BIOCYBERNETICS AND BIOMEDICALENGINEERING 41 (2021) 456–473

Available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/bbe

Review Article

Application of Artificial Intelligence techniques for the detection of Alzheimer's disease using structural MRI images

Xinxing Zhao [a], *Candice Ke En Ang* [b], *U. Rajendra Acharya* [c,d,e], *Kang Hao Cheong* [a,*]

[a] *Science, Mathematics and Technology Cluster, Singapore University of Technology and Design, 8 Somapah Road, S487372, Singapore*
[b] *MOH Holdings Pte Ltd, 1 Maritime Square, S099253, Singapore*
[c] *Department of Electronics and Computer Engineering, Ngee Ann Polytechnic, Clementi 599489, Singapore*
[d] *Department of Biomedical Engineering, School of Science and Technology, SUSS University, Clementi 599491, Singapore*
[e] *International Research Organization for Advanced Science and Technology (IROAST), Kumamoto University, Kumamoto, Japan*

ARTICLE INFO

Article history:
Received 14 October 2020
Received in revised form
6 February 2021
Accepted 22 February 2021
Available online 5 April 2021

Keywords:
Artificial Intelligence
Alzheimer's disease
Machine Learning
Deep Learning
Computer-aided diagnosis
Structural MRI

ABSTRACT

Alzheimer's disease (AD) is an irreversible, progressive brain disorder that slowly destroys memory and thinking skills. It is one of the leading types of dementia for persons aged above 65 worldwide. In order to achieve accurate and timely diagnosis, and for detection of AD in its early stages, numerous Artificial Intelligence (AI) based Computer-aided Diagnostic (CAD) approaches have been proposed using data from brain imaging. In this paper, we review the recent application of AI based CAD systems on AD and its stages, with a particular focus on the use of structural MRI due to its cost effectiveness and lack of ionizing radiation. We will review important factors of different AI techniques pertinent to AD, summarize contributions from different research groups, critically discuss challenges involved and propose directions for future research. Ultimately, it would be ideal for development of a diagnostic framework that could be applicable to not only AD, but to different types of dementia as well in the future.

1. Introduction

Dementia is the loss of cognitive functioning—thinking, remembering, and reasoning—and behavioral abilities to an extent that it interferes with a person's daily life and activities. These functions include memory, language skills, visual perception, problem solving, self-management, and the ability to focus and pay attention [1]. Alzheimer's disease is an irreversible, progressive brain disorder that slowly destroys memory and thinking skills [2]. AD accounts for most [3] dementia cases, and it tends to occur in older persons, with age being one of the risk factors. According to a report from

* Corresponding author at: Science, Mathematics and Technology Cluster, Singapore University of Technology and Design, Block 1, Level 4, Room 08, 8 Somapah Road, S487372, Singapore.
E-mail address: kanghao_cheong@sutd.edu.sg (K.H. Cheong).
https://doi.org/10.1016/j.bbe.2021.02.006

the Alzheimer's Association [4], in 2019 an estimated 5.8 million Americans of all ages are living with AD; Among them about 200,000 people are aged under 65, 0.9 million aged 65 to 74, 2.6 million aged 75 to 84, 2.1 million aged 85+. However, the pathophysiology of AD is still under research and other important risk factors in the development of AD include family history and certain genes in one's genome [5]. From the 2019 World Alzheimer Report [6], there are about 50 million people suffering from dementia globally, and the number is likely to triple by 2050. Although at current stage, the final and confirmed diagnosis of Alzheimer's disease can only be achieved through post-mortem identification of the histopathological hallmarks that actually define AD, such as neurofibrillary tangles and/or abnormal plaque deposits on the brain, recent research shows that AD stages can be traceable through some indicators, and therefore possible ways for early detection and intervention. According to Sperling et al. [7], there are three different stages for the development of AD. Firstly, in the pre-clinical AD stage, patients are clinically asymptomatic, however measurable changes are present in the brain in the form of cerebral amyloidosis. Further research is required to identify specific biomarkers that can be quantified, so as to facilitate detection and early intervention. In the next stage, Mild Cognitive Impairment (MCI) develops. There is evidence of brain changes and subtle symptoms that may only be noticeable to family members of affected patients. Of note, not all patients with MCI will progress to develop dementia; some are able to revert to normal [8], while others remain stable. Despite the above, it is more likely for MCI patients with memory problems to develop dementia than individuals without MCI [9]. In a study by Mitchell et al. [10], 32% of MCI subjects developed AD within 5 years. The prediction of which subjects are more likely to develop AD from MCI is a major goal of current AD research. In the third stage, patients develop dementia due to AD, with symptoms significant to impair one's ability to function in daily life. At this stage, there is often overt evidence of brain changes, such as reduced hippocampal volume or cortical thickness changes. The severity of symptoms and rate of development from pre-clinical stages to AD differ amongst subjects.

Two common tools used for cognition assessment are the Mini-Mental State Examination (MMSE) [11] and the Clinical Dementia Rating (CDR) [12]. These tools assess different domains of one's cognition through simple questions. Apart from these clinical tools, other methods for detection and diagnosis of AD include medical imaging, Electroencephalography (EEG) signals, quantifiable proteins in cerebrospinal fluid, blood and urine. Medical imaging is a very important tool in the evaluation of patients with brain pathology [13]. There are mainly three types of imaging technologies used, namely structural imaging, including Magnetic Resonance Imaging (MRI) and Computed Tomography (CT) scans; functional imaging, including Positron Emission Tomography (PET) and functional MRI (fMRI) scans; and molecular imaging, including PET, fMRI and single photon emission computed tomography (SPECT). Further details on the different types of imaging are presented in Table 1, more information about modality types [14] is in the following: CT scans use X-rays to show the two dimensional structure of the brain, with details such as blood perfusion. MRI scans use a strong magnetic field and radio waves to create pictures of the tissues and other structures inside the brain. fMRI imaging records the movement of blood flow, can indicate the active or functioning part of the brain, in response to the patient performing a given task. PET is a nuclear medicine imaging technique, which produces a three-dimensional image of functional processes in the brain. However, patients need to receive a small injection of radio-active material into the bloodstream. SPECT imaging also needs radio-active material injection to the patients, it records the signals from gamma rays with two or more synchronised gamma cameras, and the multiple 2-D images are computed to 3-D. It can trace blood flow within the brain that identifies where metabolic activity is occurring, enabling assessment of brain functions.

