Machine Learning Techniques for the Diagnosis of Alzheimer's Disease: A Review

Article *in* ACM Transactions on Multimedia Computing, Communications and Applications · April 2020

DOI: 10.1145/3344998

CITATIONS
14

READS
4,622

7 authors, including:

M. Tanveer
Indian Institute of Technology Indore
80 PUBLICATIONS **861** CITATIONS
SEE PROFILE

Bharat Richhariya
Indian Institute of Technology Indore
17 PUBLICATIONS **193** CITATIONS
SEE PROFILE

Riyaj Uddin Khan
Indian Institute of Technology Guwahati
4 PUBLICATIONS **14** CITATIONS
SEE PROFILE

A.H. Rashid
National Institute of Science & Technology
2 PUBLICATIONS **23** CITATIONS
SEE PROFILE

Machine learning techniques for the diagnosis of Alzheimer's disease: A review

M. TANVEER*, B. RICHHARIYA, and R.U. KHAN, Discipline of Mathematics, Indian Institute of Technology Indore, Simrol, Indore, India

A.H. RASHID, Discipline of Mathematics, Indian Institute of Technology Indore, Simrol, Indore & School of Computer Science and Engineering, National Institute of Science and Technology, Berhampur, Odisha, India

P. KHANNA, PDPM Indian Institute of Information Technology, Design and Manufacturing, Jabalpur, India

M. PRASAD and C.T. LIN, Centre for Artificial Intelligence, School of Computer Science, FEIT, University of Technology Sydney, Sydney, Australia, Australia

Alzheimer's disease is an incurable neurodegenerative disease primarily affecting the elderly population. Efficient automated techniques are needed for early diagnosis of Alzheimers. Many novel approaches are proposed by researchers for classification of Alzheimer's disease. However, to develop more efficient learning techniques, better understanding of the work done on Alzheimers is needed. Here, we provide a review on 165 papers from 2005-2019 using various feature extraction and machine learning techniques. The machine learning techniques are surveyed under three main categories: support vector machine (SVM), artificial neural network (ANN), and deep learning (DL) and ensemble methods. We present a detailed review on these three approaches for Alzheimers with possible future directions.

CCS Concepts: • **Dementia** → Alzheimer's disease; • **Machine learning** → Support vector machine; Artificial neural network; Deep learning; Ensemble methods.

Additional Key Words and Phrases: Magnetic resonance imaging (MRI), positron emission tomography (PET), diffusion tensor imaging (DTI), mild cognitive impairment (MCI).

ACM Reference Format:
M. Tanveer, B. Richhariya, R.U. Khan, A.H. Rashid, P. Khanna, M. Prasad, and C.T. Lin. 2020. Machine learning techniques for the diagnosis of Alzheimer's disease: A review. *ACM Trans. Multimedia Comput. Commun. Appl.* 16, 1s, Article 30 (April 2020), 28 pages. https://doi.org/https://doi.org/10.1145/3344998

*Corresponding author

Authors' addresses: M. Tanveer, mtanveer@iiti.ac.in; B. Richhariya, phd1701241001@iiti.ac.in; R.U. Khan, riyaz@iiti.ac.in, Discipline of Mathematics, Indian Institute of Technology Indore, Simrol, Indore, 453552, India; A.H. Rashid, ashrafrashid102@gmail.com, Discipline of Mathematics, Indian Institute of Technology Indore, Simrol, Indore & School of Computer Science and Engineering, National Institute of Science and Technology, Berhampur, Odisha, 761008, India; P. Khanna, pkhanna@iiitdmj.ac.in, PDPM Indian Institute of Information Technology, Design and Manufacturing, Jabalpur, 482005, India; M. Prasad, mukesh.prasad@uts.edu.au; C.T. Lin, chintenglin@gmail.com, Centre for Artificial Intelligence, School of Computer Science, FEIT, University of Technology Sydney, Sydney, Australia, Australia.

1 INTRODUCTION

Alzheimer's disease (AD) is one of the most common cause of dementia in today's world. According to World Alzheimer Report (2018) [126], around 50 million people were affected by this disease in 2018, which is expected to triple by 2050. Usually, the symptoms of Alzheimers are visible after 60 years of age [43]. However, some forms of AD develop very early (30-50 years) for individuals having gene mutation [10]. Alzheimer's disease gives rise to structural and functional changes in the brain. In AD patients, the time between healthy state to Alzheimers spans over many years [180]. First, patients develop mild cognitive impairment (MCI), and gradually progress to Alzheimers. However, all MCI patients do not convert to Alzheimers [37]. So, the main focus of current research is to predict the conversion of MCI to AD. These changes can be measured using medical imaging [138] and other techniques like blood plasma spectroscopy [39, 125].

Many open source databases for Alzheimers have accelerated research in this field [67, 181]. The most widely used databases are ADNI [174] (adni.loni.usc.edu), AIBL (aibl.csiro.au), OASIS (www.oasis-brains.org). A new publicly available database for clinical Alzheimer data is J-ADNI database [44, 66] containing data from longitudinal studies in Japan. Further, processing of MRI images requires a lot of effort. To facilitate analysis of MRI images open source softwares like Statistical Parametric Mapping (SPM) have been developed by Wellcome Centre for Human Neuroimaging for public use. SPM is used for voxel based morphometry (VBM) [77] of MRI data. Another very popular open source software i.e., Freesurfer [36] is developed for volume based morphometry and is used by many researchers [4, 167].

Machine learning techniques are found to be very useful for the diagnosis of Alzheimers [118, 128, 170] in the last decade. The most widely used classification techniques are support vector machine (SVM), artificial neural network (ANN), and deep learning. The primary difference between SVM and ANN is the nature of the optimization problem. SVM gives a globally optimal solution [15], while ANN gives locally optimal solution. In both SVM and ANN, feature extraction is an important step. Shi et al. [154] suggested that combination of neural networks and intelligent agents can be useful for medical image analysis. However, deep learning incorporates the feature extraction step in the learning model itself [151, 159]. For large datasets, deep learning is found to be useful especially for image data [151]. Some researchers also used ensemble methods to improve the classification accuracy for Alzheimers [28, 100, 131].

For classification of Alzheimer's data, the accuracy is dependent on the type of problem. For example, the accuracy is highest for Control normal (CN) vs AD, lesser for CN vs MCI, and least for MCI vs AD [11]. Moreover, the classification of MCI converters (MCIc) vs non-converters (MCInc), and amnestic MCI (aMCI) vs non-amnestic MCI (naMCI) is also a challenging task [32, 119]. Moreover, the data generated from MRI scanners is 3-D in nature and thus amounts to large sized datasets. So, efficient feature extraction and classification techniques are needed to analyze this data [46, 200].

In recent years, researchers provided analysis on the works done using machine learning for Alzheimers. In 2017, Litjens et al. [96] presented a review on deep learning methods for medical image analysis. It is mentioned that although deep learning models are considered as 'black boxes', some statistical techniques can be used to estimate uncertainty of the network. Shen et al. [151] performed a survey on deep learning for Alzheimers. It also supported this fact of uncertainty in prediction by deep learning models. In 2018, Jose et al. provided a review on neuroimaging techniques for brain disorders. It is stated that machine learning techniques can be useful for finding the underlying neurological causes of brain disorders [112]. Pellegrini et al. [128] discussed the machine learning techniques used for dementia and cognitive impairment from 2006-2016 in 111 papers. It stressed on the development of novel machine learning models from interdisciplinary approach. Rathore et al. [138] also provided a review on feature extraction and classification of Alzheimers and its prodromal stages.

Among the different machine learning techniques, we selected three main approaches namely, support vector machine (SVM), artificial neural network (ANN), and deep learning (DL) methods and ensemble methods. In this review, we present a separate analysis for each of these techniques on diagnosis of Alzheimers. Section 3, 4 and 5 provide a survey on application of the three techniques on classification of Alzheimer's disease. Section 6 dwells on the possible future directions, while section 7 presents the conclusions of the review.

2 SEARCH STRATEGY

We searched prominent papers in the field from Google Scholar (https://scholar.google.co.in/) and Sciencedirect (https://www.sciencedirect.com/). We excluded the studies which did not use accuracy measures for classification performance. This resulted in a total of 165 papers. Out of 165, 60 papers used SVM, 45 used a combination of ANN, multi-task learning, transfer learning, multi-kernel learning and certain feature selection techniques. We also included 60 papers based on deep learning and ensemble methods for Alzheimers. Papers using SVM are from the period 2005-2019, ANN are from 2008-2019, and deep learning and ensemble methods are from 2007-2019.

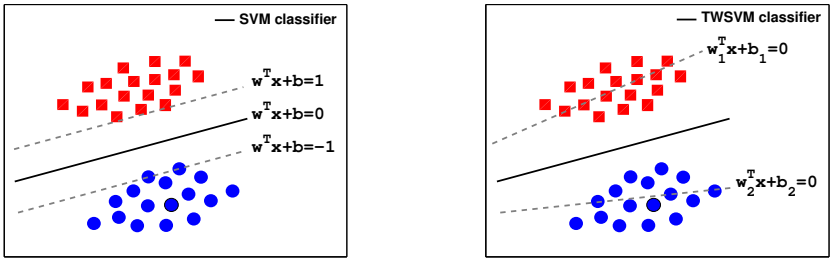
3 SUPPORT VECTOR MACHINE

Support vector machine (SVM) [29] is a very stable [194] and widely used technique for classification and regression problems [47, 140, 141, 189]. By including the structural risk minimization principle (SRM), SVM gives good generalization performance. SVM uses the maximum margin principle to classify the data points as shown in fig. 1(a). After solving a convex optimization problem, the decision function of SVM is written as,

$$f(x) = sign\ (w^T x + b), \tag{1}$$

where w is the weight vector and b is bias.

To classify non-linearly separable data, kernel functions [30] have been used to transform the data to higher dimensions. Moreover, various variants of SVM have been proposed to increase its performance w.r.t. generalization ability and training time [171]. Some computationally efficient variants of SVM are twin support vector machine (TWSVM) [69] shown in fig. 1(b), and least squares based twin support vector machine (LSTSVM) algorithms [86, 168].



(a) (b)

Fig. 1. Plot showing (a) SVM, and (b) TWSVM classifiers.

The details of all papers on SVM surveyed in this work are presented in Table 1. We present a detailed analysis of application of SVM for Alzheimers on various parameters in the following subsections.

3.1 Image modality

Image modality is a prominent factor for classification of MRI images. In case of structural MRI (sMRI) images, most of the researchers used T1-weighted images while only few researchers used T2 images [18, 75, 87]. This is because the delineation of ventricular surface of brain due to atrophy is clearly visible in T1-weighted images [176]. Fig. 2 shows the use of different modalities of data in classification of Alzheimers using SVM.

Fan et al. [41] suggested that positron emission tomography (PET) scans provide complementary information to sMRI scans, thus improving the classification accuracy of CN vs MCI using SVM. Dukart et al. [40] supported this fact that fluorodeoxyglucose-PET (FDG-PET) features are more discriminative as compared to sMRI. Further, better accuracy is found for CN vs AD [7] with PET images (100 %) as compared to single photon emission computed tomography (SPECT) images (97.5 %). Similar finding is observed for CN vs AD [150] with better accuracy for PET images (96.67 %) as compared to SPECT images (94.5 %). Kamathe et al. [75] used combination of T1, T2 and proton density (PD) scans for classification of CN vs AD. Hojjati et al. [58] used resting state functional MRI (rs-fMRI) to find the connectivity changes in brain for classification of MCIc vs MCInc, while Sheng et al. [152] used connectivity information from fMRI data. Fig. 2 shows the usage of image modality for SVM in our survey.

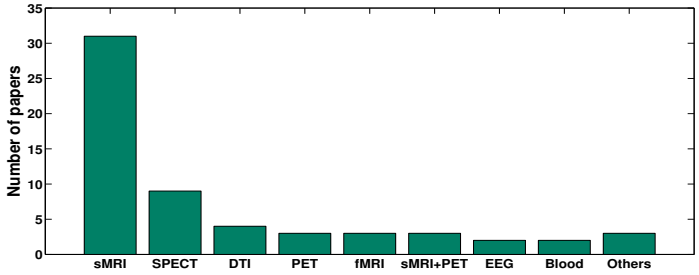


Fig. 2. Plot showing different image modalities and other data used with SVM for Alzheimers.

Diffusion tensor imaging (DTI) is also explored by various researchers for Alzheimer's disease [91, 119, 194]. Haller et al. [53] found that SVM based analysis of white matter DTI parameters is helpful in classification of different types of MCI patients.

3.2 Feature selection and extraction with SVM

Feature selection plays an important part in the classification of data. Different features are combined to form the feature vector in many works [58, 109, 120, 124, 193]. Vemuri et al. [175] found that including demographic and genetic information with sMRI scans improved the classification accuracy of CN vs AD. A refined parcellation method is proposed [115] for detecting subtle changes in gray matter (GM). Magnin et al. [109] presented a feature selection method based on histogram of regions of interests (ROIs) for CN vs AD. Gerardin et al. [48] used shape features of hippocampus to discriminate CN, MCI and AD, and found that shape deformation features are better than volumetric features. Normalized mean square error (NMSE) features are used [19] to discriminate CN with early AD. A clustering based approach is proposed [131] to group adjacent voxels for classification of CN, MCI and AD. Fisher discriminate ratio (FDR) is used [136] to extract useful voxels as features (VAF) from SPECT images.

Gaussian mixture model (GMM) is used in [150] for CN vs AD. It is stated that the proposed GMM based feature extraction makes the data linearly separable. Ortiz et al. [122] used PET and sMRI data to find the most discriminative features using sparse inverse covariance estimation (SICE) method with SVM. Non-negative matrix factorization (NMF) based features with SVM are found to give better performance than PCA with SVM to classify CN vs AD [124]. Moreover, Abdulkadir et al. [1] illustrated the affects of hardware heterogeneity on classification accuracy of SVM. They also found high confidence level of classification performance for large samples. Cuingnet et al. [32] stated that DARTEL based features are better than SPM features for CN, MCIc, MCInc, and AD. Moreover, it is concluded that feature selection techniques for sMRI images may lead to less classification accuracy due to addition of hyperparameters. Further, Schmitter et al. [147] found volume based features to be more useful than voxel based morphometry (VBM). Morphological features of brain regions are used by Plocharsky et al. [132] to classify CN vs AD, while Long et al. used shape differences in the subjects' brains for classification of CN, AD, sMCI, and pMCI. Fuzzy based classes for hippocampus volume are used by Tangaro et al. [167] for classification of CN vs AD, and MCIc vs MCInc.

Wavelet based features are used in various works. Chaplot et al. [18] used discrete wavelet transform (DWT) features, while Zhang et al. [193] found that 3-D DWT and SVM are useful for classification of CN, MCI and AD subjects. Segovia et al. [149] discovered that partial least squares (PLS) components have a higher FDR score as compared to principal component analysis (PCA) for CN vs AD using SPECT images. Ortiz et al. used self organizing maps (SOMs) [120] for unsupervised segmentation of sMRI images in classification of CN vs AD. However, Chaplot et al. [18] found that SVM performs better than SOM for classification of Alzheimer's patients using T2-weighted images.

Other techniques like SVM-RFE [52, 139] are used as an optimized feature selection technique in [56] to select prominent brain features for CN vs AD. Independent component analysis (ICA) is used in many works [79, 80] for classification of CN vs AD using SVM. EEG data is also used [85] for classification of CN vs AD using SVM. Mazaheri et al. [113] used EEG recordings of word comprehension by subjects to classify MCIc from MCInc and CN. Some researchers also focused on blood based biomarkers for Alzheimers [39, 125]. Gostolya et al. [50] used speech patterns of subjects and classified using linear SVM.

3.3 Kernel function

Different kernels have been used with SVM for classification of Alzheimers. Various researchers have used linear kernel with SVM to classify Alzheimer data as shown in fig. 3(a). This is due to the fact that in linear kernel, there is no kernel parameter to tune. Some researchers also utilized multiple kernels for SVM [4]. Moreover, in most papers, the sample size is also very small as shown in Table 1, which may lead to overfitting of the data with radial basis function (RBF) kernel [139]. This leads to the use of linear kernel due to its simplicity. The usage of different kernels as per our survey is shown in fig. 3(a).

Kloppel et al. [82] used linear SVM to classify pathologically confirmed cases of AD with CN, and suggested that SVM can help in the diagnosis of AD. It has been stated in [79, 136] that linear kernel provides better classification performance for high dimensional data as compared to polynomial or RBF kernel. However, polynomial kernel is also used by researchers. Lahmiri et al. [87] used polynomial kernel for multiclass classification of CN, MCI, and AD. Zhang et al. [191] found that polynomial kernel is useful in classification of CN vs AD using PCA features. In 2018, Lahmiri et al [88] used volumetric features with cognitive test scores for classification of CN vs AD with polynomial kernel.

Some researchers also used an ensemble of kernels. Multiple kernel SVM is used by Alam et al. [4] for classification of CN, MCI, and AD. Kamathe et al. [75] used linear, polynomial and RBF kernel for classification of CN vs AD.

Table 1. Comparison of research on classification of Alzheimer data using SVM.

