accuracy of AD versus MCI versus NC. The results presented in Table 5 depict the accuracy and F1 score for all 116 ROI features of MRI and PET, while Table 6 describes the accuracy and F1 score for the selected 41 ROI features of MRI and 21 ROI features of PET, chosen using LASSO. Fusion of MRI and PET with other modalities was also compared, and the results in Table 6 validate the proposed feature selection technique. Notably, an improvement of 2% to 6% is achieved with the proposed LASSO feature selection in

both single and multi-modality fusion of features.

The results also highlight that genetic and clinical features as single modalities are not optimal for the classification of AD versus MCI versus NC. The analysis reveals that the fusion of MRI+PET+CSF combination yields the best accuracy for all three categories of classification. However, it is noteworthy that higher accuracy can be achieved by fusing clinical or genetic features. The analysis was conducted on the ADNI-1 dataset, and the selected patient IDs are provided in Table 10. The results presented in Table 5 and 6 are computed as an average accuracy of 100 iterations.

**Table 7.** Comparison of accuracy in classification of Alzheimer’s disease (AD) and cognitively normal (NC) subjects

Sl. no	Author	Data	AD/NC (count)	Algorithm	Validation method	ACC (%)
1.	Westman, E., et. al. 2012[36]	CSF	96/111	Orthogonal partial least squares to latent structures (OPLS)	7-Fold	81.6
		MRI				87.0
		MRI + CSF				91.8
2.	Zhang, D et.al.2012 [38]	MRI	45/50	SVM	10-Fold	84.80±2.6
		PET				84.5±3.5
		CSF				80.5±2.2
		MRI+PET +CSF				92.0±3.3
		MRI+PET +CSF		Multi-modal multi-task learning		93.3±2.2
3.	Suk, H. I.,2014 [29]	MRI	93/101	Hierarchical learning Scheme	10-Fold	92.38±5.32
		PET				92.20±6.70
		MRI+PET				95.35±5.23
4.	Lei, B et.al.,2016 [19]	MRI	93/101	Simple modality fusion and SVM	10-Fold	91.76± 6.14
		PET				90.89±5.81
		MRI + PET		Hybrid level fusion		94.4± 5.65
		MRI + PET				96.93±2.65
5.	Lei, B. et. al. 2017 [20]	MRI + PET + CSF	226/186	Support Vector Classification (SVC) by sigmoid kernel	10-fold	94.68
6.	Tong, T. et.al.,2017 [31]	MRI	37/35	Nonlinear Graph Fusion, Random Forest classifier	-	82.6
		PET				88.6
		MRI + PET				89.5
		MRI+ PET + CSF+ Genetic				91.8
		MRI + PET + CSF+ Genetic		SVM		91.4
7.	Wang, Z, et. al., 2017 [33]	MRI	93/101	Progressive Graph-based transductive learning (pGTL)	10-fold	88.6± 1.69
		PET				87.3± 1.47
		MRI + PET				92.6± 0.65

8.	Proposed method	MRI	78/96	LASSO feature selection, feature normalization and SVM classifier	5-fold	99.49± 0.37
		PET				89.45± 4.38
		MRI + PET				98.49± 0.6
		MRI + CSF				98.35± 0.45
		MRI + PET + CSF				99.32± 0.23
		MRI + PET + CSF+ Genetic				99.11± 0.3
		MRI + PET + CSF +Genetic + Cognitive				99.99± 0.08

**Table 8.** Comparison of accuracy in classification of Mild cognitive impairment (MCI) and cognitively normal (NC) subjects