Among the different modalities, MRI is an efficient way that provides excellent definition of the different tissue types without the use of ionizing radiation and is of moderate cost [15]. It can be used to reveal anatomical differences between AD patients and healthy individuals, for evaluation of different stages of AD such as the development of MCI [16]. Furthermore, it can be used for detection of MCI to AD conversion, for there exists a close temporal link between MRI abnormalities and the onset of the cognitive impairment [17]. It can also be used to differentiate different types of dementia, such as between AD and frontotemporal dementia which may have similar clinical presentations [18].

AI and ML have been used for many medical and healthcare purposes, such as disease detection [19] [20], health services [21] [22] [23] [24] and industry applications [25]. In this review, we mainly focus on the use of both traditional Machine Learning and more advanced Deep Learning techniques that make use of structural MRI due to its vast use in current literature [26]. With the gradual acquisition standardization process and more advanced analysis techniques, structural MRI has the potential to contribute more to the development of robust algorithms for the automated assessment of AD and its stages.

The high volume and complexity of brain imaging for analysis puts a huge workload for clinicians in the diagnosis and subsequent management [27]. As a result, the path to diagnosis is often time-consuming and error-prone, as such diagnoses are based upon the reporting physician's knowledge and experience which may be subjective and occasionally unreliable. Quantitative measures are therefore essential [28] and would be ideal. With the advancement of image processing technologies, Computer-aided Diagnostic (CAD) tools can provide targeted solutions. They provide a complementary way of analysing and interpreting medical images, aiding in enhancing diagnostic accuracy and reducing the workload of radiologists with reproducible results. The joint efforts between clinicians and CAD tools can amplify substantially the diagnostic capabilities as shown in Fig. 1.

The detection and diagnosis of AD and its stage, with CAD methods, is classically an image-based pattern recognition problem. This consists of a few sequential steps: (1) preprocessing of images, followed by (2) feature(s) extraction, (3) dimensionality reduction, (4) feature optimization, and finally, (5) classification. Of these, significant steps involving

Table 1 – Imaging Technologies used in Research of AD.

Imaging Techniques	Description	Modality Types
Structural Imaging	Gives volume, shape, position of brain tissue information	CT & MRI
Functional Imaging	Illustrates the activity of various brain regions via cell metabolism of sugars or oxygen	PET & fMRI
Molecular Imaging	Uses targeted radiotracers to detect changes that linked to diseases	PET, fMRI & SPECT.

Machine Learning include steps 2, 3 and 5. During the training stage, characteristic features of MRI images from different groups (normal, MCI and AD patients) are derived based on quantitative calculations. These features will then be selected and/or combined to reduce the dimensionality, and a classifier will be trained with certain supervised learning techniques. The general processes are shown in Fig. 2.

On the other hand, the more advanced Deep Learning (Fig. 3) approaches based on layered artificial neural networks (ANN) usually combine the steps of feature extraction, dimensionality reduction and classification together automatically [30,31], with supervised or unsupervised learning techniques. As Deep Learning models attempt to discover the hidden representations and links among various regions of the scanned images, their performance in identifying AD patterns are generally better than the traditional Machine Learning techniques. With advanced processing powers, and more available imaging samples, Deep Learning is gaining popularity in the analysis of medical images.

2. Machine Learning approaches for diagnosis of AD

Before going into detail of the available AD research, we present information about some public datasets and the general performance measurements used in the literature.

2.1. *Open established medical imaging sets for ageing analysis*

Datasets are widely available and are chosen at the discretion of the different research groups. Nevertheless, most tend to develop algorithms based upon established medical imaging databases for added research value. Two most well-known brain imaging databases are presented in Table 2.

Alzheimer's Disease Neuroimaging Initiative (ADNI): It is an ongoing, multi-site comprehensive (ADNI 1 2 3 and Go) study, which tracks the progression of AD and its stages in the human brain with clinical, imaging, genetic and biospecimen biomarkers (http://adni.loni.usc.edu). Another well established database is the Open Access Series of Imaging Studies (OASIS); for more information, refer to https://www.oasis-brains.org.

2.2. *Performance measurements*

In CAD systems for AD research, different research groups make use of different algorithms with different modalities. For purposes of comparison, various performance metrics were used by various research, and some commonly used measurements are accuracy, sensitivity, specificity, precision and f1-score.

2.3. *Conventional Machine Learning approaches for AD diagnosis*

Clinically, the ability to diagnose AD at early or asymptomatic stages brings about great impact, as it allows for early intervention. This would benefit not only the patient but also their

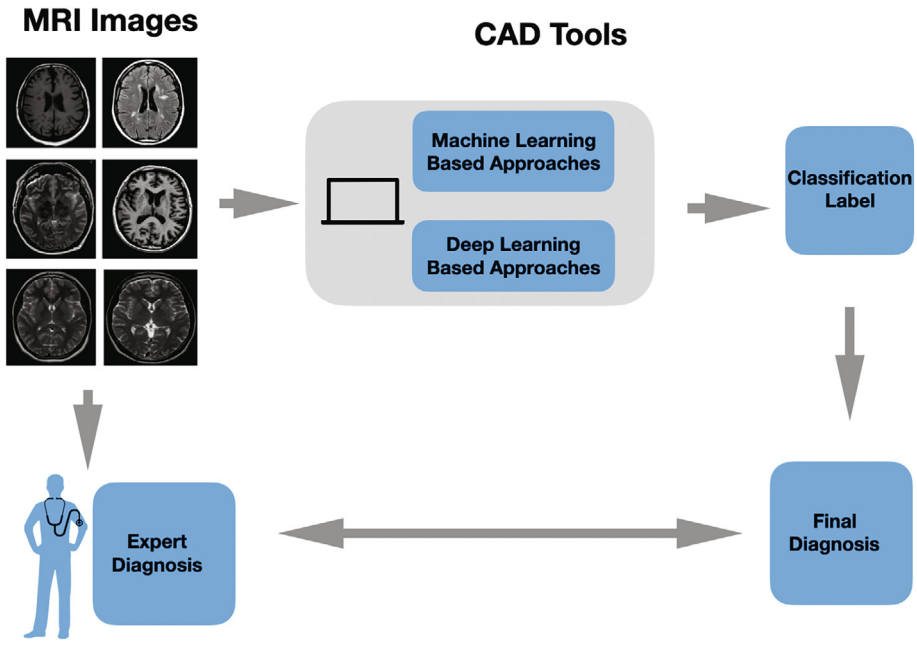

Fig. 1 – Use of CAD techniques as complementary tools for AD diagnosis.

loved ones, who often bear great caregiver stress when dealing with AD patients. This could also bring about reduced costs in the long term in terms of treatment. Therefore, in CAD systems research for AD, many focus on differentiating individuals in the early stages of AD from normal controls (NC) [55,65,78]. The classification of AD vs. MCI vs. NC [79,62,80] is also another popular research direction.