Sr. No.	Year	Authors	Target	Modality	Feature extraction	Machine learning	Dataset	Validation	Performance		
									Acc (%)	Sens (%)	Spec (%)
1	2005	Stoeckel et al. [158]	CN vs AD	SPECT	VBM+Spatial normalization	CSVM (Linear)	130 (31 CN, 99 AD)	LOOCV	72.5	93	-
2	2006	Chaplot et al. [18]	CN vs AD	sMRI (T2)	DWT	SVM (RBF)	52 (6 CN, 46 AD)	-	98	-	-
3	2007	Fung et al. [45]	CN vs AD	SPECT	Subsampling	CSVM (Linear)	130 (31 CN, 99 AD)	LOOCV	-	84.4	90.9
4	2008	Alvarez et al. [7]	CN vs AD	SPECT	Component based feature extraction	Ensemble SVM	79 (41 CN, 38 AD)	LOOCV	97.5	-	-
				PET			60 (18 CN, 42 AD)		100		
5	2008	Ramirez et al. [136]	CN vs AD	SPECT	FDR	SVM (RBF)	52 (23 CN, 29 AD)	-	90.38	-	-
6	2008	Gonzalez et al. [143]	CN vs AD	SPECT	Component based feature extraction	SVM (Linear)-Classification tree	79 (41 CN, 38 AD)	LOOCV	90	-	-
7	2008	Vemuri et al. [175]	CN vs AD	sMRI (T1)	VBM	SVM (Linear)	380 (190 CN, 190 AD)	4-fold	89.30	-	-
8	2008	Fan et al. [41]	CN vs MCI	sMRI+PET	RAVEN maps+ODC	SVM (Linear)	30 (15 CN, 15 MCI)	LOOCV	100	-	-
9	2008	Kloppel et al. [82]	CN vs AD	sMRI (T1)	VBM	SVM (Linear)	68 (34 CN, 34 AD)	LOOCV	95.6	97.1	94.1
			CN vs mAD				90 (57 CN, 33 AD)		85.6	75.8	91.2
10	2008	Mesrob et al. [115]	CN vs AD	sMRI (T1)	SVM-RFE	SVM (RBF)	61 (28 CN, 33 AD)	LOOCV	90.2	-	-
11	2009	Magnin et al. [109]	CN vs AD	sMRI (T1)	VBM+Histogram	SVM (RBF)	38 (22 CN, 16 AD)	LOOCV	94.5	91.5	96.6
12	2009	Gerardin et al. [48]	CN vs AD	sMRI (T1)	SPHARM	SVM (RBF)	46 (23 CN, 23 AD)	LOOCV	94	96	92
			CN vs MCI						83	83	84
13	2009	Chaves et al. [19]	CN vs AD	SPECT	VBM+NMSE	SVM (Linear)	79 (41 CN, 38 AD)	LOOCV	98.3	-	-
14	2010	Plant et al. [131]	CN vs AD	sMRI (T1)	VBM+Clustering	SVM (Linear)	74 (18 CN, 24 MCI, 32 AD)	LOOCV	90	96.88	77.78
			CN vs MCI						97.62	95.83	100
			MCI vs AD						95.83	88.89	100
15	2010	Segovia et al. [150]	CN vs AD	SPECT	GMM	SVM (Linear)	91 (41 CN, 50 AD)	LOOCV	94.5	-	-
				PET					96.67		
16	2010	Padilla et al. [124]	CN vs AD	SPECT	FDR+NMF	SVM (RBF)	97 (41 CN, 56 AD)	LOOCV	94.9	96.4	92.8
17	2011	Abdulkadir et al. [1]	CN vs AD	sMRI (T1)	VBM	SVM (Linear)	417 (226 CN, 191 AD)	LOOCV	87	-	-
18	2011	Illan et al. [65]	CN vs AD	SPECT	Image factorization	Ensemble SVM (RBF)	79 (41 CN, 38 AD)	LOOCV	96.91	94.64	100
19	2011	Cuingnet et al. [32]	CN vs AD	sMRI (T1)	VBM/ VolBM/ STAND- score	SVM (RBF)	509 (162 CN, 76 MCIc, 134 MCInc, 137 AD)	LOOCV	-	81	95
			CN vs MCIc						-	73	85
			MCInc vs MCIc						-	70	61
20	2012	Doecke et al. [39]	CN vs AD	Blood plasma	Statistical methods	SVM (Linear)	1131 (812 CN, 319 AD)	CV	-	80	85
21	2012	O'Dwyer et al. [119]	CN vs MCI	DTI	TBSS	SVM (RBF)	73 (40 CN, 19 naMCI, 14 aMCI)	10-fold	93	92.8	-
			CN vs aMCI vs naMCI						92.2	93.4	-
22	2013	Segovia et al. [149]	CN vs AD	SPECT	PLS	SVM (Linear)	97 (41 CN, 56 AD)	LOOCV	91.75	92.68	91.07
23	2013	Ramirez et al. [134]	CN vs AD	SPECT	FDR	SVM (RBF)	52 (23 CN, 29 AD)	LMOCV (M=5)	90.38	93.1	86.96
24	2013	Dukart et al. [40]	CN vs AD	FDG-PET+sMRI	VBM	SVM (Linear)	56 (28 CN, 28 AD)	LOOCV	100	100	100
25	2013	Ortiz et al. [120]	CN vs AD	sMRI (T1)	SOM clustering +FDR	SVM (RBF)	50 (25 CN, 25 AD)	LOOCV	92	96	98
26	2013	Haller et al. [53]	md-aMCI vs sd-aMCI	DTI	TBSS	SVM (RBF)	66 (18 sd-aMCI, 13 sd-fMCI, 35 md-aMCI)	10-fold	98.4	-	-
			md-aMCI vs sd-aMCI						97.7		
			sd-fMCI vs sd-aMCI						99.67		
27	2013	Lee et al. [91]	CN vs MCI	DTI	TBSS	SVM (RBF)	84 (39 CN, 45 MCI)	10-fold	100	100	100
28	2014	Hidalgo-Munoz et al. [56]	CN vs AD	sMRI (T1)	VBM +SVM-RFE	SVM (Linear)	370 (185 CN, 185 AD)	10-fold	100	-	-
29	2014	Lahmiri et al. [87]	CN vs MCI	sMRI (T2)	MSA	SVM (Polynomial)	33 (11 CN, 11 MCI, 11 AD)	10-fold	97.08	98.09	96.07
			MCI vs AD						97.5	100	94.93
30	2015	Khedher et al. [80]	CN vs AD	sMRI (T1)	VBM+Fast ICA	SVM (Linear)	818 (229 CN, 401 MCI, 188 AD)	2-fold	87.12	89.92	83.98
			CN vs MCI						77.62	80.27	74.49
			MCI vs AD						85.41	85.59	85.11
31	2015	Schmitter et al. [147]	CN vs AD	sMRI (T1)	VBM+VolBM	SVM (Linear)	818 (229 CN, 401 MCI, 188 AD)	LOOCV	-	86	91
			CN vs MCI							78	68
			MCI vs AD							69	67
			ADc vs ADnc							75	66
32	2015	Zhang et al. [193]	CN vs MCI vs AD	sMRI (T1)	VolBM+3D-DWT+PCA	SVM (RBF)	178 (97 CN, 57 MCI, 24 AD)	5-fold	81.5	-	-
33	2015	Zhang et al. [192]	CN vs AD	sMRI (T1)	DF+PCA	TWSVM	126 (98 CN, 28 AD)	10-fold	92.75	90.56	93.37
34	2015	Xu et al. [183]	CN vs AD	sMRI	Lasso features	S-LSTSVM (RBF)	329 (191 CN, 138 AD)	5-fold	92.1	92.52	92.07
35	2015	Ortiz et al. [122]	CN vs AD	FDG-PET+sMRI	VBM+SICE	SVM (Linear)	249 (68 CN, 111 MCI, 70 AD)	10-fold	92	96	86
			CN vs MCI						86	90	82
			MCI vs AD						84	87	81
36	2015	Retico et al. [139]	CN vs AD	sMRI (T1)	VBM+SVM-RFE	SVM (Linear)	635 (189 CN, 136 MCIc, 166 MCInc, 144 AD)	20-fold	88.9	-	-
			MCIc vs MCInc						70.7	-	-
37	2015	Zhang et al. [191]	CN vs AD	sMRI (T1)	PCA	SVM (Polynomial)	126 (98 CN, 28 AD)	10-fold	92.36	-	-
38	2016	Zhu et al. [200]	MCIc vs MCInc	sMRI	PIS	TS-SVM	151 (70 MCIc, 81 MCInc)	10-fold	82.5	-	-

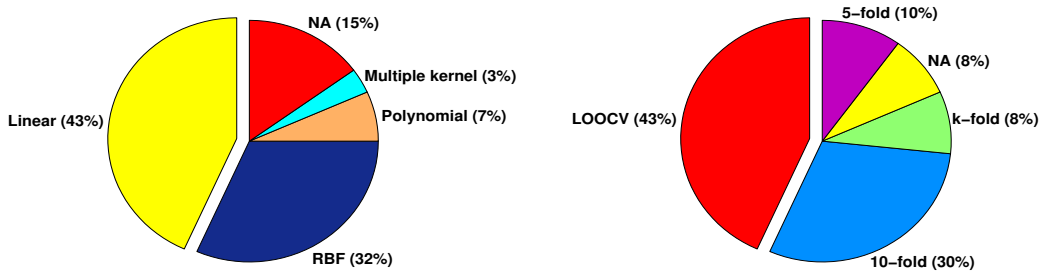
Table 1 (Contd.)

Sr. No.	Year	Authors	Target	Modality	Feature extraction	Machine learning	Dataset	Validation	Performance		
									Acc (%)	Sens (%)	Spec (%)
39	2016	Moller et al. [117]	CN vs AD	sMRI (T1)	VBM	SVM (Linear)	178 (94 CN, 84 AD)	LOOCV	85	83	87
40	2016	Plocharsky et al. [132]	CN vs AD	sMRI (T1)	Morpholgical features (length, area, depth)	SVM (Linear)	210 (100 CN, 110 AD)	10-fold	87.9	90	86.7
41	2017	Alam et al. [4]	CN vs AD	sMRI (T1)	VolBM+KPCA	SVM (Multiple kernel)	293 (102 CN, 102 MCI, 89 AD)	10-fold	93.85	92.1	94.45
			CN vs MCI						86.54	84.85	87.74
			MCI vs AD						75.12	73.92	77.24
42	2017	Khedher et al. [79]	CN vs AD	sMRI (T1)	ICA	SVM (RBF)	818 (229 CN, 401 MCI, 188 AD)	k-fold	89	92	86
			CN vs MCI						79	82	76
			MCI vs AD						85	85	86
43	2017	Beheshti et al. [11]	CN vs AD	sMRI (T1)	VBM+GA	SVM (Linear)	458 (162 CN, 65 sMCI, 71 pMCI, 160 AD)	10-fold	93.01	89.13	96.8
			pMCI vs sMCI						75	76.92	73.23
44	2017	Long et al. [104]	CN vs AD	sMRI (T1)	MDS+PCA	SVM (Linear)	427 (135 CN, 132 sMCI, 95 pMCI, 65 AD)	10-fold	96.5	93.85	97.78
			CN vs pMCI						97.1	87.37	94.82
			sMCI vs pMCI						88.99	86.32	90.91
45	2017	Tangaro et al. [167]	CN vs AD	sMRI (T1)	VolBM	SVM (Linear)	372 (117 CN, 86 MCIc, 71 MCInc, 98 AD)	10-fold	100	-	-
			MCIc vs MCInc						83.4	-	-
46	2017	Lu et al. [108]	CN vs MCI	FDG-PET	VBM	RF-RSVM	272 (152 CN, 120 MCI)	3-fold	90.53	90.63	93.33
47	2017	Alam et al. [5]	CN vs AD	sMRI	DTCWT/LDA	TWSVM	237 (130 CN, 137 AD)	10-fold	96.88	97.72	95.61
48	2017	Hojjati et al. [58]	MCIc-MCInc	rs-fMRI	PCC+F-score	SVM (Linear)	80 (18 MCIc, 62 MCInc)	9-fold	91.4	83.24	90.1
49	2017	Kulkarni et al. [85]	CN vs AD	EEG	ICA/ Wavelet/ Spectral	SVM	100 (50 CN, 50 AD)	LOOCV	96	-	-
50	2018	Sun et al. [166]	CN vs AD	sMRI (T1)	VBM+PCC	Group lasso SVM	509 (162 CN, 134 sMCI, 76 pMCI, 137 AD)	5-fold	95.1	93.8	83.8
			CN vs MCI						70.8	72.1	69.1
			sMCI vs pMCI						65.4	67.6	64.2
			MCI vs AD						65.7	63.2	67.3
51	2018	Zeng et al. [187]	CN vs AD	sMRI (T1)	PCA+PSO and SDPSO	SVM (RBF)	361 (92 CN, 82 sMCI, 95 pMCI, 92 AD)	10-fold	82.5	-	-
			sMCI vs pMCI						69.23		
			CN vs sMCI						76.92		
			CN vs pMCI						85.71		
			sMCI vs AD						72.94		
			pMCI vs AD						57.14		
52	2018	Lahmiri et al. [88]	CN vs AD	sMRI (T1)	VolBM	SVM (Polynomial)	70 (35 CN, 35 AD)	10-fold	100	100	100
53	2018	Kamathe et al. [75]	CN vs AD	sMRI (T1, T2)+PD	ICA	SVM (Polynomial)	20 (15 CN, 5 AD)	-	100	-	-
54	2018	Bi et al. [14]	CN vs AD	rs-fMRI	PCC	RSVM (RBF)	61 (36 CN, 25 AD)	-	94.44	-	-
55	2018	Mazaheri et al. [113]	MCIc vs (CN + MCInc)	EEG	TFRs	SVM (RBF)	36 (11 CN, 10 MCInc, 15 MCIc)	LOOCV	-	80	95
56	2018	Paraskevaidi et al. [125]	CN vs AD	Blood plasma	PCA-LDA	SVM	41 (15 CN, 11 eAD, 15 lAD)	LOOCV	-	84	86
			CN vs lAD							84	77
			eAD vs lAD							66	83
57	2018	Zhang et al. [194]	CN vs AD	DTI	LDH +SVM-RFE	SVM (Linear)	213 (51 CN, 75 eMCI, 39 lMCI 48 AD)	LOOCV	89.9	-	-
			CN vs eMCI						88.1		
			CN vs lMCI						100		
			eMCI vs lMCI,						92.98		
			eMCI vs AD						84.55		
			lMCI vs AD						97.7		
58	2019	Peng et al. [129]	CN vs AD	sMRI+PET+SNP	Volume + Mean intensity features	SVM (Multiple kernel)	189 (47 CN, 93 MCI, 49 AD)	10-fold	96.1	97.3	94.9
			CN vs MCI						80.3	85.6	69.8
			MCI vs AD						76.9	65.9	82.7
59	2019	Sheng et al. [152]	CN vs eMCI	fMRI	RF-score	SVM	96 (24 CN, 24 eMCI, 24 lMCI, 24 AD)	5-fold	93.8	-	-
			CN vs LMCI						95.8		
			CN vs AD						95.8		
			eMCI vs lMCI						87.5		
			lMCI vs AD						91.7		
60	2019	Gosztolya et al. [50]	CN vs MCI	Acoustic signal	MFCC	SVM (Linear)	75 (25 CN, 25 MCI, 25 mAD)	5-fold	80	-	75.9
			CN vs mAD						86	-	87.5
			MCI vs mAD						80	-	85.7

Abbreviations: **LMOCV**- Leave M out cross validation, **CV**- Cross validated, **CT**- Classification tree, **ODC**- Optimally differentiating clusters, **SPHARM**-Spherical harmonics coefficients, **VolBM**- Volume based morphometry, **eMCI**- Early MCI, **lMCI**- Late MCI, **mAD**- Mild AD, **ADc**- AD converter, **ADnc**-AD non-converter, **MDS**- Multi dimensional scaling, **PCC**- Pearson correlation coefficient, **SNP**- Single nucleotide polymorphisms, **MFCC**- Mel frequency cepstral coefficients, **DF**- Displacement field, **MSA**- Multiscale analysis, **S-LSTSVM**-Structural least squares twin support vector machine, **RSVM**-Random SVM, **rs-fMRI**- Resting state fMRI, **PIS**- Partial image sequence, **TS-SVM**- Temporally structured SVM, **RF-score**- Relief feature score, **STAND**-score- Structural abnormality index score, **sd-aMCI**- Single domain amnestic MCI, **md-aMCI**- Multiple domain amnestic MCI, **sd-fMCI**- Single domain frontal MCI, **TBSS**- Tract-based spatial statistics, **TFR**- Time frequency representation, **KPCA**- Kernel PCA, **LDH**- Local diffusion homogeneity.

Peng et al. [129] used MRI and genetic data for features, and used multiple kernel learning with SVM to classify the subjects. The selection of optimal hyperparameters is a major step in the classification of SVM. Among the various

methods, leave one out cross validation (LOOCV) has been widely used for classification of Alzheimer using SVM. The details of the cross-validation methods are shown in fig. 3(b).



(a) (b)

Fig. 3. (a) Plot showing usage of different types of kernels and (b) cross validation methods used with SVM for Alzheimers. NA means information about kernel is not available, and $k = 2, 3, 4, 9$ and 20.

3.4 Variants of SVM used for Alzheimers

Many variants of SVM are developed for different types of classification problems. Standard SVM do not use spatial information of brain image in the optimization problem [33, 158]. To provide spatial information, contiguous SVM (CSVM) is used to classify SPECT images of Alzheimers and control subjects [45, 158]. CSVM uses the information about voxel connectivity to give a more robust classifier. For reducing the computation cost, Zhang et al. [192] used twin support vector machine (TWSVM) for classification of CN vs AD, while structural least squares twin support vector machine (S-LSTSVM) is used in [183]. For optimized feature selection, Beheshti et al. [11] used genetic algorithm (GA) with linear SVM for classification of CN vs AD and pMCI vs sMCI.

For early diagnosis of Alzheimers, Zhu et al. [200] used a temporally structured SVM (TS-SVM) for classification of longitudinal MR images of MCI converters and non-converters. Lu et al. [108] proposed a random forest robust SVM (RF-RSVM) for classification of CN vs MCI using FDG-PET images. TWSVM is used for classification of CN vs AD [5] using dual-tree complex wavelet transform (DTCWT), LDA and PCA features. Sun et al. [166] introduced spatial anatomical regularization with SVM for classification of CN, AD, sMCI, and pMCI. To optimize the SVM parameters, Zeng et al. [187] proposed a switching delayed particle swarm optimization SVM (SDPSO-SVM). In order to reduce the complexity of SVM, Bi et al. [14] used random support vector machine clusters for classification of CN vs AD using rs-fMRI. In RSVM clusters approach, samples and features randomly are chosen randomly from the dataset and trained accordingly. It helps to reduce the size of training data leading to less computational complexity. Some researcher also applied ensemble based SVM for better prediction accuracy in Alzheimer's disease (AD) classification [7, 65, 143].

The usage of different cross validation strategies is shown in fig. 3(b). LOOCV comes out to be the frequently used method. This may be attributed to small sample size in the works shown in Table 1.

4 ARTIFICIAL NEURAL NETWORKS, TRANSFER LEARNING & MULTI-KERNEL LEARNING

Artificial neural networks (ANN) shown in fig. 4, are widely used for machine learning models that can model highly nonlinear patterns of data. In this section, we review the approaches using ANNs and other approaches like

Table 2. Comparison of recent studies using artificial neural network, transfer learning, multi-task learning and connectivity features.