Sl. no	Author	Data	MCI/NC (count)	Algorithm	Validation method	ACC (%)
1.	Westman, E et. al.,2012 [36]	CSF	162/111	orthogonal partial least squares to latent structures (OPLS)	7-Fold	70.3
		MRI				71.8
		MRI+CSF				77.6
2.	Zhang, D et.al.2012 [38]	MRI	91/50	SVM	10-Fold	73.9±2.8
		PET				79.7 ±2.3
		CSF				53.6±4.4
		MRI + PET +CSF		80.0±2.4		
		MRI + PET +CSF		83.2±1.5		
3.	Suk, H. I.,2014 [29]	MRI	204/101	Hierarchical learning scheme	10-Fold	84.24± 6.26
		PET				84.29±7.22
		MRI+PET				85.67±5.22
4.	Lei, B et.al.,2016 [19]	MRI	204/101	Simple modality fusion and SVM	10-Fold	83.52± 5.38
		PET				82.95±6.37
		MRI + PET		83.67±5.49		
		MRI + PET		86.57±4.72		
5.	Lei, B. et. al. 2017 [20]	MRI+PET+CSF	393/186	support vector classification (SVC) by sigmoid kernel	10-fold	80.32
6.	Tong, T. et.al.,2017 [31]	MRI	75/35	Nonlinear Graph Fusion, Random Forest classifier	-	73.3
		PET				75.4
		MRI+PET				76.7
		MRI + PET + CSF + Genetic				79.5
		MRI + PET + CSF + Genetic		SVM		77.4
7.	Wang, Z, et. al., 2017 [33]	MRI	102/101	Progressive Graph-based transductive learning (pGTL)	10-fold	70.7 ± 0.81
		PET				72.5 ± 0.76
		MRI+PET				78.9 ± 1.80
8.	Proposed method	MRI	100/96	LASSO feature selection,	5-fold	96.35± 1.99
		PET				87.89± 0.74
		MRI + PET				98.57± 0.87

		MRI + CSF		feature normalization and SVM classifier		96.58± 1.09	
		MRI + PET + CSF					97.51± 0.76
		MRI + PET + CSF + Genetic					96.79± 0.82
		MRI + PET + CSF + Genetic + Cognitive					99.49 ± 0

**Table 9.** Comparison of accuracy in classification of Alzheimer Disease (AD), Mild cognitive impairment (MCI) and cognitively normal (NC) subjects.

Sl no	Author	Data	AD/MCI/NC (count)	Algorithm	Validation method	ACC (%)
1.	Tong, T. et.al.,2017 [31]	MRI	37/75/35	Random forest classifier	-	56.3
		PET			-	56.5
		MRI+PET			-	58.2
		MRI + PET + CSF + Genetic			-	60.2
		MRI + PET + CSF + Genetic		SVM	-	59.6
2.	Proposed method	MRI	78/100/96	LASSO feature selection, feature normalization and SVM classifier	5-fold	85.09±1.22
		PET				84.81±0.81
		MRI + PET				88.08±1.16
		MRI + CSF				87.82±0.98
		MRI + PET + CSF				90.09±0.9
		MRI + PET + CSF + Genetic				90.68±0.89
		MRI + PET + CSF + Genetic + Cognitive				94.99 ± 0.8

## 6. Conclusion

This paper explores and evaluates the LASSO feature selection method to enhance the classification of MRI and PET Region of Interest (ROI) features for Alzheimer's Disease (AD) versus Mild Cognitive Impairment (MCI) versus Normal Controls (NC). The analysis involves the feature selection of ROI features from MRI and PET, coupled with the fusion of Cerebrospinal Fluid (CSF),

Genetic, and Clinical features. The fusion of features, wherein clinical and genetic data are combined with conventional modalities such as MRI, PET, and CSF, demonstrates an improvement in classification accuracy.

The proposed LASSO-based feature selection and multimodality fusion result in an 11% improvement in the

classification accuracy of AD versus MCI and a 2% improvement in AD versus MCI versus NC, compared to scenarios without employing any feature selection technique on the data.

## Acknowledgments

adni.loni.usc.edu. The ADNI database provided the data for this article, and it's important to note that the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data. However, they did not participate in the analysis or writing of this report. For a comprehensive list of ADNI investigators, please refer to the website: [adni.loni.usc.edu/http://adni.loni.usc.edu/wpcontent/uploads/how\\_to\\_apply/ADNI\\_Acknowledgement\\_List.pdf](http://adni.loni.usc.edu/http://adni.loni.usc.edu/wpcontent/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf)<http://adni.loni.usc.edu/wp->