Another research focus is in identifying these MCI subjects who are more likely to develop into AD [63,73,41]. As mentioned previously, not all MCI patients will eventually develop dementia. Those MCI subjects who develop into AD over the duration of the study are categorized as MCI converters (cMCI) or progressive MCI (pMCI), while those who do not convert are categorized as MCI non-converters (ncMCI) or stable MCI (sMCI). Other researchers attempt to predict the duration required for subjects with MCI to convert to AD [47,37], some research predicted long term conversions, for example, Popuri et al. [76] predicted up to 7 years of time-to-conversion for the classification on stable versus progressive MCI groups, achieved a good performance with an AUC 0.73. These research can aid in early intervention [37,52] with benefits as described earlier, hence are highly valuable.

In a classical MRI-based system for automated AD diagnosis, there are usually two main components: feature extraction from MRI images and classifier(s) that uses these features. There are many popular classifiers, such as Support Vector Machines (SVM), Random Forests, Logistic Regression, Discriminant Analysis, the Bayesian classifiers. Some of them are shown in Table 3, and many more are used in research shown in Tables 4–6.

With regards to feature extraction techniques [81,58], there are mainly three types [82], namely, the voxel-based [58], Region Of Interest (ROI) based [52], and patch-based approach [57]. Examples of features include the shape [43] [83], texture [84,85], or volume [86,62] that can be measured from brain imaging. These features can be derived from the various brain regions, or even from certain patterns derived. For example, features can be derived from gray matter [39,5], the similarities among the images [47], the topologies [46], the extrema in the original image or acquired from the first higher-order derivative of the original image [42], the correlations derived from ROI [57,52], or even visual features related to the hippocampal area [53].

Studies have demonstrated that hippocampal atrophy can be used as a good indicator of disease as AD progresses [87,43,88]. Research groups have attempted to assess in vivo hippocampal atrophy, such as Zhou et al. [54], who made use of hippocampal volumetric variables that were calculated using Freesurfer [89] from MRI scans. These volumetric variables were ranked and selected, and combined with MMSE

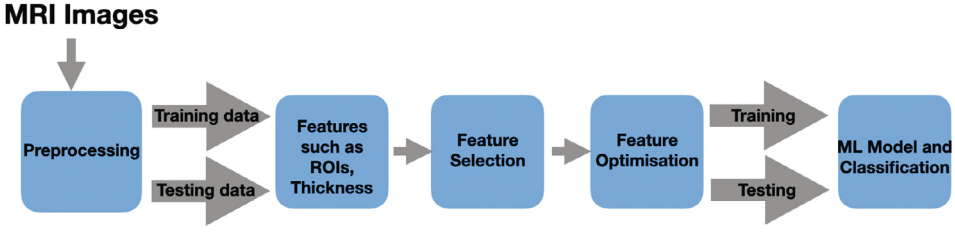

Fig. 2 – General structure of Machine Learning based diagnostic approaches, adapted from [29].

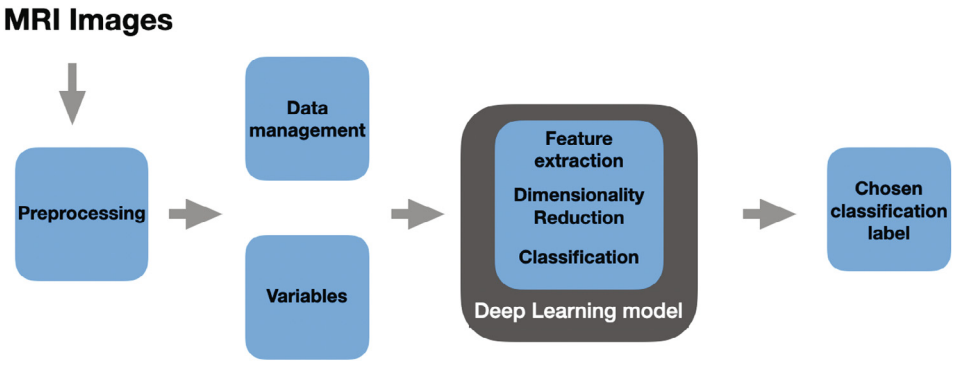

Fig. 3 – General structure of Deep Learning based diagnostic approaches.

scores as features. It showed that with the help of MMSE scores, discriminatory power for different stages of AD increased by over 10%. For AD vs NC, a classification accuracy of 92.4% (sensitivity: 84.0%; specificity: 96.1%) was achieved. For amnestic MCI vs NC, the accuracy was 74.9% (sensitivity: 61.1%; specificity: 83.4%), and for non-amnestic MCI vs NC, the accuracy was 74.1% (sensitivity: 55.2%; specificity: 82.3%). Apart from the MMSE score, some studies showed that the combination of age [90], sex and education level [91] of the subjects can serve as discriminating factors to form better aggregated biomarkers. The presence of apoE [51,92] genotype as a means of stratification can also help to boost the classification performance.