Sr. No.	Year	Authors	Target	Modality	Feature extraction	Machine learning	Dataset	Validation	Performance		
									Acc (%)	Sens (%)	Spec (%)
1	2008	Huang et al. [63]	CN vs AD	sMRI (T1)	VBM	ANN	22 (10 AD, 12 CN)	-	100	-	-
2	2009	Savio et al. [146]	CN vs AD	sMRI (T1)	VBM	LVQNN	98 (49 AD, 49 CN)	10-Fold	83	-	-
3	2010	Ahmadlou et al. [2]	CN vs AD	EEG	WT + VG	PCA+RBFNN	27 (7 CN, 20 AD)	-	97.7 ± 3.2	100	91.08
4	2010	Lopes et al. [34]	CN vs AD	EEG	-	PANN	67 (33 AD, 34 CN)	-	-	82	61
5	2010	Yang et al. [185]	CN vs AD	sMRI (T1)	VBM	PCA+BPNN	52 (24 probable AD, 28 CN)	LOOCV	92.17	79.91	88.61
6	2010	Long et al. [105]	CN vs AD	sMRI (T1)	Atlas	Quick shift clustering	75 (35 very mild-to-mild dementia, 40 CN)	-	94.67 – 97.33	-	-
7	2010	Joshi et al. [74]	AD vs PD	Health records	Feature ranking	Decision Tree Bagging BF Tree RF Tree RBF NN MLP ANN	890 (578 AD, 312 PD)	-	97.3415 98.5685 99.182 99.591 97.137 98.9775 99.591	-	-
8	2011	Sankari et al. [144]	CN vs AD	EEG	CC; WC	PNN	27 (7 CN, 20 probable AD)	-	100	-	-
9	2011	Zhang et al. [188]	CN vs AD	sMRI (T1)+PET+CSF	Atlas	MlapRls	202 (52 CN, 99 MCI, 51 AD)	10-Fold	94	94	94
10	2011	Rodrigues et al. [142]	CN vs AD	EEG	WT; STFT	ANN	34 (14 CN, 20 probable AD)	LOOCV	91.5	92.1	90.8
11	2011	Quintana et al. [133]	CN vs AD CN vs MCI CN vs MCI vs AD	NPR	NPF, Age & Education	ANN	522 (346 CN, 79 MCI, 97 AD)	-	100 98.33 66.67	-	-
12	2012	Mahanandet al. [110]	CN vs AD	sMRI (T1)	VBM	ICGA+SRAN	60 (30 CN, 30 mild-to-mod AD)	10-Fold	92.13 ± 2.71	-	-
13	2013	Chyzyk et al. [27]	CN vs AD	sMRI (T1)	VBM	GA + ELM	98 (49 CN, 49 very mild-to-mild AD)	10-Fold	86	87	90
14	2013	Jie et al. [72]	CN vs AD CN vs MCI MCIc vs MCInc	sMRI (T1) + PET	Atlas	M2TFS+MK-SVM	202 (52 CN, 43 MCIc, 56 MCInc, 51 AD)	10-Fold	95.03 79.27 68.94	94.9 85.86 64.65	95 66.54 71.79
15	2013	Liu et al. [97]	CN vs AD	sMRI (T1)	Atlas	SMSE	331 (77 CN, 169 MCI, 85 AD)	5-Fold	75.89	-	-
16	2013	Mahmood et al. [111]	CN vs AD	sMRI (T1)	Atlas	PCA+ANN	457 (357 CN, 100 AD)	-	89.22	-	-
17	2013	Naami et al. [3]	CN vs AD	sMRI (T2)	-	ANN	37 (17 AD, 20 CN)	-	100	-	-
18	2014	Jie et al. [73]	CN vs aMCI	fMRI	Atlas (NBF)	SVM	37 (25 CN, 12 aMCI)	LOOCV	91.9	100	88
19	2014	Min et al. [116]	CN vs AD pMCI vs sMCI	sMRI (T1)	VBM	SVM	459 (128 CN, 117 pMCI, 117 sMCI, 97 AD)	10-Fold	91.64 72.41	88.56 72.12	93.85 72.58
20	2014	Ortiz et al. [121]	CN vs AD MCI vs AD	sMRI (T1)	VBM	SOM	818 (229 CN, 401 MCI, 188 AD)	10-Fold	90 ± 6 83 ± 6	87 ± 7 82 ± 7	92 ± 9 87 ± 7
21	2014	Suk et al. [163]	CN vs AD MCI vs AD CN vs MCI MCIc vs MCInc	sMRI (T1) + PET +CSF	Atlas	SMTL+MK-SVM	202 (52 CN, 43MCIc, 56 MCInc, 51 AD)	10-Fold	95.27 ± 6.58 74.60 ± 9.57 80.07 ± 8.42 72.02 ± 13.80	94 46.67 86.78 58	96.33 89 67.33 82.67
22	2015	Wang et al. [179]	CN vs AD	sMRI (T2)	DWT	ANN	481 (73 CN, 408 AD)	5-Fold	100	100	100
23	2015	Cheng et al. [23]	MCIc vs MCInc	sMRI (T1) + PET + CSF	Atlas	M2TL	202 (52 CN, 43 MCIc, 56 MCInc, 51 AD)	10-Fold	80.1	85.3	73.3
24	2015	Gorji et al. [49]	CN vs AD MCI vs AD CN vs MCI	sMRI (T1)	Pseudo Zernike Moment	PRNN; LVQNN	500 (148 CN, 172 MCI, 180 AD)	10-Fold	97.27 94.88 95.59	96.64 94.18 95.89	97.79 95.55 95.34
25	2016	Khazaee et al. [78]	CN vs MCI & AD MCI vs CN & AD AD vs CN & MCI	rs-fMRI	Atlas (NBF)	SVM	168 (45 CN, 89 MCI, 34 AD)	Hold-out	87.29 72.03 97.46	-	-
26	2016	Suk et al. [160]	CN vs AD CN vs MCI pMCI vs sMCI CN vs MCI vs AD CN vs sMCI vs pMCI vs AD	sMRI (T1) + PET + CSF	Atlas	DW-S^2MTL + SVM(Linear)	202 (52 CN, 43 pMCI, 56 sMCI, 51 AD)	10-Fold	95.09 ± 2.28 80.11 ± 2.64 74.15 ± 3.35 55 53.72	92 93.89 50.5 - -	98 53.67 92.67 - -
27	2016	Tong et al. [172]	CN vs AD CN vs MCI CN vs MCI vs AD	sMRI (T1) +FDG-PET+CSF +Genetics	Atlas	NGF+SVM	147 (35 CN, 37 AD, 75 MCI)	Leave-p-out	91.8 79.5 60.26	88.9 85.1 -	94.7 67.1 -
28	2016	Aljovic et al. [6]	CN vs AD	Biomarkers	-	ANN		-	-	95.5	91.43
29	2017	Liu et al. [99]	CN vs AD MCI vs AD CN vs MCI MCIc vs MCInc	sMRI (T1)	Atlas	MKBoost+SVM	710(230 CN, 120 MCIc, 160 MCInc, 200 AD)	10-Fold	94.65 89.63 85.79 72.08	95.03 91.55 88.91 75.11	91.76 86.25 80.34 71.05
30	2017	Hon et al. [59]	CN vs AD	sMRI (T1)		CNN(TL)	200 (100 CN, 100 AD)	5-Fold	96.25	-	-
31	2017	Lama et al. [89]	CN vs AD CN vs MCI vs AD	sMRI (T1)	Cortical	PCA+RELM	214 (70 CN, 74 MCI , 70 AD)	10-Fold	77.30 61.58	62.12 54	79.85 62.25
32	2017	Jha et al. [70]	CN vs AD	sMRI (T1)	VBM	PCA + FFNN	126 (98 CN, 28 AD)	10-Fold	90.06 ± 0.01	92.0 ± 0.04	87.78 ± 0.04
33	2017	Zheng et al. [197]	CN vs AD	sMRI (T1) + PET	Atlas	(SVM+) + (RBM+)	103 (52 CN, 51 AD)	10-Fold	88.52 ± 8.61	84.60 ± 15.24	92.20 ± 12.61
34	2017	Beheshti et al. [12]	CN vs AD CN vs MCI pMCI vs sMCI AD vs sMCI AD vs pMCI	sMRI (T1)	VBM (NBF)	SVM	231 (61 CN, 42 sMCI, 45 pMCI, 83 AD)	10-Fold	84.17 70.38 61.05 67.59 62.84	88.83 78.17 52.65 79.25 76.38	79.00 60.22 70.52 45.47 39.57
35	2017	Cheng et al. [22]	CN vs AD CN vs MCI pMCI vs sMCI	sMRI (T1)	Atlas	MDTFS+MDTC	807 (226 CN, 167 pMCI, 228 sMCI, 186 AD)	10-Fold	94.7 81.5 73.8	94.1 85.8 69	94.8 73.3 77.4
36	2018	Liu et al. [98]	CN vs AD MCI vs AD CN vs MCI MCIc vs MCInc	sMRI (T1)	Atlas	MKBoost+ SVM	680 (200 CN, 120 MCIc, 160 MCInc, 200 AD)	10-Fold	95.37 90.41 86.56 73.95	92.49 92.83 90.74 76.13	96.08 88.82 84.83 72.24
37	2018	Kim et al. [81]	CN vs AD CN vs MCI	sMRI (T1) +PET+CSF	Atlas + CSF	MSH-ELM	202 (52 CN, 99 MCI, 51 AD)	10-Fold	97.2 87.01	98.08 75	94.12 91.92
38	2018	Bi et al. [13]	CN vs AD	rs-fMRI	Atlas	Random NN cluster	60 (25 AD, 35 CN)	-	92.3	-	-
39	2018	Li et al. [94]	CN vs AD	rs-fMRI	VBM	Subspace alignment	318 (129 AD, 189 CN)	LOOCV	84.6	92	79

Table 2 (Contd.)

Sr. No.	Year	Authors	Target	Modality	Feature extraction	Machine learning	Dataset	Validation	Performance		
									Acc (%)	Sens (%)	Spec (%)
40	2018	Cheng et al. [24]	CN vs AD	sMRI (T1)	VBM	rMLTFL	406 (112 CN, 86 pMCI, 106 sMCI, 102 AD)	10-Fold	95.2	95.2	95.3
			CN vs MCI						82.4	86.7	73.8
			pMCI vs sMCI						76.3	73.4	78.6
			MCI vs AD						76.7	61.4	81.8
41	2018	Zhou et al. [198]	CN vs AD	sMRI (T1)	Atlas	TrAdaBoost	543 (103 CN, 245 MCI, 195 AD)	-	93.75	87.5	100
42	2018	Veen et al. [173]	CN vs PD	FDG-PET	VBM	LGMLVQ	304 (82 CN, 146 PD, 76 AD)	-	-	79.23	68.15
			CN vs AD							91.47	91.45
			PD vs AD							84.7	86.63
43	2018	Cui et al. [31]	CN vs AD	sMRI (T1)	VBM	ANN+BGRU	427 (128 AD, 229 CN)	5-Fold	89.69	86.87	92.58
44	2019	Kar et al. [76]	CN vs AD	sMRI (T1), sMRI (T2), CT	ROI based	ANN	20 (9 AD, 11 CN)	-	100	100	100
45	2019	Fritsch et al. [42]	CN vs AD	Linguistic data	n-gram	LSTM	266 (168 AD, 98 CN)	LOOCV	85.6	-	-

Abbreviations: **SMSE**- Multi-view multimodal spectral embedding, **RNN**- Recurrent neural network, **PRNN**- Pattern recognition neural network, **LGMLVQ**- Generalized matrix learning vector quantization with local relevance matrix, **ERP**- Event related potential, **KNN**- K - nearest neighbours, **ICGA**- Integer coded genetic algorithm, **M2TL**- Multimodal manifold regularized transfer learning, **M2TFS**- Manifold regularized multi-task feature selection, **RELM**- Regularized extreme learning machine **MDTFS**- Multi-domain transfer feature selection, **MDTC**- Multi-domain transfer classification, **MKBoost**- Multi-kernel boosting, **MSH-ELM**- Multimodal sparse hierarchical extreme learning machine, **DBM**- Deformation based morphometry, **DW-S^{2}MTL**- Deep weighted subclass based sparse multi-task learning, **CT**- Computed tomography, **NPR**- Neurophysiological records, **NPF**- Neurophysiological features, **CC**- Conventional coherence, **WC**- Wavelet coherenece, **EEG**- Electroencephalogram, **NGF**- Nonlinear graph fusion, **ePANN**- Paraconsistent artificial neural network, **RF**- Random forest, **LC**- Lattice computing. **Atlas**- Atlas based parcellation of brain regions, **n-gram**- n-gram language model, **LSTM**- Long short term memory.

multi-task learning (MTL), transfer learning (TL) and multi-kernel learning (MKL) for AD classification. We also review certain feature selection (FS) techniques like network connectivity features, multi-view features and other machine learning approaches used for AD classification. The works using ANN for Alzheimers are presented in Table 2.

4.1 Transfer learning

Conventional machine learning models utilize samples from only a single domain which greatly affects their performance when the number of samples available is very low. Transfer learning is an approach that not only utilizes samples from the target domain but also from various auxiliary (related) domains. Cheng et al. [23] proposed a multimodal manifold regularized transfer learning (M2TL) to transfer knowledge learned from an auxiliary domain (CN vs AD) to aid in learning the target domain (MCIc vs MCInc).

In their subsequent work, Cheng et al. [22] went a stride ahead and proposed a novel multi-domain transfer learning model (MDTL). To deal with the small sample size problem in training deep neural networks, Hon et al. [59] used deep transfer learning by fine-tuning two popular pre-trained networks - visual geometry group 16 (VGG16) and Inception on the target task of classifying AD and CN subjects. Zhou et al. [198] used the TrAdaBoost algorithm to transfer knowledge learned from ADNI to AD samples from a local hospital. Transferring knowledge from all source domains without ignoring the unrelated domains degrades the performance of the model. Also, the class labels assigned to various samples might be erroneous. To deal with these problems, Cheng et al. [24] proposed a robust multi-label transfer learning (rMLTFL) approach that transformed original labels to multi-bit label coding vector, and simultaneously acquired common features to identify unrelated domains. Li et al. [94] also transferred knowledge gained from ADNI samples to the samples acquired locally through the subspace alignment algorithm.

4.2 ANN based machine learning models

Different variants of ANN were used on EEG signals of CN and AD subjects by [34, 144]. Savio et al. [146] compared different variants of ANN like radial basis function neural network (RBFNN), probabilistic neural network (PNN), learning

vector quantization neural network (LVQNN), and found that LVQNN performed the best. An unsupervised approach is employed by Long et al. [105] for classification of CN and AD. Joshi et al. [74] compared various classification algorithms like RBFNN, random forest (RF), best-first decision tree (BF tree), decision tree, bagging, multilayer perceptron (MLP) and ANN for classification of AD and Parkinson's disease (PD) subjects. In real world scenario, all the neuroimaging data required to train a model might not be available at hand. In these scenarios, the batch training scheme will prove ineffective. To this end, Mahanand et al. [110] used a self-adaptive resource allocation network (SRAN) classifier for AD classification. To deal with high dimensional data, Mahmood et al. [111] used PCA for dimension reduction and used ANN for classification. Naami et al. [3] also used ANN to classify CN from AD. To exploit the intrinsic correlation among features and obtain robust performance, Veen et al. [173] compared the performance of generalized matrix learning vector quantization (GMLVQ) with local and global relevance with SVM for classification of Parkinson's disease (PD) and AD subjects. To utilize longitudinal information from MRI scans, Cui et al. [31] cascaded an MLP and bidirectional gated recurrent unit (BGRU) together. Fritsch et al. [42] used the long short-term memory network to utilize information from sequential auditory data for classification of AD subjects. Kar et al. [76] used a fuzzy approach with ANN for discriminating CN and AD subjects using DTI images.

4.3 Feature selection techniques

Many techniques are proposed for better feature selection (FS) from neuroimaging data. Ahmadlou et al. [2] stated that non-linear features may reveal notable differences in certain EEG sub-bands but may not aid in separating the two groups (CN and AD) in band-limited EEG. To address this, visibility graphs (VG) were used for extracting features from EEG signals. Rodrigues et al. [142] used short time fourier transform (STFT) and WT features for classification using ANN. A characteristic of AD is the loss in GM and WM tissues which results in shrinkage of various regions and WM tracts thereby cause loss of GM and WM volume. This process is also accompanied by increase in size of ventricles and CSF volume. To utilize this information, Yang et al. [185] proposed a classification framework that combines GM, WM and CSF volumetric features with ventricular 2D and 3D shape features.

A different approach for extracting features from MR images is to extract cortical features using surface based morphometry. Cho et al. [26] proposed a robust incremental classification method using cortical thickness data, The data was transformed into the spatial frequency domain through a manifold harmonic transform. Along with MRI scans, various other types of data like neuro-psychological and physiological tests, genetic makeup and demographics can be analysed to get further information about the condition of a subject. Quintana et al. [133] leveraged data from neuropsychological tests, age and education for classifying MCI and AD using ANN. Chyzyk et al. [27] proposed a novel wrapper feature selection technique that combines GA with extreme learning machine (ELM). In their subsequent work, Yang et al. [184] used the combination of volumetric and shape features. To address the issues of prior knowledge in manual ROI selection and high dimensionality in whole brain multivariate analysis, Ortiz et al. [121] proposed a novel method for adaptively selecting important ROIs using SOM. Wang et al. [179] used a hybrid PSO with the artificial bee colony (ABC) optimization algorithm along with a feed forward neural network (FFNN). A popular approach to deal with the problem of high dimensionality in whole brain analysis is to extract features from specific ROIs of the brain. However, it comes with a disadvantage of requiring some prior knowledge. Gorji et al. [49] leveraged image moments to extract features from MR images. Jha et al. [70] proposed a dual-tree complex wavelet transform (DTCWT) for extracting features from MR images. It is stated that the property of translation invariance and directional selectivity is a better choice than DWT. Different classification tasks can also be jointly considered together for feature selection

as done by [72, 101, 163, 164]. The brain is an intrinsic complex network and to exploit this information many works have also focused on using network theory for AD classification [12, 73, 78, 98, 99].

4.4 Multiple kernel learning & multimodal data

Data from multiple modalities provides complementary information which can boost the performance of machine learning algorithms. To this end, Ye et al. [186] proposed a multiple kernel learning (MKL) framework that integrated ROI and tensor features into a common feature space. Zhang et al. [188] proposed a novel multimodal Laplacian regularized least squares (mLapRLS) model to utilize unlabelled samples in aiding the classification. Liu et al. [97] argued that selecting features separately from each modality ignores the strong inter-modal correlation within each subject which results in sub-optimal performance. To overcome this problem, a novel multi-task learning based feature selection approach was proposed that jointly selected sparse features from all the modalities. The MKL methods discussed till now combine multiple kernels from different modalities linearly, and is very sensitive to weights assigned to each modality. To address these problems, Tong et al. [172] proposed to combine data from multiple modalities non-linearly using non-linear graph fusion (NGF) technique. Zheng et al. [197] used the learning using privileged information (LUPI) approach during the training phase. A novel restricted Boltzmann machine (RBM+) was proposed for leveraging the privileged information. An ensemble of RBM+ and a LUPI variant of SVM (SVM+) were also used for the classification. In order to obtain high level features from different modalities and fuse them for classification, Kim et al. [81] proposed a multimodal sparse hierarchical extreme learning machine (MSH – ELM).

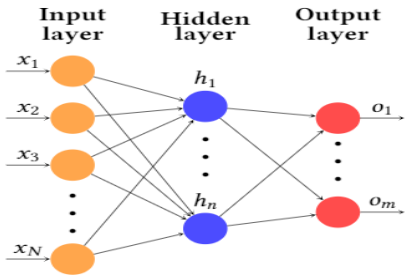


Fig. 4. Architecture of neural network.