## Appendix

**Table 10.** Chosen Patient IDs of the ADNI-1 Dataset

Categories	ID's of Subjects
<b>AD (78)</b>	1059, 1257, 221, 929, 1341, 316, 1339, 3, 1205, 991, 286, 682, 213, 343, 642, 1109, 219, 543, 1171, 1307, 850, 1254, 836, 1056, 321, 554, 147, 400, 1037, 889,1281, 1283, 1285, 341, 577, 760, 1001, 627, 1368, 1391, 1044, 474, 1371, 1379, 535, 690, 730, 565, 1164, 1397, 1402, 149, 470, 492, 1144, 747, 1062, 777, 1157, 228, 374, 979, 370, 891, 1221, 431, 754, 1382, 167, 216, 266, 740, 1409, 1290, 497, 438, 841, 1041
<b>MCI (100)</b>	222, 546, 675, 128, 293, 344, 1030, 1199, 326, 362, 861, 1282, 634, 917, 932, 1033, 1175, 240, 325, 860, 1120, 1186, 1275, 80, 142, 155, 282, 314, 407, 446, 549, 598,679, 721, 1010, 1425, 1346, 464, 941, 957, 1007, 1217, 1294, 1299, 1211, 1380, 746, 909, 1357, 641, 282, 314, 407, 446, 549, 598, 679, 721, 1010, 1425, 941, 957, 1007, 1217, 1265, 1294, 1299, 1211, 1380,1357, 641, 531, 958, 1034,892, 930, 995, 950, 1114, 1343, 378, 410, 1103, 1106, 1118, 361, 1243, 1315, 1322, 708, 709, 865, 1077,112, 394, 925, 1210, 1419, 1427, 135
<b>NC (96)</b>	223, 610, 484, 731, 751, 842, 862, 67, 419, 420, 2, 5, 8, 16, 21, 23, 637, 1133, 502, 359, 43, 55, 97, 883, 647, 14, 96, 130, 1063, 74, 120, 843, 845, 866, 618, 95, 734, 741, 48, 555, 576, 672, 813, 327, 454, 262, 779, 818, 934, 575, 1023, 467, 768, 1099, 315, 311, 312, 386, 363, 489, 526, 171, 90, 352,533, 534, 47, 967,1013, 173, 416, 360, 648, 657, 506, 680, 259, 230, 245, 272, 500, 522, 863, 778, 232, 1200, 319, 301,459, 686, 972,1194, 1195, 1197, 1202, 1203