In-vivo cortical thickness measurements performed via MRI is another important technique for AD diagnosis. Cortical thinning patterns in the cerebrum can be used to detect MCI to AD conversion. In Eskildsen et al. [37], atrophy patterns of cortical thickness were explored; by avoiding double dipping, a more accurate predictive power (about 9%–14% higher accuracy as compared with [87,49,47]) was obtained. While compared to sMCI, an accuracy of 80.9% (sensitivity 78.7%, specificity 82.8%) was achieved with pMCI6 (converted to AD in 6 months), 74.5% (sensitivity 75.2% and specificity 73.9%) with pMCI12 (converted to AD in 12 months), and 77.3% (sensitivity 69.0%, specificity 79.1%) with pMCI36 (converted to AD in 36 months). For MCI vs NC, the performance achieved an accuracy of 86.7% (sensitivity 80.4%, specificity 92.0%). Pachauri et al. [46] provided a non-parametric method which characterized important topological features with cortical thickness values. Cho et al. [49] provided a spatial frequency representation of cortical thickness data to classify subjects for AD based on incremental learning. Cortical thickness data were mapped onto a spatial frequency domain with manifold harmonic transformed from the surface of a cortex. They used PCA and LDA as the classifiers. They achieved higher sensitivity and specificity for discrimination of NC vs AD (82%/93%), NC vs cMCI (66%/89%), as well as ncMCI vs cMCI (63%/76%), as compared to all the 10 methods in Cuingnet et al. [48], including three thickness based methods.

Some research groups focused on combinations of different methods, and with merged biomarkers, and some groups used multivariate data analysis [45,51] to study many variables simultaneously and observe inherent patterns in the data to increase discrimination powers in classifying AD, MCI and NC. Wolz et al. [47] used multi-method analysis (such as hippocampal volume based, cortical thickness based, tensor-based morphometry) in the early diagnosis of AD. They attempted to find and combine different features with multiple classifiers. LDA achieved the highest discrimination power for NC vs. AD, with accuracy of 89% (sensitivity 93%, specificity 85%), while for sMCI vs pMCI, accuracy of 68% (sensitivity 67%, specificity 69%) achieved. Gupta et al. [40] proposed a method combining three different features extracted from structural MRI using voxel-based morphometry, hippocampal volume, and cortical and subcortical segmented region techniques. For AD vs. HC, the accuracy was 93.06 % (sensitivity 87.87%, specificity 95.58%) with SVM as the classifier. Ahmed et al. [93] used circular harmonic functions on hippocampus and posterior cingulate cortex to build descriptors. For the challenging classification task (AD vs MCI), they reached an accuracy of 62.07%, (specificity of 75.15%, sensitivity of 49.02%). Westman et al. [50] used Freesurfer to get regional subcortical and cortical thickness measures, and with multivariate analysis for 60 variables, including MRI and CSF measures. The MRI measures alone achieved a reasonable classification power, for AD vs. MCI

Table 2 – Open Source Imaging Databases for AD Research.

Database Name	Subjects	Modalities and Data Types
ADNI	ADNI-1:400 MCI, 200 early AD, 200 elderly NCs; ADNI-2: 150 NCs, 100 EMCI, 150 LMCI, and 150 mild AD; ADNI-3: 430–880 NCs, 425–835 mild MCI, 215–335 mild AD; ADNI-Go: 200 EMCI	PET, MRI, genetic data, clinical and cognitive assessments
ASIS	OASIS-1: 416 in all, 100 mild to moderate AD, 20 NCs; OASIS-2: 72 NCs, 64 AD; OASIS-3: 609 NCs, 489 AD at different stages	MRI, PET

Table 3 – Conventional Classifiers in CAD Systems.

Classifier	Description	Studies
Support Vector Machines (SVM)	For classification and regression, supervised, multivariate learning methods	[32,33]
Artificial Neural Networks	Inspired by biological neural networks, requires less statistical training, using nodes and layers for non-linear fitting. Performs well for tasks that are not easily solved with ordinary rule-based programming	[55,68,84]
Logistic Regression (LR)	To describe data and explain the relationship between one dependent binary variable and one or more independent variables	[34,35]
Discriminant Analysis	Predicting categories among groups based on continuous variables	[36,37]
Bayesian Classifier	Bayes' theorem based and with strong independent assumptions; can outperform other classifiers	[38,39]
Random Forests (RF)	An ensemble learning method for classification, by constructing a multitude of decision trees at training time and outputting the average of them.	[40,41]
K-Nearest Neighbour (KNN)	It stores all available labels and classifies new labels based on a similarity measure, belongs to the family of instance-based, competitive learning, and lazy learning algorithms.	[40,42]

with accuracy of 87.0% (sensitivity 83.3%, specificity 90.1%), for MCI vs. NC with accuracy of 71.8% (sensitivity %66.7, specificity 79.3%) and for cMCI vs sMCI with accuracy of 65.4% (sensitivity 65.4%, specificity 65.4%). Cortical and subcortical regions were segmented, cortical thickness and curvature were calculated in [38], ROIs with its volume, thickness and mean cortical curvature were used as features, SVM and Naïve Bayes classifiers were used. For mild AD vs NC classification, with both subcortical and cortical volumes feature vectors used with the SVM classifier, 86% accuracy (82% sensitivity, 90% specificity) was achieved, while for mild AD vs MCI, 75% accuracy (73% sensitivity, 77% specificity) was acheived, which was slightly better than just based on subcortical volumes with 74% accuracy (67% sensitivity, 80% specificity). In Coupé et al. [36], a simultaneous patch-based segmentation was presented for AD vs NC classification. It segmented and graded anatomical structures of hippocampus and entorhinal cortex at same time, and used hippocampal volume, grade (defined as the similarity of the patch surrounding the voxel), entorhinal volume and grade, and their combinations as biomarkers to find atrophic patterns. With LDA and QDA as classifiers, it was revealed that hippocampal based measures have more discriminating powers than entorhinal cortex based measures, with grading being more powerful than volume-based measures. The achieved accuracy is 90% (sensitivity 86%, specificity 94%), and higher classification accuracy of 93% with age factor included. Lama et al. [71] performed a greedy score-based feature selection technique that chose important feature vectors (cortical thickness, surface area, folding indices, mean curvature indices and volume); and with RELM as the classifier, they achieved an accuracy of 61.58% for multi-class (AD, MCI and NC) differentiation.