5 DEEP LEARNING AND ENSEMBLE METHODS

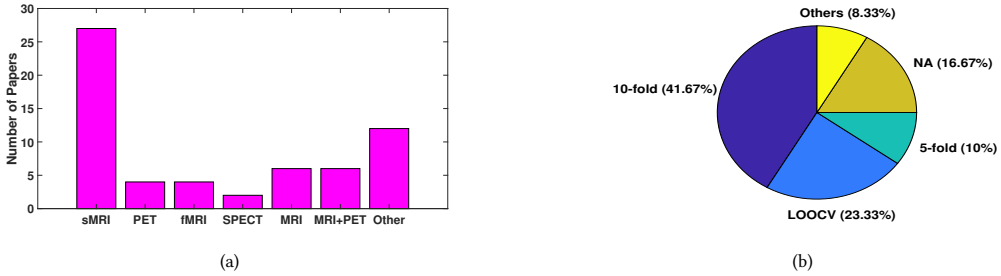
Deep learning (DL) is a multi-layered neural network shown in fig. 4, capable of learning complex structures of data to achieve high degree of abstraction. In DL, there are two kinds of models. One is known as feed forward networks (FFNs) where information flows from input to output, and other is known as recurrent networks (RN) with information from past inputs affecting the present input through feedback connections. The ensemble methods includes techniques other than SVM, ANN, and DL for classification of dementia or AD. Fig. 5(a) shows the different types of modalities used by researchers in the classification or analysis of AD. Fig. 5(b) shows the types of cross-validation for DL and ensemble methods used by researchers.

In the upcoming subsections we discuss DL for classification of AD. We include the simple deep neural network (DNN) architecture, more complex networks such as stacked autoencoders (SAEs), and deep belief networks (DBN). We

also include the popular convolutional neural networks (CNN). Finally, we consider the ensemble algorithms used for classification of AD. Table 3 shows the review of papers using DL and ensemble methods for classification of AD.

5.1 Review of DL studies on Alzheimer's disease

Suk et al. [162] applied SAE and showed that non-linear correlations within the features can improve diagnosis accuracy of AD, MCI, and MCIc. Suk et al. [159] used patch based features from MRI, and PET for classification of AD and MCI from CN by deep Boltzmann machine (DBM). Payan et al. [127] combined SAE, and CNN, and found that 3D-CNN performs better in classification of AD, MCI from CN. Suk et al. [160] used DL-based latent feature representation with a SAE based classifier for classification of AD, MCI, from CN, and MCIc from MCInc. Best accuracies were obtained by multi-kernel SVM (MK-SVM) with low-level features (LLF) and SAE-learned features (SAEF). Liu et al. [102] used SAE based DL architecture for diagnosis of AD in four stages by obtaining high level features and soft-max logistic regressor (LR). Hosseini et al. [62] proposed a deeply supervised adaptive 3D-CNN (DSA-3DCNN) to classify AD and MCI from CN, which was pre-trained by 3D convolutional autoencoder (3D-CAE) using features from MRI. Ortiz et al. [123] extracted 3D patches defined by automated anatomical labeling (AAL) atlas and trained DBNs. The best classification accuracy is obtained by extracting features using DBN (FEDBN-SVM) architecture.



(a) (b)

Fig. 5. (a) Plot showing usage of different types of image modalities, and (b) cross validation methods with DL and ensemble methods. NA means information not available.

Suk et al. [164] combined sparse regression models with a deep CNN known as deep ensemble sparse regression network (DeepESRNet), and were able to clinically diagnose AD. Suk et al. [165] combined DL and state-space modelling and functional dynamics with time of rs-fMRI to form a biomarker for the diagnosis of MCI. Sarraf et al. [145] are the first to use fMRI data with DL. Suk et al. [161] combined two conceptually different models of sparse regression multi-output linear regression (MOLR), joint linear and LR (JLLR) and deep CNN. Better performance in classification of AD is achieved with JLLR-Deep ensemble spare regression network (JLLR-DeepESRNet).

Shi et al. [153] introduced a multimodal stacked deep polynomial networks (MM-SDPN) algorithm consisting of two-stage SDPNs which fuses and learns from multimodal neuroimaging data i.e MRI and PET. It is stated that MM-SDPN based algorithms are better than SAE based algorithms. Basaia et al. [9] used 3D sMRI (T1) and build a DL algorithm that predicts MCI who will convert to AD in 36 months. It is showed that the algorithm is not dependent on dataset, and their CNN based architecture is able to differentiate AD, MCI patients from CN. Leracitano et al. [64] used DL method based on power spectral density (PSD) of brain states using EEG. Patients affected by AD, MCI are differentiated from CN by using PSD spectrograms of EEG. Lu et al. [107] proposed DL multiscale DNN (MDNN) using measures from a single modality FDG-PET and showed that discriminative ability of DL architecture are better for diagnosis AD in case

of single modality. Ensemble of multiple classifiers are more robust, stable, and improve classification performance using different validation settings.

Li et al. [92] used multiple cluster dense CNNs (DenseNets) in which 3D patches were extracted from sMRI (T1) and clustered using K-Means clustering technique. This technique does not require segmentation and rigid registration of MRI for the classification and diagnosis of AD by DenseNets. Jain et al. [68] used transfer learning approach in which first sMRI (T1) slices are passed through a pretrained ImageNet VGG16 network for extraction of features. Spasov et al. [156] used deep CNN for extracting the descriptive features from sMRI (T1) based on 3D separable and grouped convolutions. The parameters are selected efficiently using CNN for classification of AD. Wang et al. [177] developed an ensemble of 3D densely connected CNNs (3D-DenseNets), which maximizes the information flow from one layer to next layer. It avoids over-fitting and enhances the performance for diagnosis of AD and MCI using sMRI images.

5.2 Review of ensemble and other methods for AD classification

Fan et al. [41] used features from MRI and PET and applied high-dimensional pattern classification (PC) to classify AD from MCI. Functional connectivity from rs-fMRI is used by Wang et al. [178] to discriminate AD from CN using LR. An increase in positive correlation is observed between parietal, prefrontal, and occipital lobe while decrease in parietal and prefrontal lobes. Davatzikos et al. [37] used voxel based analysis for the classification of AD and frontotemporal dementia (FTD) from CN using high-dimensional PC method. Better diagnostic accuracy is obtained by high dimensional multi-variate discriminant analysis than conventional measurement. Hinrichs et al. [57] used linear program (LP) for classification of AD using sMRI (T1) and FDG-PET with regularization using spatial smoothness. Lopez et al. [106] used PCA for dimensionsion reduction, and used Bayesian classifiers for CN vs AD. Horn et al. [61] used ROI based features which were reduced by partial least squares (PLS) regression, and yield best accuracy using KNN for CN vs AD.

Chincarini et al. [25] used RF classifier to select informative features from intensity and textural based MRI features. In order to build more robust classifier, Termenon et al. [169] used relevance vector machine (RVM) to build the discriminant function, and applied two stage sequential ensemble of classifiers. Employing an RF classifier on all the available features can degrade the classifier performance due to overfitting. To address this, Li et al. [93] combined weak hierarchical lasso feature selection (wHLFS) with RF for MCIc vs MCInc classification. wHLFS selects only the most important subset of features improving the classification accuracy of RF. To exploit the complementary information provided by multiple modalities, Hor et al. [60] introduced tree-based feature transforms of MRI and PET scans and used RF for classification. Dominguez et al. [54] compared the performance of RF and SVM to classify AD and MCI from CN, on linguistic and phonetic metrics. RF also performed remarkably on features from combination of different modalities like MRI, PET, genetic data and CSF measures [28, 51, 90].

Techniques like LR, LDA, PLS and their variants are also used for diagnosis of AD. In [38, 114], LR was used on features from sMRI. Rao et al. [137] found better performance with sparse LR (SLR) and regularized sparse LR (SRSLR) on voxel wise GM volumes as compared to MLD, and penalized LR (PLR) models for CN vs AD. Liu et al. [100] found that sparse representation-based classifier (SRC) performed better than SVM with high dimensional features. Techniques based on sparse feature representations are also used in [71, 195]. Liu et al. [103] introduced locally linear embedding (LLE) in combination with RLR that leveraged information from patches of local brain regions. It resulted in significant improvement in classification accuracy. Casanova et al. [17] used RLR for CN vs AD, and found that AD pattern similarity (AD-PS) scores are powerful parameters to measure conversion of MCI to AD. Chen et al. [20] used Fisher LDA to classify AD and MCI from CN by large scale networks (LSN) analysis using brain pattern connectivity. Cho et al. [26] introduced incremental learning in neuroimaging and used a spatial

Table 3. Comparison of recent studies using deep learning and ensemble methods.

Sr. No.	Year	Authors	Target	Modality	Feature extraction	Machine learning	Dataset	Validation	Performance		
									Acc (%)	Sens (%)	Spec (%)
1	2007	Wang et al. [178]	CN vs AD	rs-fMRI	ROI	LR	44(18 AD, 26 CN)	-	81.1	-	-
2	2008	Fan et al. [41]	AD vs MCI	MRI, PET	VBM	PC	30(15 CN, 15 MCI)	LOOCV	93	-	-
3	2008	Davatzikos et al. [37]	AD-FTD CN vs FTD CN vs AD	sMRI (T1)	VBM	PC	98 (37 CN, 37 AD, 12 FTD, 12 CN)	LOOCV	84.3 100 100	- - -	- - -
4	2009	Hinrichs et al. [57]	CN vs AD	sMRI (T1), FDG-PET	GMPs WMPs	LP	183 (89 AD, 94 CN)	LMOCV	84	84	82
5	2009	Lopez et al. [106]	CN vs AD	SPECT	VAF	RF	147 (37 AD, 35 CN, 75 MCI 35 CN)	10-fold	89 74.6	87.5 77.5	90 67.9
6	2009	Horn et al. [61]	AD vs FTD	SPECT	ROI	KNN	173 (82 AD, 91 FTD)	LOOCV	88	93	85
7	2009	McEvoy et al. [114]	CN vs AD	sMRI (T1)	Morphometric measures	LR	398 (139 CN, 84 AD, 175 MCI)	LOOCV	89	83	93
8	2009	Desikan et al. [38]	CN vs AD CN vs MCI	sMRI (T1)	Anatomical Regions	LR	216 (94 CN, 65 AD 57 MCI)	-	95 95	- -	- -
9	2010	Plant et al. [131]	CN vs AD AD vs MCI CN vs MCI AD vs MCI vs CN	sMRI (T1)	GM WM CSF	SVM+Bayes+ VBI	74 (18 CN, 24 MCI, 32 AD)	LOOCV	92 95.83 97.62 75.0	93.75 100 95.83 55.56	88.89 93.33 100 86.67
10	2010	Ramirez et al. [135]	CN vs AD	SPECT	ROI	RF	97 (41 CN, 30 AD1, 22 AD2, 4 AD3)	LOOCV	89.69	100	92.7
11	2011	Chen et al. [20]	CN vs AD CN vs MCI	fMRI	116 ROI	Fisher LDA	55 (20 CN, 20 AD, 15 MCI)	LOOCV	87 91	85 95	80 93.90
12	2011	Dai et al. [35]	CN vs AD	sMRI, rs-fMRI	ALFF,GMD,RFCS	MLDA	38 (22 CN, 16 AD)	LOOCV	89.47	87.5	90.91
13	2011	Westman et al. [182]	CN vs AD MCInc vs MCIc	PET	ROI	OPLS	1067 (295 AD, 335 CN, 353 MCInc, 84 MCIc)	7-fold	-	83.4 71.4	87.8 60.1
14	2011	Chincarini et al. [25]	CN vs AD CN vs MCIc MCInc vs MCIc	sMRI (T1)	CSF, GM, WM	RF+SVM	635 (144 AD, 302 MCI, (136 MCIc 166 MCInc) 189 CN)	20-fold	97 92 74	89 89 72	94 80 65
15	2011	Rao et al. [137]	CN vs AD	sMRI (T1)	GM Density	LR	140(75 AD, 65 CN)	10-fold	85.25	90.77	80.26
16	2012	Liu et al. [100]	CN vs AD CN vs MCI	sMRI (T1)	GMPMs	SRC ensemble	652(198 AD, 225 MCI, (112 sMCI 113 sMCI) 229 CN)	10-fold	90.80 87.85	86.32 85.12	94.76 90.40
17	2012	Termenon et al. [169]	CN vs AD	sMRI	VBM	RVM+SVM	98 (49 AD, 49 CN)	10-fold	82	83	81
18	2012	Cho et al. [26]	CN vs AD MCIc vs MCInc	sMRI	Cortical Thickness	LDA	491 (128 AD, 160 CN, (72 MCIc, 131 MCInc))	-	-	82 63	93 73
19	2013	Spulber et al. [157]	CN vs AD CN vs MCI	sMRI (T1)	ROI	OPLS	1064 (295 AD, 335 CN, 261 sMCI, 173 MCIc)	7-fold	84.4 67.7	86.1 69.6	90.4 66.6
20	2013	Suk et al. [162]	CN vs AD CN vs MCI MCIc vs MCInc	MRI PET CSF	ROI	SAE	202 (51 AD, 99 MCI, 43 MCIc, 56 MCInc, 52 CN)	10-fold	89.9 73.7 60.2	- - -	- - -
21	2013	Gray et al. [51]	CN vs AD CN vs MCI pMCI vs sMCI	MRI, PET, CSF	83 ROI	RF	147 (37 AD, 75 MCI, (34 pMCI, 41 sMCI), 35 CN)	10-fold	89 75 58	87.9 77.5 57.1	90 67.9 58.7
22	2013	Chen et al.[21]	CN vs AD	MRI	GM maps	HMM	150 (75 AD, 75 CN)	LOOCV	80.7	81.3	80.0
23	2013	Liu et al. [103]	CN vs MCIc CN vs MCInc CN vs AD MCInc vs MCIc MCInc vs AD MCIc vs AD	sMRI (T1)	94 ROI	LR	413 (86 AD, 190 MCI, 97 MCIc, 93 MCInc, 137 CN)	LOOCV	-	65 81 86 80 77 56	63 82 93 56 73 61
24	2013	Casanova et al. [17]	CN vs AD CN vs MCInc CN vs MCIc MCInc vs MCIc	sMRI (T1), DNA, Cognitive data	Voxel intensities of GM, WM and CSF	RLR	689 (171 AD, 330 MCI, (153 MCIc, 182 MCInc), 188 CN)	10-fold	87.1 64.5 72.3 63.0	84.3 57.9 66.3 58.6	88.9 70.1 78.9 68.1
25	2014	Suk et al. [159]	CN vs AD CN vs MCI MCIc vs MCInc	MRI, FDG-PET	GM Intensity	CNN	780 (148 AD,403 MCI, 167 MCIc, 236 MCInc, 229 CN)	RP	92.87 76.21 72.44	- - -	- - -
26	2014	Singh et al. [155]	MCInc vs MCIc	PET CSF	AS and NMA	QDA	127 (73 MCInc, 54 MCIn)	LOOCV	66.14	64.81	67.12
27	2014	Lebedev et al. [90]	CN vs AD MCInc vs MCIc	sMRI (T1)	41 ROI	RF	575 (185 AD, 225 CN, 35 MCIc, 130 MCInc)	5-fold	- -	88.6 81.3	92 83.3
28	2014	Li et al. [93]	MCInc-MCIc	MRI	ROI	RF	293 (161 MCIc, 132 MCInc)	5-fold	74.7	66.7	81.4
29	2014	Jiang et al. [71]	CN vs AD	PET	83 ROI	KNN	103 (51 AD, 52 CN)	10-fold	91	92	-
30	2014	Lillemark et al. [95]	CN vs AD AD vs MCI	sMRI (T1)	ROI	LDA	524 (170 CN, 114 AD, 240 MCI)	-	87.70 76.60	- -	- -
31	2015	Korolev et al. [83]	MCIc vs pMCI	sMRI	ROI	pMKL	259 (120 MCIn, 139 pMCI)	10-fold	80.0	83.4	76.4
32	2015	Cabral et al. [16]	sMCI vs pMCI	FDG-PET	VI	SVM, GNB	100 (44 pMCI, 56 sMCI)	10-fold	85	-	-
33	2015	Payan et al. [127]	CN vs AD AD vs MCI CN vs MCI	MRI	3D patch	SAE	2265 (755 AD, 755 CN, 755 MCI)	-	95.38 86.84 92.11	- - -	- - -
34	2015	Suk et al. [160]	CN vs AD MCI vs CN AD vs MCI MCIc vs MCInc	sMRI (T1), PET CSF	GM, WM	SAE	202 (51 AD, 99 MCI (43 MCIc, 56 MCInc), 52 CN)	10-fold	98.8 90.70 83.70 88.30	- - - -	- - - -
35	2015	Liu et al. [102]	CN vs AD CN vs MCI CN vs MCInc vs MCIn vs AD	sMRI (T1), PET	ROI	SAE	758 (204 CN, 180 AD, 214 MCIc, 160 MCInc)	10-fold	91.40 82.10 53.79	92.32 60.00 52.14	90.42 92.32 86.98
36	2016	Hosseini et al. [62]	CN vs AD AD vs MCI CN vs MCI	sMRI	VBM	CNN	210 (70 AD, 70 CN, 70 MCI)	10-fold	99.30 100 94.20	100 100 97.40	98.60 100 91.40
37	2016	Suk et al. [165]	CN vs MCI	rs-fMRI	ROI	CNN	37 (12 MCI, 25 CN)	LOOCV	81.08	77.78	82.14

Table 3 (Contd.)