### References

- [1] 2021 Alzheimer's disease facts and figures. (2021). *Alzheimer's and Dementia*, 17(3), 327–406. <https://doi.org/10.1002/alz.12328>
- [2] Alzheimer, E., Mcevoy, L. K., Fennema-notestine, C., Jennings, R. G., Brewer, J. B., Hoh, C. K., & Dale, A. M. (2010). Relative Capability of MR Imaging and FDG PET to Depict Changes Associated with Purpose: Methods: Results. *Radiology*, 256(3), 932–942. <https://doi.org/10.1148/radiol.10091402/-/DC1>
- [3] Ashburner, J. (2009). Computational anatomy with the SPM software. *Magnetic Resonance Imaging*, 27(8), 1163–1174. <https://doi.org/10.1016/j.mri.2009.01.006>
- [4] Blennow, K., & Hampel, H. (2003). Review CSF markers for incipient Alzheimer's disease CSF markers for incipient AD. 2(October), 605–613.
- [5] Buerger, K., Ewers, M., Pirttilä, T., Zinkowski, R., Alafuzoff, I., Teipel, S. J., DeBernardis, J., Kerkman, D., McCulloch, C., Soininen, H., & Hampel, H. (2006). CSF phosphorylated tau protein correlates with neocortical neurofibrillary pathology in Alzheimer's disease. *Brain*, 129(11), 3035–3041. <https://doi.org/10.1093/brain/awl269>
- [6] Cho, Y., Seong, J. K., Jeong, Y., & Shin, S. Y. (2012). Individual subject classification for Alzheimer's disease based on incremental learning using a spatial frequency representation of cortical thickness data. *NeuroImage*, 59(3), 2217–2230. <https://doi.org/10.1016/j.neuroimage.2011.09.085>
- [7] Coley, N., Gardette, V., Cantet, C., Gillette-Guyonnet, S., Nourhashemi, F., Vellas, B., & Andrieu, S. (2011). How Should We Deal with Missing Data in Clinical Trials Involving Alzheimer's Disease Patients? *Current Alzheimer Research*, 999(999), 1–12. <https://doi.org/10.2174/1567211212443482050>
- [8] Cuingnet, R., Gerardin, E., Tessieras, J., Auzias, G., Lehéricy, S., Habert, M. O., Chupin, M., Benali, H., & 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(2), 766–781. <https://doi.org/10.1016/j.neuroimage.2010.06.013>
- [9] De Magalhães Oliveira, P. P., Nitrini, R., Busatto, G., Buchpiguel, C., Sato, J. R., & Amaro, E. (2010). Use of SVM methods with surface-based cortical and volumetric subcortical measurements to detect Alzheimer's disease. *Journal of Alzheimer's Disease*, 19(4), 1263–1272. <https://doi.org/10.3233/JAD-2010-1322>
- [10] Gray, K. R., Wolz, R., Heckemann, R. A., Aljabar, P., Hammers, A., & Rueckert, D. (2012). Multi-region analysis of longitudinal FDG-PET for the classification of Alzheimer's disease. *NeuroImage*, 60(1), 221–229. <https://doi.org/10.1016/j.neuroimage.2011.12.071>
- [11] Grundke-Iqbal, I., Iqbal, K., & Tung, Y. C. (1986). Abnormal phosphorylation of the microtubule-associated protein  $\tau$  (tau) in Alzheimer cytoskeletal pathology. *Proceedings of the National Academy of Sciences of the United States of America*, 83(13), 44913–44917. <https://doi.org/10.1097/00002093-198701030-00020>