Other research focused on choosing or arranging and combining different classifiers to improve performance. Kim et al. [68] utilized extreme learning machine (ELM) to differentiate AD, MCI from NC and compared the performance between ELM and linear kernel SVM. For AD vs NC, accuracy of 92.84% was achieved (sensitivity 88.54%, specificity 96.12%); for MCI vs NC with accuracy of 78.28% (sensitivity 86.26%, specificity 81.55%); while for cMCI vs. ncMCI, accuracy was 87.72% (sensitivity 99.64%, specificity 42.42%). Huang et al. [94] used the thickness of cortex regions as the features, and attempted to increase the discriminant power by combining of SVM and Adaboost. Adaboost is an ensemble learning method [95], combining multiple "weak classifiers" into a single "strong classifier", iterative placing more weight on the misclassified training samples and decreasing the weights of the correctly classified ones. For AD vs NC, they achieved an accuracy of 84.38%, 4–10% higher than SVM, LDA and GMM alone. Farhan et al. [55] used an ensemble-of-classifiers technique for AD early diagnosis. Ensemble of classifiers [96,72] is a set of classifiers whose individual decisions are combined to classify a new sample based on voting. They used the volume of GM, WM, CSF and hippocampal size as the features. With combined features and ensemble of SVM, RF and ANN, an accuracy of 93.75% (sensitivity 87.5%, specificity 100%) was achieved for classifying AD vs NC.

The typologies and correlations derived from regions can be used as features. Wee et al. [52] used ROI-based cortical GM, cortical associated WM volumes and the correlation of regional (ROIs) mean cortical thickness as features. They used a multi-kernel SVM classifier, and achieved an accuracy of 92.35% (sensitivity 90.35%, specificity 94.31%), 83.75% (sensitivity 83.55%, specificity 83.95%) and 72.94% (sensitivity 78.03%, specificity 80.46%) in AD vs. NC, MCI vs. NC and AD vs. MCI classification respectively. For the prediction of pMCI (to AD) at different conversion times, at 36 months, they achieved an accuracy of 71.76% (sensitivity 73.64%, specificity 69.89%) for pMCI vs. sMCI, and for 12, 18 and 24 months, similar results were achieved. Bron et al. [97] used the a significance map to select features for which each voxel was derived from gray matter morphometry. As the significance map can quantify the relevance of individual features from the SVM weights, it may provide a better way in ranking the features and in estimating the relevance of individual fea-

tures. Liu et al. [70] provided a method for AD classification that used individual hierarchical networks for the feature representation. GM based ROIs were used for feature extraction and classification. Heterogeneous network was built based on ROIs (nodes) and links (edges) between ROIs. Different texture properties, such as energy, contrast, entropy were used to represent each ROI. SVM classifier was used and for AD vs MCI, accuracy of 90.41% was achieved (sensitivity 92.83%, specificity 88.82%), NC vs MCI 86.56% (sensitivity 90.74%, specificity 84.83%) and ncMCI vs. cMCI 73.95% (sensitivity 76.13%, specificity 72.24%). Tong et al. [57] used a graph to exploit the relationships among the patches and solved the multiple instance learning (MIL) problems to aid the classification. For NC vs AD, they achieved an accuracy of 89.2% (sensitivity 85.1%, specificity 92.6%), while for pMCI vs sMCI an accuracy of 69.3% (sensitivity 66.7%, specificity 71.2%).

One of the challenges faced in AD related research is that while predicting cMCI, it is difficult to train accurate classifiers, because of the small sample number of cMCI and ncMCI for training with high dimensionality of data. Cheng et al. [61] used domain transfer learning [98] for predictions of MCI to AD conversion. They provided a solution that uses auxiliary domains from NC and AD groups to differentiate ncMCI and cMCI in the target domain. They achieved an accuracy of 73.4% (sensitivity 74.3%, specificity 72.1%) when using MRI modality alone. While in Liu et al. [63], multi-templates were used for processing medical images. Their selected features corresponded to multiple templates and induced relationship from them. Ensemble of SVM classifiers were used to combine outputs. This learning method for automatic classification of AD and MCI reached a good performance, for AD vs NC, accuracy of 93.06% (sensitivity 94.85%, specificity 90.49%), while for sMCI vs pMCI with accuracy of 79.25% (sensitvity 87.92%, specificity 75.54%). While in Liu et al. [67], a similar method was utilised with multiple templates to generate multi-view morphological patterns as features. Ensembled view-specific SVM classifiers were used and results fused for classification. NC vs. AD achieved an accuracy of 93.83% (sensitvity 92.78%, specificity 95.69%) while pMCI vs. sMCI achieved an accuracy of 80.9% (sensitivity 85.95%, specificity 78.41%). Not only multiple-templates, but multiple atlases can also be used to improve performance. Min et al. [56] derived optimal representation from multiple atlases using maximum-margin based representation learning technique as compared with the traditional way of using one atlas space. With SVM as classifiers, for AD vs NC, an accuracy of 90.69% was achieved (sensitivity 87.5%, specificity 93.01%) and for sMCI vs pMCI, with accuracy of 73.69% (sensitivity 76.44%, specificity 70.76%).

Other research emphasis were placed on building more powerful aggregated biomarkers and by integrating complementary imaging modalities in a single learning framework if they are available. In Tong et al. [15], sparse representation techniques were used and several factors were studied. For MCI to AD conversion predictions, a grading biomarker was formed for every MCI subject, and combined it with age and cognitive measures. The prediction performance of conversion of an MCI subject to AD within 3 years was of 84.1% accuracy (sensitivity 88.7%, specificity 76.5%). There was a similar idea of building aggregate biomarkers for MCI to AD conversion prediction in Moradi et al. [60]. A semi-supervised learning was used to derive a biomarker for the conversion and then combined with age and cognitive measures to form an aggregate biomarkers. Ahmed et al. [99] used diffusion tensor imaging (another modality) plus the MRI to build multimodal AD signature to boost the discrimination power. The classification accuracies achieved by the proposed method are 90.2%, 79.42% and 76.63% for respectively AD versus NC, MCI versus NC and AD versus MCI binary classification problems.

Conventional Machine Learning research in the classification of AD and its stages in patients commonly utilizes structural MRI, some of them are presented in Tables 4–6. The most popular classifier SVM is a supervised learning model used for solving both regression and mostly classification problems. SVM excels in extracting effectively these high-dimensional representative features, which can be used by effective classifiers that for the automation of diagnosis. It is effective when number of features are more than training examples and it is the best algorithm when classes are separable. However, in conventional Machine Learning research, there are some innate problems. First of all, there is a high level of error susceptibility, that all events subsequent to an error may be problematic or just simply undesirable. Such errors do occur easily, and can take many forms. Secondly, as feature definition and extraction in the conventional Machine Learning typically rely on laborious work of segmenting of brain structures, complex image pre-processing techniques are required. Next, with complex and time consuming pre-possessing techniques, different research groups may use different techniques leading to huge amounts of variability with lack of standardization. These problems can be addressed by Deep Learning techniques which may have more to offer.