Sr. No.	Year	Authors	Target	Modality	Feature extraction	Machine learning	Dataset	Validation	Performance		
									Acc (%)	Sens (%)	Spec (%)
38	2016	Ortiz et al.[123]	CN vs AD CN vs MCIc CN vs MCIs MCI vs AD MCInc vs MCIc	sMRI (T1)	VBM	DBN	241 (68 CN, 70 AD, 39 MCIc, 64 MCInc)	10-fold	90 83 80 84 78	86 67 60 79 61	94 95 90 89 88
39	2016	Suk et al. [164]	CN vs AD MCI vs CN MCIc vs MCIs	MRI	ROI	SAE	805 (186 AD, 393 MCI (167 MCIc, 226 MCIs), 226 CN)	10-fold	91.02 73.02 74.82	92.72 77.60 70.93	89.94 68.22 78.82
40	2016	Zhang et al. [190]	CN vs AD MCInc vs MCIc	MRI	TBM	Adaboost	810 (194 AD, 228 CN, 246 MCInc, 142 MCIc)	LOOCV	81 77	83 82	78 76
41	2016	Hor et al. [60]	AD vs MCIc AD vs MCInc MCInc vs MCIc	MRI, PET	ROI	RF	608 (108 AD, 322 MCI, (96 MCIc, 126 MCInc), 178 CN)	5-fold	78.8 79.5 81.5	74.7 73.7 83.1	80.5 89.7 80.3
42	2016	Zheng et al. [196]	CN vs AD	MRI, PET	93 ROI	MMSDPN	103 (51 AD, 52 CN)	10-fold	97.27	97.32	98.33
43	2016	Clark et al. [28]	CN vs MCI	sMRI	CT and volume	Ensembled classifier	158 (51 CN, 24 MCIc, 83 MCInc)	-	83.2	62.5	89.2
44	2016	Schouten et al. [148]	CN vs AD	rs-fMRI, MRI	GM density	RLR	250 (77 AD, 173 CN)	10-fold	89.6	82.6	92.7
45	2017	Sarraf et al. [145]	CN vs AD	rs-fMRI	Slice based	DL-CNN	144 (52 AD, 92 CN)	5-fold	100	-	-
46	2017	Suk et al. [161]	CN vs AD CN vs MCI pMCI vs sMCI	sMRI	93 ROI	JLLR DeepESRNet	805 (186 AD, 393 MCI, (167 pMCI, 226 sMCI), 226 CN)	10-fold	91.02 73.02 74.82	92.72 77.66 70.93	89.94 68.22 78.82
47	2016	Zhu et al. [199]	CN vs AD CN vs MCI MCIc vs MCInc CN vs MCI vs AD CN vs MCIc vs MCInc vs AD	sMRI (T1)+PET +CSF	Atlas	LDA+LPP	202 (52 CN, 43 MCIc vs 56 MCInc, 51 AD)	-	95.5 ± 1.05 79.7 ± 0.21 71.2 ± 1.22 73.35 ± 1.53 61.06 ± 1.40	-	-
48	2017	Leracitano et al. [64]	CN VS AD AD VS MCI CN vs MCI CN vs AD vs MCI	EEG	PSD, Epoch	CNN	189 (63 AD, 63 MCI, 63 CN)	-	92.95 84.62 91.99 83.33	- - - -	- - - -
49	2017	Asgari et al. [8]	AD vs MCI	Word count	LIWC	SVM+RF	41(27 CN, 14 MCI)	5-fold	74.7	6.51	72.3
50	2018	Liu et al. [101]	CN vs AD CN vs pMCI vs sMCI vs AD	sMRI (T1)	Patch based	Multi task multi channel deep neural network	1984 (881 CN, 459 sMCI, 202 pMCI , 442 AD)	-	93.7 51.8	94.6 -	93.2 -
51	2018	Basaia et al. [9]	CN vs AD CN vs MCIc CN vs MCIs AD vs MCIc MCIs vs AD MCIc vs MCIs	sMRI (T1)	GM, WM, CSF	CNN	1566 (418 AD, 741 MCI, (208 MCIc, 533 MCIs), 407 CN)	10-fold	98.2 87.7 76.4 75.8 86.3 74.9	98.1 87.3 75.1 74.8 84.0 75.8	98.3 88.1 77.1 77.1 88.7 74.1
52	2018	Lu et al. [107]	CN vs AD sMCI vs pMCI	FDG-PET	ROI	MDNN	1051 (226 AD, 304 CN, 409 sMCI, 112 pMCI)	10-fold	93.58 81.55	91.54 73.33	95.06 83.83
53	2018	Li et al. [92]	CN vs AD CN vs MCI	sMRI (T1)	Patch based	DenseNet	831 (199 AD, 229 CN, 403 MCI)	5-fold	89.5 73.8	87.9 86.6	90.8 51.5
54	2018	Zheng et al. [195]	CN vs AD CN vs MCI AD vs MCI sMCI vs pMCI	sMRI (T1)	ROI	MFN	528 (142 AD, 221 MCI, 126 pMCI, 75 sMCI, 165 CN)	10-fold	98.7 97.93 7383 67.92	98.59 98.64 64.08 72.22	98.79 96.97 80.09 61.05
55	2018	Dominguez et al. [54]	CN vs AD CN vs MCI	Audio transcription	Cov+lin	RF	517 (257 AD, 217 CN, 43 MCI)	10-fold	94 87	100 87	86 86
56	2018	Hett et al. [55]	CN vs AD CN vs pMCI AD vs sMCI pMCI vs sMCI	sMRI (T1)	Hippocampal segments	Fusion SVM	651(226 CN, 186 AD, 223 sMCI, 16 pMCI)	10-fold	- - - -	93.4 91.6 77.6 77.0	87.6 83.0 71.0 64.1
57	2018	Jain et al. [68]	CN vs AD AD vs MCI CN vs MCI	sMRI (T1)	Entropy based	CNN	150 (50 CN, 50 AD, 50 MCI)	-	99.14 99.30 99.22	- - -	- - -
58	2019	Krishnakumar et al. [84]	CN vs AD AD vs MCI CN vs MCI MCIc-MCInc	sMRI (T1)	GM, WM	SVM+RF+KNN	812 (227 CN, 396 MCI, (165 MCIc, 223 MCInc), 189 AD)	k-fold k=1 to 50	87.39 63.41 64.74 66.38	89.58 57.29 45.61 60.24	85.82 65.35 72.44 69.80
59	2019	Spasov et al. [156]	CN vs AD sMCI vs pMCI	sMRI (T1)	ROI, APOe4	CNN	785 (192 AD, 184 CN, 228 sMCI, 181 pMCI)	10-fold	100 92.5	100 86.5	100 85
60	2019	Wang et al. [177]	AD vs MCIs CN vs MCIs CN vs AD CN vs AD vs MCI	sMRI (T1)	Volume	3D-CNN	833 (221 AD, 297 MCI, 315 CN)	10-fold	93.61 98.42 98.83 97.52	- - - -	- - - -

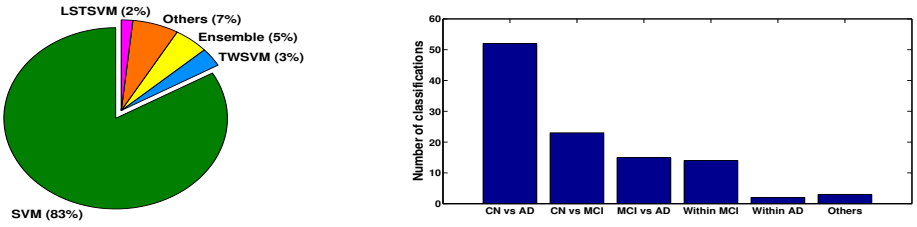
Abbreviations: **RLR**- Regularized logistic regression, **BM**- Biological markers, **GMPs**- Grey matter probability maps, **WMPs**- White matter probability maps, **LMO**- Leave many out, **FDG**- Fluorodeoxyglucose, **MDNN**- Multi scale deep neural network, **ALLF**- Amplitude of low-frequency fluctuations, **ReHo**- Regional homogeneity, **RFCS**- Regional functional connectivity strength, **MLDA**- Maximum uncertainty linear discriminant analysis, **Cov+lin**- Coverage and linguistic measures, **RFCS**- Regional functional connectivity strength, **GMPMs**- Gray matter probability maps, **AS**- Anatomical shapes,

NMA- Neuronal metabolic activity, **pMKL**- Probabilistic multiple kernel learning, **MFN**- Multifeature-based network, **MMSDPN**- Multimodal stacked deep polynomial network.

frequency representation of cortical thickness (CT) data using LDA. LDA was also used on multimodal data, surface connectivity data and MRI in [35, 84, 95]. Chen et al. [21] utilized hidden markov model (HMM) on features from MRI slices for early AD vs CN classification. Westman et al. [182] combined MRI data from two different cohorts and used orthogonal partial least square to latent structures (OPLS) models for combined and individual cohorts. Spulber et al. [157] also used OPLS to generate a severity index and found that there were significant differences in ROIs of MCIs (MCI stable) and MCIc. Bayesian classification methods were also used for AD classification in [16, 131]. Hett et al. [55] used texture-based grading framework which is capable of capturing structural alterations caused by AD using fusion SVM. Some other works like [8, 135, 148, 155, 190] have used different features which have shown significant results.

6 FUTURE DIRECTIONS

In the classification of dementia related data, there are various categories or targets. One classification target is MCI vs AD, which is one of the most important targets for early diagnosis of AD. It can be observed in fig. 6(b), 7(b), and 9(b) that most of the work has been done in classification of CN vs AD and CN vs MCI. Moreover, classifications like MCI vs AD are very less. This needs to be addressed in future research for early detection of AD. Other categories like MCIc vs MCInc, and MCIc vs MCIs are also addressed in very few papers. Therefore, researchers can focus on these particular problems for early detection of dementia caused by Alzheimer's disease.



(a) (b)

Fig. 6. Plot showing usage of (a) different variants of SVM and (b) different target groups for Alzheimers.

In AD diagnosis, an important focus point for research is the development of individual specific diagnosis models. For this, multimodal clinical data can be utilized as per the population. Moreover, novel learning techniques need to be developed for small datasets, since in real world scenarios the sample size from some population may not be large for training of model. Further, the data collection for Alzheimers includes noise from various sources. So, noise insensitive techniques must be applied for AD classification. In the next subsections, we suggest some future directions specific to the different machine learning approaches used for Alzheimers.

6.1 SVM

As discussed in previous sections, different variants of SVM have been employed for classification of Alzheimers. The usage of different types of SVM in our survey is shown in fig. 6(a). One can notice that among the different variants of SVM, 83% of the papers used standard SVM. This shows the popularity and robustness of SVM in the classification of

MRI data [112]. In 3% of the papers, TWSVM is used [5, 192], whereas LSTSVM [183] is used in only 1 paper. CSVM is used in [45, 158].

Some papers used ensemble of SVMs to classify Alzheimer's data [7, 65]. However, it can be observed that only 7 % of the papers are in the others category. This category involves the algorithms based on SVM which are modified especially for Alzheimers. One can observe that very few variants of SVM have been applied for AD. This shows that research is needed in application of other variants of SVM for Alzheimers. Moreover, other than the existing models, some novel variants of SVM also need to be developed for Alzheimer's disease as was done in [33]. In this paper, the spatial regularization on MRI image is included with SVM using graph Laplacian approach. Also, one can develop and use novel kernel functions for diagnosis of AD using SVM. Such kind of novel models can increase the classification performance of SVM.

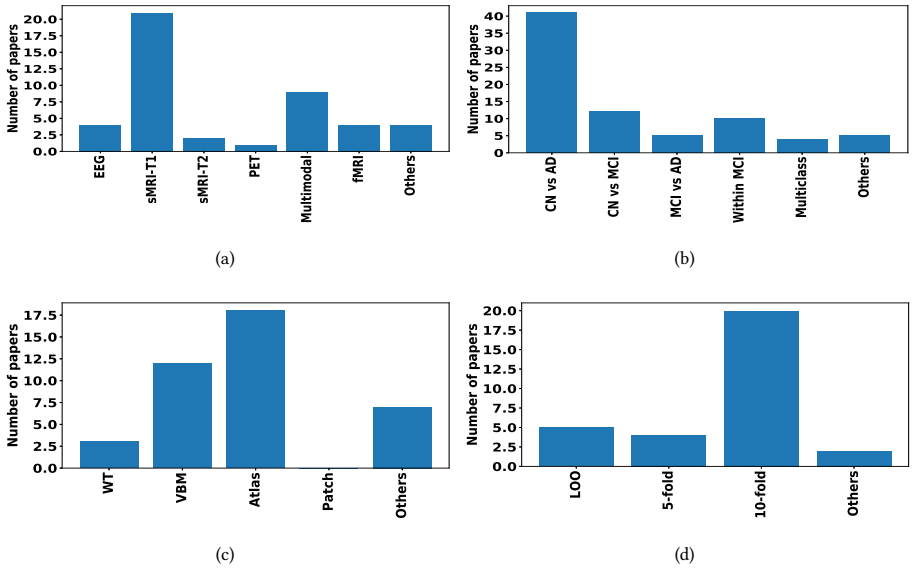


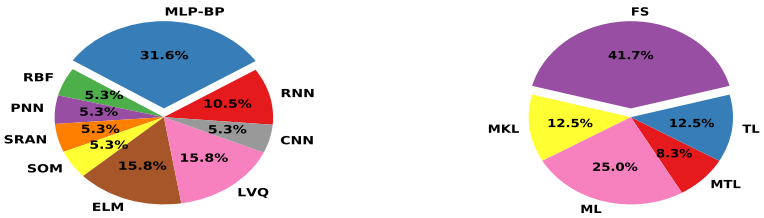
Fig. 7. Plots for application of ANN on different (a) modalities (b) classification tasks (c) feature extraction methods, and (d) cross validation methods

6.2 ANN, TL and MKL techniques

Techniques like transfer learning and multikernel learning can be used on multimodal data. We can see in fig. 7(a) that T1-weighted structural MRI is the most used modality. For feature extraction from MR images, atlas based methods have been mostly used by the researchers as shown in fig. 7(c). Atlas based methods employ some type of atlas to parcellate the brain regions from which different types of features are extracted. In future, other feature selection techniques can be utilized with ANN. For validating the model performance, 10-fold cross validation was the mostly used technique in fig. 7(d). However, one can also use methods like LOOCV, since it is the mostly used cross-validation method for SVM based algorithms in fig. 3(b).

The usage of different architecture of ANNs are shown in fig. 8(a). Mostly ANNs with backpropagation were used. Fig. 8(b) shows different learning paradigms used by researchers. Most of the effort has been given on selection of

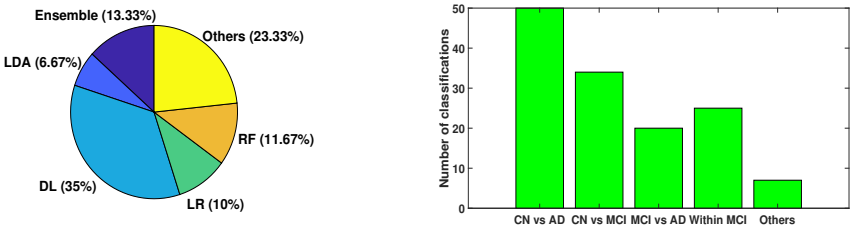
better and more informative features. More informative features alone are insufficient for improving performance on classification of AD. Thus, more effort needs to be given on producing novel classifiers specifically designed for handling neuroimaging data. Also, more efforts need to be given on approaches like TL and MTL as they can aid in dealing with the problem of small sample size and high dimensionality.



(a) (b)

Fig. 8. Pie charts showing usage of (a) variants of ANN and (b) different learning paradigms.

It can be seen from fig. 7(a) that less efforts have been given on utilizing data from multiple modalities as compared to data from single modality. Hence, more efforts should be given on developing models that can leverage multimodal data.



(a) (b)

Fig. 9. (a) Plot showing usage of different types of machine learning algorithms, and (b) targets used for classification of AD data.

6.3 DL and ensemble methods

In most of the papers using DL and ensemble techniques, T1-weighted structural MRI is used for classification of AD as shown in Table 3. Most of the researchers used ROI based features to classify or diagnose AD. DL techniques based on CNN architectures perform better for classification of AD. Moreover, 10-fold cross-validation is mostly used for validation of learning algorithms as shown in fig. 5(b).

The usage of DL, ensemble, and other classification techniques in our survey is shown in fig. 9(a). Among the different techniques, 35% of the papers used DL, 10% of the papers used LR, 11.67% of the papers used RF, 6.67% of papers used LDA, and 13.33% of papers used ensemble methods, and about 23.33% of the papers used other methods. The DL, ensemble and other techniques reviewed in our paper have issues related to model interpretability for routine use by clinicians. It can be observed in fig. 5(a) that most of the machine learning algorithms use MRI scans. Hence, more focus

should be given on leveraging data from various other modalities. In future, DL models can be trained on large sized Alzheimers datasets for better classification performance.

7 CONCLUSION

In this work, papers using three major machine learning techniques - SVM, ANN and DL are analyzed for diagnosis of Alzheimers. Research on other learning techniques like transfer, ensemble, and multi-kernel learning is also discussed. This can be useful for researchers working on any of these techniques to work on Alzheimers. As per this survey, it can be stated that SVM based models have been widely used for Alzheimer's disease showing its robustness. This is because techniques like ANN suffers from the drawbacks of local minima, which is not the case with SVM. However, ANNs are more versatile and robust when it comes to incremental learning [110], modelling sequential data [130], and quantizing high dimensional spaces [121]. Therefore, novel variants of ANN can be used for Alzheimer's in future. Deep learning and ensemble learning techniques give promising results by modeling highly complex data with high accuracy. The abundant usage of SVM also stems from the fact that it is easier to interpret as compared to deep neural networks which act as black box models. This problem should be addressed in future by focusing on clinical interpretability of deep learning models. Also, more work is needed in proper integration of feature selection techniques and machine learning model for a particular modality of data.

It is also observed that researchers have given more importance to the feature extraction phase, and not much to the classification phase. This can be addressed in future research, since novel models can give some new insight in the diagnosis of Azheimers. Moreover, more work is required in formulation of machine learning models which can integrate information from various modalities for early diagnosis of Alzheimer's disease.

ACKNOWLEDGMENTS

This work was supported by Department of Science and Technology, INDIA as Ramanujan fellowship grant no. SB/S2/RJN-001/2016. It is also supported by Council of Scientific & Industrial Research (CSIR), New Delhi, INDIA under Extra Mural Research (EMR) Scheme grant no. 22(0751)/17/EMR-II. We gratefully acknowledge the Indian Institute of Technology Indore for providing facilities and support. We are thankful to the Indian Institute of Technology Indore for providing Institute fellowship to Mr. Bharat Richhariya. We are also thankful to CSIR, New Delhi, INDIA for providing research fellowship to Mr. Riyaj Uddin Khan.

REFERENCES

[1] Ahmed Abdulkadir, Bénédicte Mortamet, Prashanthi Vemuri, Clifford R Jack Jr, Gunnar Krueger, Stefan Klöppel, Alzheimer's Disease Neuroimaging Initiative, et al. 2011. Effects of hardware heterogeneity on the performance of SVM Alzheimer's disease classifier. *Neuroimage* 58, 3 (2011), 785–792.

[2] Mehran Ahmadlou, Hojjat Adeli, and Anahita Adeli. 2010. New diagnostic EEG markers of the Alzheimer's disease using visibility graph. *Journal of neural transmission* 117, 9 (2010), 1099–1109.

[3] B Al-Naami, N Gharaibeh, and A AlRazzaq Kheshman. 2013. Automated detection of Alzheimer disease using region growing technique and artificial neural network. *World Acad. Sci. Eng. Technol. Int. J. Biomed. Biol. Eng* 7, 5 (2013).

[4] Saruar Alam, Goo-Rak Kwon, and Alzheimer's Disease Neuroimaging Initiative. 2017. Alzheimer disease classification using KPCA, LDA, and multi-kernel learning SVM. *International Journal of Imaging Systems and Technology* 27, 2 (2017), 133–143.

[5] Saruar Alam, Goo-Rak Kwon, Ji-In Kim, and Chun-Su Park. 2017. Twin SVM-based classification of Alzheimer's disease using complex dual-tree wavelet principal coefficients and LDA. *Journal of healthcare engineering* 2017 (2017).

[6] Almir Aljović, Almir Badnjević, and Lejla Gurbeta. 2016. Artificial neural networks in the discrimination of Alzheimer's disease using biomarkers data. In *2016 5th Mediterranean Conference on Embedded Computing (MECO)*. IEEE, 286–289.

[7] Ignacio Álvarez, Míriam López, Juan Manuel Górriz, Javier Ramírez, Diego Salas-Gonzalez, Carlos García Puntonet, and Fermín Segovia. 2008. Automatic classification system for the diagnosis of Alzheimer disease using Component-Based SVM aggregations. In *International Conference on*

Neural Information Processing. Springer, 402–409.