- [12] Gupta, Y., Lama, R. K., & Kwon, G. R. (2019). Prediction and Classification of Alzheimer's Disease Based on Combined Features From Apolipoprotein-E Genotype, Cerebrospinal Fluid, MR, and FDG-PET Imaging Biomarkers. *Frontiers in Computational Neuroscience*, 13(October), 1–18. <https://doi.org/10.3389/fncom.2019.00072>
- [13] Hansson, O., Zetterberg, H., Buchhave, P., Andreasson, U., Londos, E., Minthon, L., & Blennow, K. (2007). Prediction of Alzheimer's disease using the CSF A $\beta$ 42/A $\beta$ 40 ratio in patients with mild cognitive impairment. *Dementia and Geriatric Cognitive Disorders*, 23(5), 316–320. <https://doi.org/10.1159/000100926>
- [14] Hao, X., Bao, Y., Guo, Y., Yu, M., Zhang, D., Risacher, S. L., Saykin, A. J., Yao, X., & Shen, L. (2020). Multi-modal neuroimaging feature selection with consistent metric constraint for diagnosis of Alzheimer's disease. *Medical Image Analysis*, 60, 101625. <https://doi.org/10.1016/j.media.2019.101625>
- [15] Heywood, W. E., Hallqvist, J., Heslegrave, A. J., Zetterberg, H., Fenoglio, C., Scarpini, E., Rohrer, J. D., Galimberti, D., & Mills, K. (2018). CSF pro-orexin and amyloid- $\beta$ 38 expression in Alzheimer's disease and frontotemporal dementia. *Neurobiology of Aging*, 72, 171–176. <https://doi.org/10.1016/j.neurobiolaging.2018.08.019>
- [16] Hinrichs, C., Singh, V., Xu, G., & Johnson, S. C. (2011). Predictive markers for AD in a multi-modality framework: An analysis of MCI progression in the ADNI population. *NeuroImage*, 55(2), 574–589. <https://doi.org/10.1016/j.neuroimage.2010.10.081>
- [17] Hu, K., Wang, Y., Chen, K., Hou, L., & Zhang, X. (2015). Multi-scale features extraction from baseline structure MRI for MCI patient classification and AD early diagnosis. *Neurocomputing*, 175(PartA), 132–145. <https://doi.org/10.1016/j.neucom.2015.10.043>
- [18] Jacquier, M., Arango, D., Villareal, E., Torres, O., Serrano, M. L., Cruts, M., Montañes, P., Cano, C., Rodriguez, M. N., Serneels, S., & Van Broeckhoven, C. (2001). APOE  $\epsilon$ 4 and Alzheimer's disease: Positive association in a Colombian clinical series and review of the Latin-American studies. *Arquivos de Neuro-Psiquiatria*, 59(1), 11–17. <https://doi.org/10.1590/s0004-282x2001000100004>
- [19] Lei, B., Chen, S., Ni, D., & Wang, T. (2016). Discriminative learning for Alzheimer's disease diagnosis via canonical correlation analysis and multimodal fusion. *Frontiers in Aging Neuroscience*, 8(MAY), 1–17. <https://doi.org/10.3389/fnagi.2016.00077>
- [20] Lei, B., Yang, P., Wang, T., Chen, S., & Ni, D. (2017). Relational-Regularized Discriminative Sparse Learning for Alzheimer's Disease Diagnosis. *IEEE Transactions on Cybernetics*, 47(4), 1102–1113. <https://doi.org/10.1109/TCYB.2016.2644718>
- [21] Li, Y., Meng, F., & Shi, J. (2019). Learning using privileged information improves neuroimaging-based CAD of Alzheimer's disease: a comparative study. *Medical and Biological Engineering and Computing*, 57(7), 1605–1616. <https://doi.org/10.1007/s11517-019-01974-3>
- [22] Maldjian, J. A., Laurienti, P. J., Kraft, R. A., & Burdette, J. H. (2003). An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *NeuroImage*, 19(3), 1233–1239. [https://doi.org/10.1016/S1053-8119\(03\)00169-1](https://doi.org/10.1016/S1053-8119(03)00169-1)
- [23] McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C. R., Kawas, C. H., Klunk, W. E., Koroshetz, W. J., Manly, J. J., Mayeux, R., Mohs, R. C., Morris, J. C., Rossor, M. N., Scheltens, P., Carrillo, M. C., Thies, B., Weintraub, S., & Phelps, C. H. (2011). The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's and Dementia*, 7(3), 263–269. <https://doi.org/10.1016/j.jalz.2011.03.005>
- [24] Riederer, I., Bohn, K. P., Preibisch, C., Wiedemann, E., Zimmer, C., Alexopoulos, P., & Förster, S. (2018). Alzheimer disease and mild cognitive impairment: Integrated pulsed arterial spin-labeling MRI and 18F-FDG PET. *Radiology*, 288(1), 198–206. <https://doi.org/10.1148/radiol.2018170575>
- [25] Shoji, M., Matsubara, E., Kanai, M., Watanabe, M., Nakamura, T., Tomidokoro, Y., Shizuka, M., Katsumi, W., Igeta, Y., Ikeda, Y., Mizushima, K., Amari, M., Ishiguro, K., Kawarabayashi, T., Harigaya, Y., Okamoto, K., & Hirai, S. (1998). Combination assay of CSF Tau, A $\beta$ 1-40 and A $\beta$ 1-42 (43) as a biochemical marker of Alzheimer's disease. *J Neurol Sci.*, 158, 134–140.
- [26] Sørensen, L., Igel, C., Pai, A., Balas, I., Anker, C., Lillholm, M., & Nielsen, M. (2017). Differential diagnosis of mild cognitive impairment and Alzheimer's disease using structural MRI cortical thickness, hippocampal shape, hippocampal texture, and volumetry. *NeuroImage: Clinical*, 13, 470–482. <https://doi.org/10.1016/j.nicl.2016.11.025>
- [27] Sperling, R. A., Aisen, P. S., Beckett, L. A., Bennett, D. A., Craft, S., Fagan, A. M., Iwatsubo, T., Jack, C. R., Kaye, J., Montine, T. J., Park, D. C., Reiman, E. M., Rowe, C. C., Siemers, E., Stern, Y., Yaffe, K., Carrillo, M. C., Thies, B., Morrison-Bogorad, M., ... Phelps, C. H. (2011). Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's*

and *Dementia*, 7(3), 280–292.  
<https://doi.org/10.1016/j.jalz.2011.03.003>

[28] S. S and B. R. Karthikeyan, "Feature Selection on MRI Data for Improved Classification of Alzheimer Disease," 2021 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT), 2021, pp. 1-6, doi: 10.1109/CONECCT52877.2021.9622551.