2.4. *Deep Learning approaches for AD diagnosis*

Deep Learning models high-level features and extract these abstractions from data with multiple non-linear transformations based on Machine Learning algorithms. The training process is considered to be deep because it consists of multiple layers, where the results from one layer will be transferred as input to the next layer. There are two phases in the Deep Learning process, a training and an inferring phase. The training phase uses huge amounts of data and searches for determining characteristics, while the other phase infers and labels the new unexposed data using the learned knowledge. As compared to conventional Machine Learning, Deep Learning algorithms do not require prior feature selection or data pre-processing (in some cases might require); they utilize either supervised or unsupervised training, with multiple stages and levels of abstraction to achieve an optimal representation. Since 2013, Deep Learning techniques in processing of medical images has gained momentum, with many deep models being developed [100–102]. Gupta et al. [103], one of the earliest works in this area, used cross–domain features that represent MRI data. They learned a set of bases from natural images and MRI scans with a sparse autoencoder, and then convoluted with MRI data from the ADNI dataset to extract features. Their 3-way classifier (NC vs AD vs MCI) achieved an accuracy of 85% and the binary classifiers (AD vs NC 94.74%, AD vs MCI 88.10% and MCI vs NC 86.35%)

Table 4 – Some of ML Methods in AD Researches with MRI.

Ref	Application	Classifier& Method	Datasets	Performance (%)
[43] 2009	classification: NC vs. MCI vs. AD	SVM (Linear + RBF); model the hippocamp shape with spherical harmonics coefficients	NC:25,MCI:23, AD:23	AD vs. NC: Acc 94, Sen 96, Spe 92; MCI vs. NC: Acc 83, Sen 83, Spe 84
[44] 2010	prediction: MCI to AD.	SVM, Naïve Bayesian and VFI; useing classifiers to derive a quantitative index of pattern matching for the prediction.	NC:18, MCI: 24, AD:32	NC vs. AD: Acc 92, Sen 93.75, Spe 88.89; MCI-MCI vs. MCI-AD: Acc 75, Sen 55.56, Spe 86.67;
[45] 2010	classification: AD vs. NC, MCI vs. AD and MCI vs. NC	OPLS analysis; segmentation of hippocampus with volume measurements (24 volume measures in all)	NC:112, AD:117, MCI:122	AD vs. NC: Sen 81–90, Spe 82–94; AD vs. MCI: Sen 71–75, Spe 70–79; MCI vs. NC: Sen 58–66, Spe 59–73
[46] 2011	classification: NC vs. AD	SVM; cortical thickness based kernels.	NC:196, AD:160	NC vs. AD: Acc 75, AUC 80.63
[47] 2011	classification: NC vs. AD, pMCI vs. sMCI	LDA and SVM; Combining features from different structural MRI analysis	NC:231,sMCI:238, pMCI:167, AD:198	AD vs. NC: Acc 89, Sen 93, Spe 85; sMCI vs. pMCI: Acc 68, Sen 67, Spe 69;
[48] 2011	classification: NC vs. AD, CN vs MCIc, MCIc vs MCInc	SVM; voxel-based or cortical thickness-based or hippocampus	NC:162,sMCI:134, pMCI:76, AD:137	AD vs. NC: Sen 81, Spe 95; pMCI vs. sMCI: Sen 70, Spe 61;pMCI vs. NC: Sen 73, Spe 85
[36] 2012	classification: NC vs. AD	QDA and LDA; using atrophic patterns of hippocampus and entorhinal cortex structures.	NC:60, AD:60	AD vs. NC: Acc 90, Sen 88, Spe 94;
[49] 2012	classification: NC vs. AD, ncMCI vs. cMCI	PCA-LDA; Using cortical thickness data.	NC:160,ncMCI:131, cMCI:72AD:128	AD vs. NC: Acc 89, Sen 82, Spe 93; sMCI vs. MCI: Acc 72, Sen 63, Spe 76; NC vs. cMCI: Acc 82,Sen 66, Spe 89
[50] 2012	classification: sMCI vs. cMCI, AD vs. MCI, MCI vs. NC	OPLS; Using Freesurfer to generate regional subcortical volumes and cortical thickness measures.	NC:111,cMCI: 81, sMCI: 81, AD:96	AD vs. MCI: ACC 87.0, Sen 83.3, Spe 90.1; MCI vs. NC: Acc 71.8, Sen 66.7, Spe 79.3;MCIc vs. MCIs: Acc 65.4, Sen 65.4, Spe 65.4;
[5] 2012	classification: pMCI vs. NC, AD vs. NC, sMCI vs. pMCI	LDA; data-driven approach for biomarker extraction.	NC:153,pMCI: 107, sMCI: 114, AD:96	AD vs. NC: ACC 90.3, Sen 87.5, Spe 92.1; pMCI vs. NC: Acc 86.9, Sen 81.2, Spe 90.9;pMCIc vs. sMCI: Acc 82.1, Sen 81.5, Spe 82.9;
[38] 2012	classification: Mild AD vs. NC, Mild AD vs. MCI	SVM and Naive Bayes; using subcortical and cortical volumes feature vectors	NC:29,MCI: 30, AD:21	NC vs. AD: Acc 86, Sen 82, Spe 90; MCI vs. AD: Acc 75, Sen73, Spe 77
[37] 2012	classification: NC vs. AD, pMCI vs. sMCI	LDA; ROI-wise cortical thickness	NC:226,sMCI:134, pMCI:340, AD:194	AD vs. NC: Acc 84.5, 79.40 82, Spe 88.9; sMCI vs. pMCI: Acc 66.70, Sen 59.00, Spe 70.20
[51] 2013	classification: AD vs. NC and ncMCI vs. AD.	OPLS, DT, ANN, SVM, 68 cortical thickness measures + 50 regional volumes	NC:110, AD:116, ncMCI:98, cMCI:21	cMCI vs. AD: Acc 81–86; AD vs. NC: Acc 78.3–88.9
[52] 2013	classification: AD vs. NC; MCI vs. NC; pMCI vs. sMCI	Multi-kernel SVM; using the correlation of regional mean cortical thickness between pairs of ROI to build a similarity map.	NC:200,sMCI:111, pMCI:89,AD:198	AD vs. NC: Acc 92.35, Sen 90.35, Spe 94.31; MCI vs. NC: 83.75, 83.55, 83.95; AD vs. MCI 72.94, 78.03, 80.46; pMCI vs. sMCI: Acc 75.05, Sen 63.52, Spe 84.41

Table 5 – Summary of ML Methods in AD Researches with MRI(continued).