[8] Meysam Asgari, Jeffrey Kaye, and Hiroko Dodge. 2017. Predicting mild cognitive impairment from spontaneous spoken utterances. *Alzheimer's & Dementia: Translational Research & Clinical Interventions* 3, 2 (2017), 219–228.

[9] Silvia Basaia, Federica Agosta, Luca Wagner, Elisa Canu, Giuseppe Magnani, Roberto Santangelo, Massimo Filippi, Alzheimer's Disease Neuroimaging Initiative, et al. 2018. Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks. *NeuroImage: Clinical* (2018), 101645.

[10] Randall J Bateman, Paul S Aisen, Bart De Strooper, Nick C Fox, Cynthia A Lemere, John M Ringman, Stephen Salloway, Reisa A Sperling, Manfred Windisch, and Chengjie Xiong. 2011. Autosomal-dominant Alzheimer's disease: a review and proposal for the prevention of Alzheimer's disease. *Alzheimer's research & therapy* 3, 1 (2011), 1.

[11] Iman Beheshti, Hasan Demirel, Hiroshi Matsuda, and Alzheimer's Disease Neuroimaging Initiative. 2017. Classification of Alzheimer's disease and prediction of mild cognitive impairment-to-Alzheimer's conversion from structural magnetic resource imaging using feature ranking and a genetic algorithm. *Computers in biology and medicine* 83 (2017), 109–119.

[12] Iman Beheshti, Norihide Maikusa, Morteza Daneshmand, Hiroshi Matsuda, Hasan Demirel, and Gholamreza Anbarjafari. 2017. Classification of Alzheimer‘s disease and prediction of mild cognitive impairment conversion using histogram-based analysis of patient-specific anatomical brain connectivity networks. *Journal of Alzheimer‘s Disease* 60, 1 (2017), 295–304.

[13] Xia-an Bi, Qin Jiang, Qi Sun, Qing Shu, and Yingchao Liu. 2018. Analysis of Alzheimer's disease based on the random neural network cluster in fMRI. *Frontiers in neuroinformatics* 12 (2018), 60.

[14] Xia-an Bi, Qing Shu, Qi Sun, and Qian Xu. 2018. Random support vector machine cluster analysis of resting-state fMRI in Alzheimer‘s disease. *PloS one* 13, 3 (2018), e0194479.

[15] Halil Bisgin, Tanmay Bera, Hongjian Ding, Howard G Semey, Leihong Wu, Zhichao Liu, Amy E Barnes, Darryl A Langley, Monica Pava-Ripoll, Himansu J Vyas, et al. 2018. Comparing SVM and ANN based Machine Learning Methods for Species Identification of Food Contaminating Beetles. *Scientific reports* 8 (2018).

[16] Carlos Cabral, Pedro M Morgado, Durval Campos Costa, Margarida Silveira, Alzheimer‘s Disease Neuroimaging Initiative, et al. 2015. Predicting conversion from MCI to AD with FDG-PET brain images at different prodromal stages. *Computers in biology and medicine* 58 (2015), 101–109.

[17] Ramon Casanova, Fang-Chi Hsu, Kaycee M Sink, Stephen R Rapp, Jeff D Williamson, Susan M Resnick, Mark A Espeland, and Alzheimer's Disease Neuroimaging Initiative. 2013. Alzheimer's disease risk assessment using large-scale machine learning methods. *PLoS One* 8, 11 (2013), e77949.

[18] Sandeep Chaplot, LM Patnaik, and NR Jagannathan. 2006. Classification of magnetic resonance brain images using wavelets as input to support vector machine and neural network. *Biomedical signal processing and control* 1, 1 (2006), 86–92.

[19] R Chaves, J Ramírez, JM Górriz, M López, D Salas-Gonzalez, I Alvarez, and F Segovia. 2009. SVM-based computer-aided diagnosis of the Alzheimer's disease using t-test NMSE feature selection with feature correlation weighting. *Neuroscience letters* 461, 3 (2009), 293–297.

[20] Gang Chen, B Douglas Ward, Chunming Xie, Wenjun Li, Zhilin Wu, Jennifer L Jones, Malgorzata Franczak, Piero Antuono, and Shi-Jiang Li. 2011. Classification of Alzheimer disease, mild cognitive impairment, and normal cognitive status with large-scale network analysis based on resting-state functional MR imaging. *Radiology* 259, 1 (2011), 213–221.

[21] Ying Chen and Tuan D Pham. 2013. Development of a brain MRI-based hidden Markov model for dementia recognition. *Biomedical engineering online* 12, 1 (2013), S2.

[22] Bo Cheng, Mingxia Liu, Dinggang Shen, Zuoyong Li, Daoqiang Zhang, and Alzheimer‘s Disease Neuroimaging Initiative. 2017. Multi-domain transfer learning for early diagnosis of Alzheimer‘s disease. *Neuroinformatics* 15, 2 (2017), 115–132.

[23] Bo Cheng, Mingxia Liu, Heung-Il Suk, Dinggang Shen, Daoqiang Zhang, and Alzheimer‘s Disease Neuroimaging Initiative. 2015. Multimodal manifold-regularized transfer learning for MCI conversion prediction. *Brain imaging and behavior* 9, 4 (2015), 913–926.

[24] Bo Cheng, Mingxia Liu, Daoqiang Zhang, Dinggang Shen, Alzheimer‘s Disease Neuroimaging Initiative, et al. 2018. Robust multi-label transfer feature learning for early diagnosis of Alzheimer‘s disease. *Brain imaging and behavior* (2018), 1–16.

[25] Andrea Chincarini, Paolo Bosco, Piero Calvini, Gianluca Gemme, Mario Esposito, Chiara Olivieri, Luca Rei, Sandro Squarcia, Guido Rodriguez, Roberto Bellotti, et al. 2011. Local MRI analysis approach in the diagnosis of early and prodromal Alzheimer's disease. *Neuroimage* 58, 2 (2011), 469–480.

[26] Youngsang Cho, Joon-Kyung Seong, Yong Jeong, Sung Yong Shin, Alzheimer's Disease Neuroimaging Initiative, et al. 2012. Individual subject classification for Alzheimer‘s disease based on incremental learning using a spatial frequency representation of cortical thickness data. *Neuroimage* 59, 3 (2012), 2217–2230.

[27] Darya Chyzhyk, Alexandre Savio, and Manuel Graña. 2014. Evolutionary ELM wrapper feature selection for Alzheimer‘s disease CAD on anatomical brain MRI. *Neurocomputing* 128 (2014), 73–80.

[28] David Glenn Clark, Paula M McLaughlin, Ellen Woo, Kristy Hwang, Sona Hurtz, Leslie Ramirez, Jennifer Eastman, Reshil-Marie Dukes, Puneet Kapur, Thomas P DeRamus, et al. 2016. Novel verbal fluency scores and structural brain imaging for prediction of cognitive outcome in mild cognitive impairment. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring* 2 (2016), 113–122.

[29] Corinna Cortes and Vladimir Vapnik. 1995. Support vector machine. *Machine learning* 20, 3 (1995), 273–297.

[30] Nello Cristianini and Bernhard Scholkopf. 2002. Support vector machines and kernel methods: the new generation of learning machines. *Ai Magazine* 23, 3 (2002), 31–31.

[31] Ruoxuan Cui, Manhua Liu, and Gang Li. 2018. Longitudinal analysis for Alzheimer's disease diagnosis using RNN. In *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*. IEEE, 1398–1401.

[32] Rémi Cuingnet, Emilie Gerardin, Jérôme Tessieras, Guillaume Auzias, Stéphane Lehéricy, Marie-Odile Habert, Marie Chupin, Habib Benali, Olivier Colliot, and Alzheimer's Disease Neuroimaging Initiative. 2011. Automatic classification of patients with Alzheimer's disease from structural MRI: a comparison of ten methods using the ADNI database. *neuroimage* 56, 2 (2011), 766–781.

[33] Rémi Cuingnet, Joan Alexis Glaunès, Marie Chupin, Habib Benali, and Olivier Colliot. 2013. Spatial and anatomical regularization of SVM: a general framework for neuroimaging data. *IEEE transactions on pattern analysis and machine intelligence* 35, 3 (2013), 682–696.

[34] Helder Frederico da Silva Lopes, Jair M Abe, and Renato Anghinah. 2010. Application of paraconsistent artificial neural networks as a method of aid in the diagnosis of Alzheimer disease. *Journal of medical systems* 34, 6 (2010), 1073–1081.

[35] Zhengjia Dai, Chaogan Yan, Zhiqun Wang, Jinhui Wang, Mingrui Xia, Kuncheng Li, and Yong He. 2012. Discriminative analysis of early Alzheimer's disease using multi-modal imaging and multi-level characterization with multi-classifier (M3). *Neuroimage* 59, 3 (2012), 2187–2195.

[36] Anders M Dale, Bruce Fischl, and Martin I Sereno. 1999. Cortical surface-based analysis: I. Segmentation and surface reconstruction. *Neuroimage* 9, 2 (1999), 179–194.

[37] Christos Davatzikos, Susan M Resnick, X Wu, P Parmpi, and Christopher M Clark. 2008. Individual patient diagnosis of AD and FTD via high-dimensional pattern classification of MRI. *Neuroimage* 41, 4 (2008), 1220–1227.

[38] Rahul S Desikan, Howard J Cabral, Christopher P Hess, William P Dillon, Christine M Glastonbury, Michael W Weiner, Nicholas J Schmansky, Douglas N Greve, David H Salat, Randy L Buckner, et al. 2009. Automated MRI measures identify individuals with mild cognitive impairment and Alzheimer's disease. *Brain* 132, 8 (2009), 2048–2057.

[39] James D Doecke, Simon M Laws, Noel G Faux, William Wilson, Samantha C Burnham, Chiou-Peng Lam, Alinda Mondal, Justin Bedo, Ashley I Bush, Belinda Brown, et al. 2012. Blood-based protein biomarkers for diagnosis of Alzheimer disease. *Archives of neurology* 69, 10 (2012), 1318–1325.

[40] Juergen Dukart, Karsten Mueller, Henryk Barthel, Arno Villringer, Osama Sabri, Matthias Leopold Schroeter, and Alzheimer's Disease Neuroimaging Initiative. 2013. Meta-analysis based SVM classification enables accurate detection of Alzheimer's disease across different clinical centers using FDG-PET and MRI. *Psychiatry Research: Neuroimaging* 212, 3 (2013), 230–236.

[41] Yong Fan, Susan M Resnick, Xiaoying Wu, and Christos Davatzikos. 2008. Structural and functional biomarkers of prodromal Alzheimer's disease: a high-dimensional pattern classification study. *Neuroimage* 41, 2 (2008), 277–285.

[42] Julian Fritsch, Sebastian Wankerl, and Elmar Nöth. 2019. Automatic Diagnosis of Alzheimer's Disease Using Neural Network Language Models. In *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 5841–5845.

[43] Rudimar L Frozza, Mychael V Lourenco, and Fernanda G De Felice. 2018. Challenges for Alzheimer's disease therapy: Insights from novel mechanisms beyond memory defects. *Frontiers in neuroscience* 12 (2018), 37.

[44] Motonobu Fujishima, Atsushi Kawaguchi, Norihide Maikusa, Ryozo Kuwano, Takeshi Iwatsubo, and Hiroshi Matsuda. 2017. Sample size estimation for Alzheimer's disease trials from Japanese ADNI serial magnetic resonance imaging. *Journal of Alzheimer's Disease* 56, 1 (2017), 75–88.

[45] Glenn Fung and Jonathan Stoeckel. 2007. SVM feature selection for classification of SPECT images of Alzheimer's disease using spatial information. *Knowledge and Information Systems* 11, 2 (2007), 243–258.

[46] Esteve Gallego-Jutglà, Jordi Solé-Casals, François-Benoît Vialatte, Mohamed Elgendi, Andrzej Cichocki, and Justin Dauwels. 2015. A hybrid feature selection approach for the early diagnosis of Alzheimer's disease. *Journal of neural engineering* 12, 1 (2015), 016018.

[47] Andrés García-Floriano, Cuauhtémoc López-Martín, Cornelio Yáñez-Márquez, and Alain Abran. 2018. Support vector regression for predicting software enhancement effort. *Information and Software Technology* 97 (2018), 99–109.

[48] Emilie Gerardin, Gaël Chételat, Marie Chupin, Rémi Cuingnet, Béatrice Desgranges, Ho-Sung Kim, Marc Niethammer, Bruno Dubois, Stéphane Lehéricy, Line Garnero, et al. 2009. Multidimensional classification of hippocampal shape features discriminates Alzheimer's disease and mild cognitive impairment from normal aging. *Neuroimage* 47, 4 (2009), 1476–1486.

[49] HT Gorji and J Haddadnia. 2015. A novel method for early diagnosis of Alzheimer's disease based on pseudo Zernike moment from structural MRI. *Neuroscience* 305 (2015), 361–371.

[50] Gábor Gosztolya, Veronika Vincze, László Tóth, Magdolna Pákáski, János Kálmán, and Ildikó Hoffmann. 2019. Identifying Mild Cognitive Impairment and mild Alzheimer's disease based on spontaneous speech using ASR and linguistic features. *Computer Speech & Language* 53 (2019), 181–197.

[51] Katherine R Gray, Paul Aljabar, Rolf A Heckemann, Alexander Hammers, Daniel Rueckert, Alzheimer's Disease Neuroimaging Initiative, et al. 2013. Random forest-based similarity measures for multi-modal classification of Alzheimer's disease. *NeuroImage* 65 (2013), 167–175.

[52] Isabelle Guyon, Jason Weston, Stephen Barnhill, and Vladimir Vapnik. 2002. Gene selection for cancer classification using support vector machines. *Machine learning* 46, 1-3 (2002), 389–422.

[53] Sven Haller, Pascal Missonnier, FR Herrmann, Cristelle Rodriguez, M-P Deiber, Duy Nguyen, Gabriel Gold, K-O Lovblad, and Panteleimon Giannakopoulos. 2013. Individual classification of mild cognitive impairment subtypes by support vector machine analysis of white matter DTI. *American Journal of Neuroradiology* 34, 2 (2013), 283–291.

[54] Laura Hernández-Domínguez, Sylvie Ratté, Gerardo Sierra-Martínez, and Andrés Roche-Bergua. 2018. Computer-based evaluation of AlzheimerâĂŹs disease and mild cognitive impairment patients during a picture description task. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring* 10 (2018), 260–268.

[55] Kilian Hett, Vinh-Thong Ta, José V Manjón, Pierrick Coupé, Alzheimer's Disease Neuroimaging Initiative, et al. 2018. Adaptive fusion of texture-based grading for Alzheimer's disease classification. *Computerized Medical Imaging and Graphics* 70 (2018), 8–16.

[56] Antonio R Hidalgo-Muñoz, Javier Ramírez, Juan M Górriz, and Pablo Padilla. 2014. Regions of interest computed by SVM wrapped method for Alzheimer's disease examination from segmented MRI. *Frontiers in aging neuroscience* 6 (2014), 20.

[57] Chris Hinrichs, Vikas Singh, Lopamudra Mukherjee, Guofan Xu, Moo K Chung, Sterling C Johnson, and Alzheimer's Disease Neuroimaging Initiative. 2009. Spatially augmented LPboosting for AD classification with evaluations on the ADNI dataset. *Neuroimage* 48, 1 (2009), 138–149.

[58] Seyed Hani Hojjati, Ata Ebrahimzadeh, Ali Khazaee, Abbas Babajani-Feremi, and Alzheimer's Disease Neuroimaging Initiative. 2017. Predicting conversion from MCI to AD using resting-state fMRI, graph theoretical approach and SVM. *Journal of neuroscience methods* 282 (2017), 69–80.

[59] Marcia Hon and Naimul Mefraz Khan. 2017. Towards Alzheimer's disease classification through transfer learning. In *2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE, 1166–1169.

[60] Soheil Hor and Mehdi Moradi. 2016. Learning in data-limited multimodal scenarios: Scandent decision forests and tree-based features. *Medical image analysis* 34 (2016), 30–41.

[61] Jean-François Horn, Marie-Odile Habert, Aurélie Kas, Zoulikha Malek, Philippe Maksud, Lucette Lacomblez, Alain Giron, and Bernard Fertil. 2009. Differential automatic diagnosis between Alzheimer's disease and frontotemporal dementia based on perfusion SPECT images. *Artificial intelligence in medicine* 47, 2 (2009), 147–158.

[62] Ehsan Hosseini-Asl, Georgy Gimel'farb, and Ayman El-Baz. 2016. Alzheimer's disease diagnostics by a deeply supervised adaptable 3D convolutional network. *arXiv preprint arXiv:1607.00556* (2016).

[63] Chengzhong Huang, Bin Yan, Hua Jiang, and Dahui Wang. 2008. Combining voxel-based morphometry with artifical neural network theory in the application research of diagnosing alzheimer's disease. In *2008 International Conference on BioMedical Engineering and Informatics*, Vol. 1. IEEE, 250–254.

[64] Cosimo Ieracitano, Nadia Mammone, Alessia Bramanti, Amir Hussain, and Francesco C Morabito. 2019. A Convolutional Neural Network approach for classification of dementia stages based on 2D-spectral representation of EEG recordings. *Neurocomputing* 323 (2019), 96–107.

[65] IA Illán, JM Górriz, MM López, Javier Ramírez, Diego Salas-Gonzalez, Fermín Segovia, Rosa Chaves, and Carlos García Puntonet. 2011. Computer aided diagnosis of Alzheimer's disease using component based SVM. *Applied Soft Computing* 11, 2 (2011), 2376–2382.

[66] Takeshi Iwatsubo. 2010. Japanese Alzheimer's Disease Neuroimaging Initiative: present status and future. *Alzheimer's & Dementia* 6, 3 (2010), 297–299.

[67] Clifford R Jack Jr, Josephine Barnes, Matt A Bernstein, Bret J Borowski, James Brewer, Shona Clegg, Anders M Dale, Owen Carmichael, Christopher Ching, Charles DeCarli, et al. 2015. Magnetic resonance imaging in alzheimer's disease neuroimaging initiative 2. *Alzheimer's & Dementia* 11, 7 (2015), 740–756.

[68] Rachna Jain, Nikita Jain, Akshay Aggarwal, and D Jude Hemanth. 2019. Convolutional neural network based Alzheimer's disease classification from magnetic resonance brain images. *Cognitive Systems Research* (2019).

[69] Jayadeva, R Khemchandani, and Suresh Chandra. 2007. Twin support vector machines for pattern classification. *IEEE Transactions on pattern analysis and machine intelligence* 29, 5 (2007), 905–910.

[70] Debesh Jha, Ji-In Kim, and Goo-Rak Kwon. 2017. Diagnosis of Alzheimer's disease using dual-tree complex wavelet transform, PCA, and feed-forward neural network. *Journal of healthcare engineering* 2017 (2017).

[71] Qikun Jiang and Jun Shi. 2014. Sparse kernel entropy component analysis for dimensionality reduction of neuroimaging data. In *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 3366–3369.