[29] Suk, H. Il, Lee, S. W., & Shen, D. (2014). Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis. *NeuroImage*, 101, 569–582.  
<https://doi.org/10.1016/j.neuroimage.2014.06.077>

[30] Tibshirani, R. (1996). Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), 267–288.  
<https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>

[31] Tong, T., Gray, K., Gao, Q., Chen, L., & Rueckert, D. (2017). Multi-modal classification of Alzheimer's disease using nonlinear graph fusion. *Pattern Recognition*, 63(May 2016), 171–181.  
<https://doi.org/10.1016/j.patcog.2016.10.009>

[32] Verghese, P. B., Castellano, J. M., & Holtzman, D. M. (2011). Apolipoprotein E in Alzheimer's disease and other neurological disorders. *The Lancet Neurology*, 10(3), 241–252. [https://doi.org/10.1016/S1474-4422\(10\)70325-2](https://doi.org/10.1016/S1474-4422(10)70325-2)

[33] Wang, Z., Zhu, X., Adeli, E., Zhu, Y., Nie, F., Munsell, B., & Wu, G. (2017). Multi-modal classification of neurodegenerative disease by progressive graph-based transductive learning. *Medical Image Analysis*, 39, 218–230. <https://doi.org/10.1016/j.media.2017.05.003>

[34] Welge, V., Fiege, O., Lewczuk, P., Mollenhauer, B., Esselmann, H., Klafki, H. W., Wolf, S., Trenkwalder, C., Otto, M., Kornhuber, J., Wiltfang, J., & Bibl, M. (2009). Combined CSF tau, p-tau181 and amyloid- $\beta$  38/40/42 for diagnosing Alzheimer's disease. *Journal of Neural Transmission*, 116(2), 203–212.  
<https://doi.org/10.1007/s00702-008-0177-6>

[35] Westman, E., Aguilar, C., Muehlboeck, J. S., & Simmons, A. (2013). Regional magnetic resonance imaging measures for multivariate analysis in Alzheimer's disease and mild cognitive impairment. *Brain Topography*, 26(1), 9–23. <https://doi.org/10.1007/s10548-012-0246-x>

[36] Westman, E., Muehlboeck, J. S., & Simmons, A. (2012). Combining MRI and CSF measures for classification of Alzheimer's disease and prediction of mild cognitive impairment conversion. *NeuroImage*, 62(1), 229–238.  
<https://doi.org/10.1016/j.neuroimage.2012.04.056>

[37] Wiltfang, J., Esselmann, H., Bibl, M., Hüll, M.,

Hampel, H., Kessler, H., Frölich, L., Schröder, J., Peters, O., Jessen, F., Luckhaus, C., Perneczky, R., Jahn, H., Fiszer, M., Maler, J. M., Zimmermann, R., Bruckmoser, R., Kornhuber, J., & Lewczuk, P. (2007). Amyloid  $\beta$  peptide ratio 42/40 but not A $\beta$ 42 correlates with phospho-Tau in patients with low- and high-CSF A $\beta$ 40 load. *Journal of Neurochemistry*, 101(4), 1053–1059.  
<https://doi.org/10.1111/j.1471-4159.2006.04404.x>

[38] Zhang, D., & Shen, D. (2012). Multi-modal multi-task learning for joint prediction of multiple regression and classification variables in Alzheimer's disease. *NeuroImage*, 59(2), 895–907.  
<https://doi.org/10.1016/j.neuroimage.2011.09.069>

[39] Zhang, D., Wang, Y., Zhou, L., Yuan, H., & Shen, D. (2011). Multimodal classification of Alzheimer's disease and mild cognitive impairment. *NeuroImage*, 55(3), 856–867.  
<https://doi.org/10.1016/j.neuroimage.2011.01.008>