Ref	Application	Classifier& Method	Datasets	Performance (%)
[35] 2013	classification: NC vs. AD, CN vs. cMCI, NC vs. ncMCI, ncMCI vs. cMCI	RLR; using large-scale regularized logistic regression to model the conditional probabilities.	NC:188,cMCI: 153, ncMCI: 182, AD:171	NC vs. AD: Acc 81.5, Sen 74.0, Spe 88.1; NC vs. cMCI: Acc 72.3, Sen 66.3, Spe78.9; NC vs. ncMCI: Acc 61.4, Sen 58.2, Spe 65.5; ncMCI vs. cMCI: Acc 61.5, Sen 45.8, Spe75.5
[53] 2014	classification: NC vs. AD, NC vs. MCI, AD vs. MCI	SVM(RBF kernel); using visual features from hippocampal area and extracting the hippocampus ROI.	NC:72,MCI: 111, AD:35	NC vs. AD: Acc 87, Sen 75.5, Spe 100; NC vs. MCI: Acc 78.22, Sen 92.83, Spe 83.34; NC vs. MCI: Acc 86.56, Sen 70.73, Spe 84.83; MCI vs. AD: Acc 72.23, Sen 70, Spe 75
[54] 2014	classification: NC vs. AD, aMCI vs. NC, naMCI vs. NC	SVM; Using FreeSurfer to calculate volumetric variables and constructing a decisional space through feature selection and ranking mechanism.	NC:127, aMCI:67, naMCI:56, AD:59	AD vs. NC: Acc 78.2, Sen 68.5, Spe 82.7; aMCI vs. NC: Acc 66.7, Sen 55.9, Spe 72.5; naMCI vs. NC: Acc 62.1, Sen 51.5, Spe 70.3
[55] 2014	classification: NC vs. AD	KNN, SVM and DT; volume of CSF, GM, WM and hippocampus size.	NC:48, AD:37	NC vs. AD: Acc 93.75, Sen 100, Spe 87.5
[56] 2014	classification: NC vs. AD, pMCI vs. sMCI	SVM; derived optimal representation from multiple atlases using maximum-margin based representation learning technique.	NC:128, sMCI:117, pMCI:117, AD:97	AD vs. NC: Acc 90.69, Sen 87.56, Spe 93.01; sMCI vs. pMCI: Acc 73.69, Sen 76.44, Spe 70.76
[57] 2014	classification: NC vs. AD, pMCI vs. sMCI	mi-Graph + LSVM; using a graph to exploit the relationships among the patches	NC:231,sMCI:238, pMCI:167, AD:198	AD vs. NC: Acc 89.2, Sen 85.1, Spe 92.6; pMCI vs. sMCI: Acc 69.3, Sen 66.7, Spe 71.2
[58] 2014	classification: AD and NC.	ELM,OP-ELM; VBM, evolutionary wrapper feature selection	NC:49, AD:49	AD vs. NC: Acc 86, Sen 87 Spe 90
[59] 2015	classification: NC vs. AD vs. MCI vs. MCIc vs. MCInc	SVM;using a significance map to select the features	NC:162,MCInc:134, MCIc:76, AD:137	AD vs. CN: 90.0, AD vs. MCI 64.8, MCI vs. CN 71.6, and MCIc vs. MCInc 60.9, all in terms of AUC
[60] 2015	discriminating: pMCI vs. sMCI.	SVM and LDS; a conversion biomarker based on semi-supervised learning, combined with with age and cognitive measures about the subjects.	NC:231,pMCI: 164, sMCI: 100, uMCI: 130, AD:200	SVM: Acc 69.15, Sen 86.73, Spe 40.34; LDS: Acc 74.74, Sen 88.85, Spe 51.59;
[61] 2015	classification: ncMCI vs. cMCI	DTSVM + DTSS + DTFS; using a domain transfer learning method for MCI conversion prediction.	NC:52,ncMCI56, cMCI: 43, AD:51	ncMCI vs. cMCI: Acc 73.4, Sen 74.3, Spe 72.1;
[33] 2015	classification: AD and NC.	SVM; maximum inter-class variance, eigenbrain with Welch's t-test, SVM with different kernels	NC:98, AD:28	AD vs. NC: Acc 86.71–92.36, Sen 83.48–90.17 Spe 86.99–94.90
[62] 2015	classification: AD, MCI and NC.	linear SVM, kernel SVM, and kernel SVM by PSO with PSOTVAC, Winner-Takes-All, Max- Wins-Voting, and Directed Acyclic Graph; 3D discrete wavelet transform, PCA,	NC:97, AD:24, MCI:57	overall accuracy of 81.5%
[63] 2016	classification: NC vs. AD, SMCI vs. PMCI	Ensemble SVM; Model the relationships among multi-templates and among individual subjects.	NC:128,sMCI:117, pMCI:117, AD:97	AD vs. NC: Acc 93.06, Sen 94.85, Spe 90.49, AUC 95.79; sMCI vs. pMCI: Acc 79.25, Sen 87.92, Spe 75.54
[39] 2016	classification: NC vs. AD	MLP, SVM and NB; analysing the shape changes of CC based on Laplace Beltrami eigenvalue features.	NC:30, AD:30	CN vs. AD: Acc 93.37, Sen 93.37, Spe 93.37
[64] 2016	classification: NC vs. AD, MCI vs. NCI	SVM;compared different feature selection method.	NC:231, MCI:400, AD:200	AD vs. NC: Acc 84.6; MCI vs. NC: Acc 79.7
[65] 2016	classification: NC vs. AD	SVM; using voxel-based morphometry with Fisher Criterion top features selection.	NC:68, AD:68	NC vs. AD: Acc 96.32, Sen 94.11, Spe 98.52

Table 6 – Summary of ML Methods in AD Researches with MRI(continued).