[72] Biao Jie, Daoqiang Zhang, Bo Cheng, and Dinggang Shen. 2013. Manifold regularized multi-task feature selection for multi-modality classification in Alzheimer's disease. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 275–283.

[73] Biao Jie, Daoqiang Zhang, Chong-Yaw Wee, and Dinggang Shen. 2014. Topological graph kernel on multiple thresholded functional connectivity networks for mild cognitive impairment classification. *Human brain mapping* 35, 7 (2014), 2876–2897.

[74] Sandhya Joshi, Deepa Shenoy, PL Rrashmi, KR Venugopal, and LM Patnaik. 2010. Classification of Alzheimer's disease and Parkinson's disease by using machine learning and neural network methods. In *2010 Second International Conference on Machine Learning and Computing*. IEEE, 218–222.

[75] Rupali S Kamathe and Kalyani R Joshi. 2018. A novel method based on independent component analysis for brain MR image tissue classification into CSF, WM and GM for atrophy detection in Alzheimer's disease. *Biomedical Signal Processing and Control* 40 (2018), 41–48.

[76] Subrata Kar and D Dutta Majumder. 2019. A Novel Approach of Diffusion Tensor Visualization Based Neuro Fuzzy Classification System for Early Detection of Alzheimer's Disease. *Journal of Alzheimer's disease reports* Preprint (2019), 1–18.

[77] K Kazemi and N Noorizadeh. 2014. Quantitative comparison of SPM, FSL, and brainsuite for brain MR image segmentation. *Journal of biomedical physics & engineering* 4, 1 (2014), 13.

[78] Ali Khazaee, Ata Ebrahimzadeh, and Abbas Babajani-Feremi. 2016. Application of advanced machine learning methods on resting-state fMRI network for identification of mild cognitive impairment and Alzheimer's disease. *Brain imaging and behavior* 10, 3 (2016), 799–817.

[79] Laila Khedher, Ignacio A Illán, Juan M Górriz, Javier Ramírez, Abdelbasset Brahim, and Anke Meyer-Baese. 2017. Independent component analysis-support vector machine-based computer-aided diagnosis system for Alzheimer's with visual support. *International journal of neural systems* 27, 03 (2017), 1650050.

[80] Laila Khedher, Javier Ramírez, Juan Manuel Górriz, Abdelbasset Brahim, and IA Illán. 2015. Independent component analysis-based classification of Alzheimer's disease from segmented MRI data. In *International Work-Conference on the Interplay between Natural and Artificial Computation*.

Springer, 78–87.

[81] Jongin Kim and Boreom Lee. 2018. Identification of Alzheimer's disease and mild cognitive impairment using multimodal sparse hierarchical extreme learning machine. *Human brain mapping* 39, 9 (2018), 3728–3741.

[82] Stefan Klöppel, Cynthia M Stonnington, Carlton Chu, Bogdan Draganski, Rachael I Scahill, Jonathan D Rohrer, Nick C Fox, Clifford R Jack Jr, John Ashburner, and Richard SJ Frackowiak. 2008. Automatic classification of MR scans in Alzheimer's disease. *Brain* 131, 3 (2008), 681–689.

[83] Igor O Korolev, Laura L Symonds, Andrea C Bozoki, and Alzheimer's Disease Neuroimaging Initiative. 2016. Predicting progression from mild cognitive impairment to Alzheimer's dementia using clinical, MRI, and plasma biomarkers via probabilistic pattern classification. *PloS one* 11, 2 (2016), e0138866.

[84] V Krishnakumar, Latha Parthiban, Alzheimer's Disease Neuroimaging Initiative, et al. 2019. A Novel Texture Extraction Technique with T1 Weighted MRI for the Classification of Alzheimer's Disease. *Journal of neuroscience methods* (2019).

[85] NN Kulkarni and VK Bairagi. 2017. Extracting salient features for EEG-based diagnosis of Alzheimer's disease using support vector machine classifier. *IETE Journal of Research* 63, 1 (2017), 11–22.

[86] M Arun Kumar and Madan Gopal. 2009. Least squares twin support vector machines for pattern classification. *Expert Systems with Applications* 36, 4 (2009), 7535–7543.

[87] Salim Lahmiri and Mounir Boukadoum. 2014. New approach for automatic classification of Alzheimer's disease, mild cognitive impairment and healthy brain magnetic resonance images. *Healthcare technology letters* 1, 1 (2014), 32–36.

[88] Salim Lahmiri and Amir Shmuel. 2018. Performance of machine learning methods applied to structural MRI and ADAS cognitive scores in diagnosing AlzheimerâĂŹs disease. *Biomedical Signal Processing and Control* (2018).

[89] Ramesh Kumar Lama, Jeonghwan Gwak, Jeong-Seon Park, and Sang-Woong Lee. 2017. Diagnosis of Alzheimer's disease based on structural MRI images using a regularized extreme learning machine and PCA features. *Journal of healthcare engineering* 2017 (2017).

[90] AV Lebedev, Eric Westman, GJP Van Westen, MG Kramberger, Arvid Lundervold, Dag Aarsland, H Soininen, I Kłoszewska, P Mecocci, M Tsolaki, et al. 2014. Random Forest ensembles for detection and prediction of Alzheimer's disease with a good between-cohort robustness. *NeuroImage: Clinical* 6 (2014), 115–125.

[91] Wook Lee, Byungkyu Park, and Kyungsook Han. 2013. Classification of diffusion tensor images for the early detection of Alzheimer's disease. *Computers in biology and medicine* 43, 10 (2013), 1313–1320.

[92] Fan Li, Manhua Liu, and Alzheimer's Disease Neuroimaging Initiative. 2018. Alzheimer's disease diagnosis based on multiple cluster dense convolutional networks. *Computerized Medical Imaging and Graphics* 70 (2018), 101–110.

[93] Han Li, Yashu Liu, Pinghua Gong, Changshui Zhang, Jieping Ye, and Alzheimers Disease Neuroimaging Initiative. 2014. Hierarchical interactions model for predicting Mild Cognitive Impairment (MCI) to Alzheimer's Disease (AD) conversion. *PloS one* 9, 1 (2014), e82450.

[94] Wei Li, Yifei Zhao, Xi Chen, Yang Xiao, and Yuanyuan Qin. 2018. Detecting Alzheimer's Disease on Small Dataset: A Knowledge Transfer Perspective. *IEEE journal of biomedical and health informatics* (2018).

[95] Lene Lillemark, Lauge Sørensen, Akshay Pai, Erik B Dam, and Mads Nielsen. 2014. Brain region's relative proximity as marker for Alzheimer's disease based on structural MRI. *BMC medical imaging* 14, 1 (2014), 21.

[96] Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen Awm Van Der Laak, Bram Van Ginneken, and Clara I Sánchez. 2017. A survey on deep learning in medical image analysis. *Medical image analysis* 42 (2017), 60–88.

[97] Feng Liu, Chong-Yaw Wee, Huafu Chen, and Dinggang Shen. 2014. Inter-modality relationship constrained multi-modality multi-task feature selection for Alzheimer's Disease and mild cognitive impairment identification. *NeuroImage* 84 (2014), 466–475.

[98] Jin Liu, Min Li, Wei Lan, Fang-Xiang Wu, Yi Pan, and Jianxin Wang. 2018. Classification of Alzheimer's disease using whole brain hierarchical network. *IEEE/ACM transactions on computational biology and bioinformatics* 15, 2 (2018), 624–632.

[99] Jin Liu, Jianxin Wang, Bin Hu, Fang-Xiang Wu, and Yi Pan. 2017. Alzheimer's disease classification based on individual hierarchical networks constructed with 3-D texture features. *IEEE transactions on nanobioscience* 16, 6 (2017), 428–437.

[100] Manhua Liu, Daoqiang Zhang, Dinggang Shen, Alzheimer's Disease Neuroimaging Initiative, et al. 2012. Ensemble sparse classification of Alzheimer's disease. *NeuroImage* 60, 2 (2012), 1106–1116.

[101] Mingxia Liu, Jun Zhang, Ehsan Adeli, and Dinggang Shen. 2018. Joint Classification and Regression via Deep Multi-Task Multi-Channel Learning for Alzheimer's Disease Diagnosis. *IEEE Transactions on Biomedical Engineering* (2018).

[102] Siqi Liu, Sidong Liu, Weidong Cai, Hangyu Che, Sonia Pujol, Ron Kikinis, Dagan Feng, Michael J Fulham, et al. 2015. Multimodal neuroimaging feature learning for multiclass diagnosis of Alzheimer's disease. *IEEE Transactions on Biomedical Engineering* 62, 4 (2015), 1132–1140.

[103] Xin Liu, Duygu Tosun, Michael W Weiner, Norbert Schuff, and Alzheimer's Disease Neuroimaging Initiative. 2013. Locally linear embedding (LLE) for MRI based Alzheimer's disease classification. *Neuroimage* 83 (2013), 148–157.

[104] Xiaojing Long, Lifang Chen, Chunxiang Jiang, Lijuan Zhang, AlzheimerâĂŹs Disease Neuroimaging Initiative, et al. 2017. Prediction and classification of Alzheimer disease based on quantification of MRI deformation. *PloS one* 12, 3 (2017), e0173372.

[105] Xiaojing Long and Chris Wyatt. 2010. An automatic unsupervised classification of MR images in Alzheimer's disease. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE, 2910–2917.

[106] M López, J Ramírez, JM Górriz, D Salas-Gonzalez, I Alvarez, F Segovia, and CG Puntonet. 2009. Automatic tool for Alzheimer's disease diagnosis using PCA and Bayesian classification rules. *Electronics Letters* 45, 8 (2009), 389–391.

[107] Donghuan Lu, Karteek Popuri, Gavin Weiguang Ding, Rakesh Balachandar, Mirza Faisal Beg, Alzheimer's Disease Neuroimaging Initiative, et al. 2018. Multiscale deep neural network based analysis of FDG-PET images for the early diagnosis of Alzheimer‘s disease. *Medical image analysis* 46 (2018), 26–34.

[108] Shen Lu, Yong Xia, Weidong Cai, Michael Fulham, David Dagan Feng, AlzheimerâĂŹs Disease Neuroimaging Initiative, et al. 2017. Early identification of mild cognitive impairment using incomplete random forest-robust support vector machine and FDG-PET imaging. *Computerized Medical Imaging and Graphics* 60 (2017), 35–41.

[109] Benoît Magnin, Lilia Mesrob, Serge Kinkingnéhun, Mélanie Pélégrini-Issac, Olivier Colliot, Marie Sarazin, Bruno Dubois, Stéphane Lehéricy, and Habib Benali. 2009. Support vector machine-based classification of Alzheimer‘s disease from whole-brain anatomical MRI. *Neuroradiology* 51, 2 (2009), 73–83.

[110] Belathur Suresh Mahanand, Sundaram Suresh, Narasimhan Sundararajan, and M Aswatha Kumar. 2012. Identification of brain regions responsible for Alzheimer‘s disease using a Self-adaptive Resource Allocation Network. *Neural Networks* 32 (2012), 313–322.

[111] Rigel Mahmood and Bishad Ghimire. 2013. Automatic detection and classification of Alzheimer's Disease from MRI scans using principal component analysis and artificial neural networks. In *2013 20th International Conference on Systems, Signals and Image Processing (IWSSIP)*. IEEE, 133–137.

[112] José María Mateos-Pérez, Mahsa Dadar, María Lacalle-Aurioles, Yasser Iturria-Medina, Yashar Zeighami, and Alan C Evans. 2018. Structural neuroimaging as clinical predictor: A review of machine learning applications. *NeuroImage: Clinical* (2018).

[113] Ali Mazaheri, Katrien Segaert, John Olichney, Jin-Chen Yang, Yu-Qiong Niu, Kim Shapiro, and Howard Bowman. 2018. EEG oscillations during word processing predict MCI conversion to Alzheimer's disease. *NeuroImage: Clinical* 17 (2018), 188–197.

[114] Linda K McEvoy, Christine Fennema-Notestine, J Cooper Roddey, Donald J Hagler Jr, Dominic Holland, David S Karow, Christopher J Pung, James B Brewer, and Anders M Dale. 2009. Alzheimer disease: quantitative structural neuroimaging for detection and prediction of clinical and structural changes in mild cognitive impairment. *Radiology* 251, 1 (2009), 195–205.

[115] Lilia Mesrob, Benoit Magnin, Olivier Colliot, Marie Sarazin, Valérie Hahn-Barma, Bruno Dubois, Patrick Gallinari, Stéphane Lehéricy, Serge Kinkingnéhun, and Habib Benali. 2008. Identification of atrophy patterns in Alzheimer‘s disease based on SVM feature selection and anatomical parcellation. In *International Workshop on Medical Imaging and Virtual Reality*. Springer, 124–132.

[116] Rui Min, Guorong Wu, Jian Cheng, Qian Wang, Dinggang Shen, and Alzheimer‘s Disease Neuroimaging Initiative. 2014. Multi-atlas based representations for Alzheimer‘s disease diagnosis. *Human brain mapping* 35, 10 (2014), 5052–5070.

[117] Christiane Möller, Yolande AL Pijnenburg, Wiesje M van der Flier, Adriaan Versteeg, Betty Tijms, Jan C de Munck, Anne Hafkemeijer, Serge ARB Rombouts, Jeroen van der Grond, John van Swieten, et al. 2015. Alzheimer disease and behavioral variant frontotemporal dementia: automatic classification based on cortical atrophy for single-subject diagnosis. *Radiology* 279, 3 (2015), 838–848.

[118] Elaheh Moradi, Antonietta Pepe, Christian Gaser, Heikki Huttunen, Jussi Tohka, and AlzheimerâĂŹs Disease Neuroimaging Initiative. 2015. Machine learning framework for early MRI-based Alzheimer‘s conversion prediction in MCI subjects. *NeuroImage* 104 (2015), 398–412.

[119] Laurence O'Dwyer, Franck Lamberton, Arun LW Bokde, Michael Ewers, Yetunde O Faluyi, Colby Tanner, Bernard Mazoyer, Desmond O'Neill, Máiréad Bartley, D Rónán Collins, et al. 2012. Using support vector machines with multiple indices of diffusion for automated classification of mild cognitive impairment. *PloS one* 7, 2 (2012), e32441.

[120] Andrés Ortiz, Juan M Górriz, Javier Ramírez, Francisco Jesús Martínez-Murcia, and AlzheimerâĂŹs Disease Neuroimaging Initiative. 2013. LVQ-SVM based CAD tool applied to structural MRI for the diagnosis of the Alzheimer‘s disease. *Pattern Recognition Letters* 34, 14 (2013), 1725–1733.

[121] Andrés Ortiz, Juan M Górriz, Javier Ramírez, Francisco J Martinez-Murcia, and Alzheimer's Disease Neuroimaging Initiative. 2014. Automatic ROI selection in structural brain MRI using SOM 3D projection. *PloS one* 9, 4 (2014), e93851.

[122] Andrés Ortiz, Jorge Munilla, Ignacio Álvarez-Illán, Juan M Górriz, Javier Ramírez, and Alzheimer‘s Disease Neuroimaging Initiative. 2015. Exploratory graphical models of functional and structural connectivity patterns for Alzheimer‘s disease diagnosis. *Frontiers in computational neuroscience* 9 (2015), 132.

[123] Andres Ortiz, Jorge Munilla, Juan M Gorriz, and Javier Ramirez. 2016. Ensembles of deep learning architectures for the early diagnosis of the Alzheimer‘s disease. *International journal of neural systems* 26, 07 (2016), 1650025.

[124] P Padilla, JM Górriz, J Ramírez, EW Lang, R Chaves, F Segovia, M López, D Salas-González, and I Álvarez. 2010. Analysis of SPECT brain images for the diagnosis of Alzheimer's disease based on NMF for feature extraction. *Neuroscience letters* 479, 3 (2010), 192–196.

[125] Maria Paraskevaidi, Camilo LM Morais, Diane E Halliwell, David MA Mann, David Allsop, Pierre L Martin-Hirsch, and Francis L Martin. 2018. Raman spectroscopy to diagnose Alzheimer‘s disease and dementia with Lewy bodies in blood. *ACS chemical neuroscience* 9, 11 (2018), 2786–2794.

[126] Christina Patterson. 2018. The state of the art of dementia research: New frontiers. *World Alzheimer's Report 2018* (2018).

[127] Adrien Payan and Giovanni Montana. 2015. Predicting Alzheimer's disease: a neuroimaging study with 3D convolutional neural networks. *arXiv preprint arXiv:1502.02506* (2015).

[128] Enrico Pellegrini, Lucia Ballerini, Maria Del C Valdes Hernandez, Francesca M Chappell, Victor González-Castro, Devasuda Anblagan, Samuel Danso, Susana Muñoz-Maniega, Dominic Job, Cyril Pernet, et al. 2018. Machine learning of neuroimaging for assisted diagnosis of cognitive impairment and dementia: A systematic review. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring* 10 (2018), 519–535.

[129] Jialin Peng, Xiaofeng Zhu, Ye Wang, Le An, and Dinggang Shen. 2019. Structured sparsity regularized multiple kernel learning for Alzheimer‘s disease diagnosis. *Pattern recognition* 88 (2019), 370–382.

[130] AA Petrosian, DV Prokhorov, W Lajara-Nanson, and RB Schiffer. 2001. Recurrent neural network-based approach for early recognition of Alzheimer‘s disease in EEG. *Clinical Neurophysiology* 112, 8 (2001), 1378–1387.

[131] Claudia Plant, Stefan J Teipel, Annahita Oswald, Christian Böhm, Thomas Meindl, Janaina Mourao-Miranda, Arun W Bokde, Harald Hampel, and Michael Ewers. 2010. Automated detection of brain atrophy patterns based on MRI for the prediction of Alzheimer's disease. *Neuroimage* 50, 1 (2010), 162–174.

[132] Maciej Plocharski, Lasse Riis Østergaard, Alzheimer's Disease Neuroimaging Initiative, et al. 2016. Extraction of sulcal medial surface and classification of Alzheimer's disease using sulcal features. *Computer methods and programs in biomedicine* 133 (2016), 35–44.

[133] María Quintana, Joan Guàrdia, Gonzalo Sánchez-Benavides, Miguel Aguilar, José Luis Molinuevo, Alfredo Robles, María Sagrario Barquero, Carmen Antúnez, Carlos Martínez-Parra, Anna Frank-García, et al. 2012. Using artificial neural networks in clinical neuropsychology: High performance in mild cognitive impairment and Alzheimer's disease. *Journal of Clinical and Experimental Neuropsychology* 34, 2 (2012), 195–208.

[134] Javier Ramírez, JM Górriz, Diego Salas-Gonzalez, A Romero, Míriam López, Ignacio Álvarez, and Manuel Gómez-Río. 2013. Computer-aided diagnosis of Alzheimer's type dementia combining support vector machines and discriminant set of features. *Information Sciences* 237 (2013), 59–72.

[135] J Ramírez, JM Górriz, F Segovia, R Chaves, D Salas-Gonzalez, M López, I Álvarez, and P Padilla. 2010. Computer aided diagnosis system for the Alzheimer's disease based on partial least squares and random forest SPECT image classification. *Neuroscience letters* 472, 2 (2010), 99–103.

[136] Javier Ramírez, Juan Manuel Górriz, Míriam López, Diego Salas-Gonzalez, Ignacio Álvarez, Fermín Segovia, and Carlos García Puntonet. 2008. Early detection of the alzheimer disease combining feature selection and kernel machines. In *International Conference on Neural Information Processing*. Springer, 410–417.

[137] Anil Rao, Ying Lee, Achim Gass, and Andreas Monsch. 2011. Classification of Alzheimer's Disease from structural MRI using sparse logistic regression with optional spatial regularization. In *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 4499–4502.

[138] Saima Rathore, Mohamad Habes, Muhammad Aksam Iftikhar, Amanda Shacklett, and Christos Davatzikos. 2017. A review on neuroimaging-based classification studies and associated feature extraction methods for Alzheimer's disease and its prodromal stages. *NeuroImage* 155 (2017), 530–548.

[139] Alessandra Retico, Paolo Bosco, Piergiorgio Cerello, Elisa Fiorina, Andrea Chincarini, and Maria Evelina Fantacci. 2015. Predictive Models Based on Support Vector Machines: Whole-Brain versus Regional Analysis of Structural MRI in the Alzheimer's Disease. *Journal of Neuroimaging* 25, 4 (2015), 552–563.

[140] B Richhariya and Muhammad Tanveer. 2018. EEG signal classification using universum support vector machine. *Expert Systems with Applications* 106 (2018), 169–182.

[141] Bharat Richhariya and Muhammad Tanveer. 2018. A robust fuzzy least squares twin support vector machine for class imbalance learning. *Applied Soft Computing* 71 (2018), 418–432.

[142] Pedro Rodrigues and João Paulo Teixeira. 2011. Artificial neural networks in the discrimination of alzheimer's disease. In *International Conference on ENTERprise Information Systems*. Springer, 272–281.

[143] Diego Salas-Gonzalez, Juan Manuel Górriz, Javier Ramírez, Míriam López, Ignacio Álvarez, Fermín Segovia, and Carlos García Puntonet. 2008. Computer aided diagnosis of Alzheimer's disease using support vector machines and classification trees. In *International Conference on Neural Information Processing*. Springer, 418–425.

[144] Ziad Sankari and Hojjat Adeli. 2011. Probabilistic neural networks for diagnosis of Alzheimer's disease using conventional and wavelet coherence. *Journal of neuroscience methods* 197, 1 (2011), 165–170.

[145] Saman Sarraf, Danielle D DeSouza, John Anderson, Ghassem Tofighi, et al. 2017. DeepAD: Alzheimer's disease classification via deep convolutional neural networks using MRI and fMRI. *BioRxiv* (2017), 070441.

[146] Alexandre Savio, Maite García-Sebastián, Carmen Hernández, Manuel Graña, and Jorge Villanúa. 2009. Classification results of artificial neural networks for alzheimer's disease detection. In *International Conference on Intelligent Data Engineering and Automated Learning*. Springer, 641–648.

[147] Daniel Schmitter, Alexis Roche, Bénédicte Maréchal, Delphine Ribes, Ahmed Abdulkadir, Meritxell Bach-Cuadra, Alessandro Daducci, Cristina Granziera, Stefan Klöppel, Philippe Maeder, et al. 2015. An evaluation of volume-based morphometry for prediction of mild cognitive impairment and Alzheimer's disease. *NeuroImage: Clinical* 7 (2015), 7–17.

[148] Tijn M Schouten, Marisa Koini, Frank de Vos, Stephan Seiler, Jeroen van der Grond, Anita Lechner, Anne Hafkemeijer, Christiane Möller, Reinhold Schmidt, Mark de Rooij, et al. 2016. Combining anatomical, diffusion, and resting state functional magnetic resonance imaging for individual classification of mild and moderate Alzheimer's disease. *NeuroImage: Clinical* 11 (2016), 46–51.

[149] Fermín Segovia, JM Górriz, Javier Ramírez, Diego Salas-Gonzalez, and Ignacio Álvarez. 2013. Early diagnosis of Alzheimer's disease based on partial least squares and support vector machine. *Expert Systems with Applications* 40, 2 (2013), 677–683.

[150] F Segovia, JM Górriz, J Ramírez, D Salas-González, I Álvarez, M López, R Chaves, and P Padilla. 2010. Classification of functional brain images using a GMM-based multi-variate approach. *Neuroscience Letters* 474, 1 (2010), 58–62.

[151] Dinggang Shen, Guorong Wu, and Heung-Il Suk. 2017. Deep learning in medical image analysis. *Annual review of biomedical engineering* 19 (2017), 221–248.

[152] Jinhua Sheng, Bocheng Wang, Qiao Zhang, Qingqiang Liu, Yangjie Ma, Weixiang Liu, Meiling Shao, and Bin Chen. 2019. A novel joint HCPMMP method for automatically classifying AlzheimerâĂŹs and different stage MCI patients. *Behavioural brain research* (2019).

[153] Jun Shi, Xiao Zheng, Yan Li, Qi Zhang, and Shihui Ying. 2018. Multimodal neuroimaging feature learning with multimodal stacked deep polynomial networks for diagnosis of Alzheimer's disease. *IEEE journal of biomedical and health informatics* 22, 1 (2018), 173–183.

[154] Zhenghao Shi, Lifeng He, Kenji Suzuki, Tsuyoshi Nakamura, and Hidenori Itoh. 2009. Survey on neural networks used for medical image processing. *International journal of computational science* 3, 1 (2009), 86.

[155] Nikhil Singh, P Thomas Fletcher, J Samuel Preston, Richard D King, JS Marron, Michael W Weiner, Sarang Joshi, and Alzheimer's Disease Neuroimaging Initiative (ADNI). 2014. Quantifying anatomical shape variations in neurological disorders. *Medical image analysis* 18, 3 (2014), 616–633.

[156] Simeon Spasov, Luca Passamonti, Andrea Duggento, Pietro Lio, Nicola Toschi, Alzheimer's Disease Neuroimaging Initiative, et al. 2019. A parameter-efficient deep learning approach to predict conversion from mild cognitive impairment to Alzheimer's disease. *NeuroImage* 189 (2019), 276–287.

[157] Gabriela Spulber, Andrew Simmons, J-S Muehlboeck, Patrizia Mecocci, Bruno Vellas, Magda Tsolaki, Iwona Kłoszewska, Hilkka Soininen, Christian Spenger, Simon Lovestone, et al. 2013. An MRI-based index to measure the severity of Alzheimer's disease-like structural pattern in subjects with mild cognitive impairment. *Journal of internal medicine* 273, 4 (2013), 396–409.

[158] Jonathan Stoeckel and Glenn Fung. 2005. SVM feature selection for classification of SPECT images of Alzheimer's disease using spatial information. In *Fifth IEEE International Conference on Data Mining (ICDM'05)*. IEEE, 8–pp.

[159] Heung-Il Suk, Seong-Whan Lee, Dinggang Shen, and Alzheimer's Disease Neuroimaging Initiative. 2014. Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis. *NeuroImage* 101 (2014), 569–582.

[160] Heung-Il Suk, Seong-Whan Lee, Dinggang Shen, and Alzheimer's Disease Neuroimaging Initiative. 2015. Latent feature representation with stacked auto-encoder for AD/MCI diagnosis. *Brain Structure and Function* 220, 2 (2015), 841–859.

[161] Heung-Il Suk, Seong-Whan Lee, Dinggang Shen, and Alzheimer's Disease Neuroimaging Initiative. 2017. Deep ensemble learning of sparse regression models for brain disease diagnosis. *Medical image analysis* 37 (2017), 101–113.

[162] Heung-Il Suk and Dinggang Shen. 2013. Deep learning-based feature representation for AD/MCI classification. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 583–590.

[163] Heung-Il Suk and Dinggang Shen. 2014. Subclass-based multi-task learning for Alzheimer's disease diagnosis. *Frontiers in aging neuroscience* 6 (2014), 168.

[164] Heung-Il Suk and Dinggang Shen. 2016. Deep ensemble sparse regression network for Alzheimer's disease diagnosis. In *International Workshop on Machine Learning in Medical Imaging*. Springer, 113–121.

[165] Heung-Il Suk, Chong-Yaw Wee, Seong-Whan Lee, and Dinggang Shen. 2016. State-space model with deep learning for functional dynamics estimation in resting-state fMRI. *NeuroImage* 129 (2016), 292–307.

[166] Zhuo Sun, Yuchuan Qiao, Boudewijn PF Lelieveldt, Marius Staring, Alzheimer's Disease NeuroImaging Initiative, et al. 2018. Integrating spatial-anatomical regularization and structure sparsity into SVM: Improving interpretation of Alzheimer's disease classification. *NeuroImage* 178 (2018), 445–460.

[167] Sabina Tangaro, Annarita Fanizzi, Nicola Amoroso, Roberto Bellotti, Alzheimer's Disease Neuroimaging Initiative, et al. 2017. A fuzzy-based system reveals Alzheimer's disease onset in subjects with Mild Cognitive Impairment. *Physica Medica* 38 (2017), 36–44.

[168] Mohammad Tanveer, Mohammad Asif Khan, and Shen-Shyang Ho. 2016. Robust energy-based least squares twin support vector machines. *Applied Intelligence* 45, 1 (2016), 174–186.

[169] M Termenon and Manuel Graña. 2012. A two stage sequential ensemble applied to the classification of Alzheimer's disease based on mri features. *Neural Processing Letters* 35, 1 (2012), 1–12.

[170] M Termenon, Manuel Grana, A Besga, J Echeveste, and A Gonzalez-Pinto. 2013. Lattice independent component analysis feature selection on diffusion weighted imaging for Alzheimer's disease classification. *Neurocomputing* 114 (2013), 132–141.

[171] Yingjie Tian and Zhiquan Qi. 2014. Review on: twin support vector machines. *Annals of Data Science* 1, 2 (2014), 253–277.

[172] Tong Tong, Katherine Gray, Qinquan Gao, Liang Chen, Daniel Rueckert, Alzheimer's Disease Neuroimaging Initiative, et al. 2017. Multi-modal classification of Alzheimer's disease using nonlinear graph fusion. *Pattern recognition* 63 (2017), 171–181.

[173] Rick van Veen, L Talavera Martinez, RV Kogan, SK Meles, Deborah Mudali, Jos BTM Roerdink, F Massa, M Grazzini, Jose A Obeso, Maria C Rodriguez-Oroz, et al. 2018. Machine Learning Based Analysis of FDG-PET Image Data for the Diagnosis of Neurodegenerative Diseases.. In *APPIS*. 280–289.

[174] Dallas P Veitch, Michael W Weiner, Paul S Aisen, Laurel A Beckett, Nigel J Cairns, Robert C Green, Danielle Harvey, Clifford R Jack Jr, William Jagust, John C Morris, et al. 2018. Understanding disease progression and improving Alzheimer's disease clinical trials: Recent highlights from the Alzheimer's Disease Neuroimaging Initiative. *Alzheimer's & Dementia* (2018).

[175] Prashanthi Vemuri, Jeffrey L Gunter, Matthew L Senjem, Jennifer L Whitwell, Kejal Kantarci, David S Knopman, Bradley F Boeve, Ronald C Petersen, and Clifford R Jack Jr. 2008. Alzheimer's disease diagnosis in individual subjects using structural MR images: validation studies. *Neuroimage* 39, 3 (2008), 1186–1197.

[176] Prashanthi Vemuri and Clifford R Jack. 2010. Role of structural MRI in Alzheimer's disease. *Alzheimer's research & therapy* 2, 4 (2010), 23.

[177] Hongfei Wang, Yanyan Shen, Shuqiang Wang, Tengfei Xiao, Liming Deng, Xiangyu Wang, and Xinyan Zhao. 2019. Ensemble of 3D densely connected convolutional network for diagnosis of mild cognitive impairment and Alzheimer's disease. *Neurocomputing* 333 (2019), 145–156.

[178] Kun Wang, Meng Liang, Liang Wang, Lixia Tian, Xinqing Zhang, Kuncheng Li, and Tianzi Jiang. 2007. Altered functional connectivity in early Alzheimer's disease: A resting-state fMRI study. *Human brain mapping* 28, 10 (2007), 967–978.

[179] Shuihua Wang, Yudong Zhang, Zhengchao Dong, Sidan Du, Genlin Ji, Jie Yan, Jiquan Yang, Qiong Wang, Chunmei Feng, and Preetha Phillips. 2015. Feed-forward neural network optimized by hybridization of PSO and ABC for abnormal brain detection. *International Journal of Imaging Systems and Technology* 25, 2 (2015), 153–164.

[180] Tingyan Wang, Robin G Qiu, and Ming Yu. 2018. Predictive Modeling of the Progression of Alzheimer‘s Disease with Recurrent Neural Networks. *Scientific reports* 8 (2018).

[181] Michael W Weiner, Dallas P Veitch, Paul S Aisen, Laurel A Beckett, Nigel J Cairns, Robert C Green, Danielle Harvey, Clifford R Jack Jr, William Jagust, John C Morris, et al. 2017. Recent publications from the Alzheimer's Disease Neuroimaging Initiative: Reviewing progress toward improved AD clinical trials. *Alzheimer's & Dementia* 13, 4 (2017), e1–e85.

[182] Eric Westman, Andrew Simmons, J-Sebastian Muehlboeck, Patrizia Mecocci, Bruno Vellas, Magda Tsolaki, Iwona Kłoszewska, Hilkka Soininen, Michael W Weiner, Simon Lovestone, et al. 2011. AddNeuroMed and ADNI: similar patterns of Alzheimer's atrophy and automated MRI classification accuracy in Europe and North America. *Neuroimage* 58, 3 (2011), 818–828.

[183] Yitian Xu, Xianli Pan, Zhijian Zhou, Zhiji Yang, and Yuqun Zhang. 2015. Structural least square twin support vector machine for classification. *Applied Intelligence* 42, 3 (2015), 527–536.

[184] Shih-Ting Yang, Jiann-Der Lee, Tzyh-Chyang Chang, Chung-Hsien Huang, Jiun-Jie Wang, Wen-Chuin Hsu, Hsiao-Lung Chan, Yau-Yau Wai, and Kuan-Yi Li. 2013. Discrimination between Alzheimer‘s disease and mild cognitive impairment using SOM and PSO-SVM. *Computational and mathematical methods in medicine* 2013 (2013).

[185] Shih-Ting Yang, Jiann-Der Lee, Chung-Hsien Huang, Jiun-Jie Wang, Wen-Chuin Hsu, and Yau-Yau Wai. 2010. Computer-aided diagnosis of Alzheimer‘s disease using multiple features with artificial neural network. In *Pacific Rim International Conference on Artificial Intelligence*. Springer, 699–705.

[186] Jieping Ye, Kewei Chen, Teresa Wu, Jing Li, Zheng Zhao, Rinkal Patel, Min Bae, Ravi Janardan, Huan Liu, Gene Alexander, et al. 2008. Heterogeneous data fusion for Alzheimer‘s disease study. In *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 1025–1033.

[187] Nianyin Zeng, Hong Qiu, Zidong Wang, Weibo Liu, Hong Zhang, and Yurong Li. 2018. A new switching-delayed-PSO-based optimized SVM algorithm for diagnosis of Alzheimer‘s disease. *Neurocomputing* 320 (2018), 195–202.

[188] Daoqiang Zhang and Dinggang Shen. 2011. Semi-supervised multimodal classification of Alzheimer‘s disease. In *2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*. IEEE, 1628–1631.

[189] Daqing Zhang, Jianfeng Xiao, Nannan Zhou, Mingyue Zheng, Xiaomin Luo, Hualiang Jiang, and Kaixian Chen. 2015. A genetic algorithm based support vector machine model for blood-brain barrier penetration prediction. *BioMed research international* 2015 (2015).

[190] Jie Zhang, Cynthia Stonnington, Qingyang Li, Jie Shi, Robert J Bauer, Boris A Gutman, Kewei Chen, Eric M Reiman, Paul M Thompson, Jieping Ye, et al. 2016. Applying sparse coding to surface multivariate tensor-based morphometry to predict future cognitive decline. In *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*. IEEE, 646–650.

[191] Yudong Zhang, Zhengchao Dong, Preetha Phillips, Shuihua Wang, Genlin Ji, Jiquan Yang, and Ti-Fei Yuan. 2015. Detection of subjects and brain regions related to Alzheimer's disease using 3D MRI scans based on eigenbrain and machine learning. *Frontiers in computational neuroscience* 9 (2015), 66.

[192] Yudong Zhang and Shuihua Wang. 2015. Detection of AlzheimerâĂŹs disease by displacement field and machine learning. *PeerJ* 3 (2015), e1251.

[193] Yudong Zhang, Shuihua Wang, Preetha Phillips, Zhengchao Dong, Genlin Ji, and Jiquan Yang. 2015. Detection of Alzheimer's disease and mild cognitive impairment based on structural volumetric MR images using 3D-DWT and WTA-KSVM trained by PSOTVAC. *Biomedical Signal Processing and Control* 21 (2015), 58–73.

[194] Ying-Teng Zhang and Shen-Quan Liu. 2018. Individual identification using multi-metric of DTI in Alzheimer‘s disease and mild cognitive impairment. *Chinese Physics B* 27, 8 (2018), 088702.

[195] Weihao Zheng, Zhijun Yao, Yuanwei Xie, Jin Fan, and Bin Hu. 2018. Identification of AlzheimerâĂŹs disease and mild cognitive impairment using networks constructed based on multiple morphological brain features. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* 3, 10 (2018), 887–897.

[196] Xiao Zheng, Jun Shi, Yan Li, Xiao Liu, and Qi Zhang. 2016. Multi-modality stacked deep polynomial network based feature learning for alzheimer's disease diagnosis. In *2016 IEEE 13th international symposium on biomedical imaging (ISBI)*. IEEE, 851–854.

[197] Xiao Zheng, Jun Shi, Qi Zhang, Shihui Ying, and Yan Li. 2017. Improving MRI-based diagnosis of Alzheimer‘s disease via an ensemble privileged information learning algorithm. In *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*. IEEE, 456–459.

[198] Ke Zhou, Wenguang He, Yonghui Xu, Gangqiang Xiong, and Jie Cai. 2018. Feature selection and transfer learning for Alzheimer‘s disease clinical diagnosis. *Applied Sciences* 8, 8 (2018), 1372.

[199] Xiaofeng Zhu, Heung-Il Suk, Seong-Whan Lee, and Dinggang Shen. 2016. Subspace regularized sparse multitask learning for multiclass neurodegenerative disease identification. *IEEE Transactions on Biomedical Engineering* 63, 3 (2016), 607–618.

[200] Yingying Zhu, Xiaofeng Zhu, Minjeong Kim, Dinggang Shen, and Guorong Wu. 2016. Early diagnosis of Alzheimer‘s disease by joint feature selection and classification on temporally structured support vector machine. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 264–272.