Ref	Application	Classifier& Method	Datasets	Performance (%)
[66] 2016	classification: AD vs. NC, cMCI vs. ncMCI	SVM; wavelet transforms based multi-scale features extraction	NC:228, cMCI: 71, AD:188, ncMCI: 62	NC vs. AD: Acc 85.10, Sen 84.57, Spe 85.53; ncMCI vs. cMCI: Acc 76.69, Sen 71.83, Spe 82.26
[67] 2016	classification: AD vs. NC, PMCI vs. sMCI	Ensemble SVM; Using multiple selected templates, extracting multiview feature representations, and clustering subjects into different subclasses	NC:128, sMCI: 117, AD:97, pMCI: 117	NC vs. AD: Acc 93.83, Sen 92.78, Spe 95.69; pMCI vs. sMCI: Acc 80.9, Sen 85.95, Spe 78.41
[15] 2017	classification: NC vs. AD, SMCI vs. PMCI	SVM& RF; Combined biomarker based on sparse representation techniques	NC:229,sMCI:129, pMCI:171,uMCI 98,AD:191	sMCI vs.pMCI: Acc 84.1 Sen 88.7 Spe 76.5
[68] 2017	classification: AD vs. NC vs. MCI	ELM; recursive feature elimination algorithm and compared with SVM.	NC:208,nMCI:281, cMCI:69, AD:160	AD vs.NC: Acc 92.84, Sen 88.54, Spec 96.12;MCI vs.NC: Acc 78.28, Sen 86.26, Spe 81.55; cMCI vs. ncMCI: Acc 87.72, Sen 99.64, Spe 42.42
[69] 2017	classification: AD vs. NC, pMCI vs. sMCI	SVM; ranking features(VOI from GM) with t-test, and a genetic algorithm to find the optimal set	NC:94, sMCI: 65, AD:92, pMCI: 71	NC vs. AD: Acc 93.01, Sen 89.13, Spe 96.80; pMCI vs. sMCI: Acc 75, Sen 76.92, Spe 73.23
[70] 2017	classification: NC vs. AD, AD vs. MCI, NC vs. MCI, ncMCI vs. cMCI	MKL and SVM; using individual hierarchical networks constructed with 3D texture features from GM.	NC:230,cMCI: 120, ncMCI: 160, AD:200	NC vs. AD: Acc 95.37, Sen 92.49, Spe 96.08; AD vs. MCI: Acc 90.41, Sen 92.83, Spe 88.82; NC vs. MCI: Acc 86.56, Sen 90.74, Spe 84.83; ncMCI vs. cMCI: Acc 73.95, Sen 76.13, Spe 72.24
[71] 2017	classification: MCI, AD and NC.	SVM, IVM, and RELM; a greedy score-based feature selection technique is employed.	NC:70, AD:70, MCI:74	AD vs. NC: Acc 59.5–78.1, Sen 62.12–75.81, Spe 62.85–88.81; AD vs. NC vs. MCI: Acc 60.2–80.32, Sen 61.70–81.70, Spe 61.20–90.63
[72] 2018	classification: AD vs. NC,MCI	SVM + GPC + Adaboot;Ensemble of classifiers.	NC:100,nMCI:100, MCI:100, AD:100	Acc 57.1; AUC 79.6
[73] 2018	classification: NC vs. MCI vs. cMCI vs. AD	Ensemble SVM (Linear + RBF), bagging and feature selection	NC:100,MCI:100, cMCI:100, AD:100	55.6/55.0
[74] 2018	discriminating: cMCI vs. ncMCI.	Linear SVM; ROI and cortical thickness based	cMCI: 18, ncMCI: 62	SVM: Acc 89.41, Sen 91.40, Spe 87.97
[41] 2019	classification: NC vs. AD, AD vs. MCI, NC vs. MCI, ncMCI vs. cMCI	KNN,RF and SVM; using specific image blocks and computes the textures from WM and GM.	NC:227,cMCI: 165, ncMCI: 231, AD:189	NC vs. AD: Acc 87.39, Sen 85.42, Spe 88.81; AD vs. MCI: Acc 63.41, Sen 57.29, Spe 65.36; NC vs. MCI: Acc 64.74, Sen 45.61, Spe 72.44; ncMCI vs. cMCI: Acc 66.38, Sen 60.24, Spe 69.8
[75] 2019	prediction: MCI to AD.	KLS-SVM; Fast DWT + F-Score + PCA + Poly Kernel, cortical thickness, curvature, gray matter volume, surface area, and sulcal depth.	NC:145, AD:158	NC vs. AD: Acc 100;
[42] 2020	classification: AD and NC.	DT, KNN,PNN, LDA, SVM;Bi-directional empirical mode decomposition, extraction oflocal binary patterns	NC:110, AD:55	AD vs. NC: Acc 93.9, Sen 90.09 Spe 95.5
[76] 2020	classification: stable NC vs. Dementia of Alzheimer's type (DAT), sMCI vs. pMCI	Variational Bayes Probabilistic Multi-Kernel Learning;	stable NC: 423, DAT: 330, validation on 8,834 unseen images from different AD databases	baseline DAT vs. baseline sNC: Acc 89.5, Sen 87.3, Spe 91.3; follow-up DAT vs. follow-up sNC: Acc 88.5, Sen 93.2, Spe 86.4; sMCI vs. pMCI (6 months to 7 years): AUC 0.729–0.803, Acc: 67.6–71.2
[34] 2020	classification: AD, NC and pMCI.	DT, RF, NB, LR, SVM; brain volume based approach	NC:72, AD:64, pMCI:14	LR: Acc 93; DT: Acc 83; RF: Acc 90; NB: Acc 93; SVM with linear kernel: Acc 95
[77] 2021	classification: MCI, AD and NC.	MCSVM; multiclass classification framework with embedding feature selection and fusion based on multimodal neuroimaging.	NC:40, AD:38, MCI:42	AD vs. MCI: Acc 83.62; AD vs. NC: Acc 87.65; MCI vs. NC: Acc 77.20; AD vs. MCI vs. NC: Acc 83.78

achieved good performance despite the simple approach. Another early work by Brosch et al. [104] proposed to use a deep belief network [105] to find patterns of similarity in 3D brain images (MRI) groups for classification of AD and NC.

Payan et al. [106], combined sparse auto-encoders and convolutional neural networks (CNNs) based on the same dataset and number of samples as in [103]. 3D convolutions on the whole MRI image were used, and their results were compared with 2D convolutions on slices in these experiments. The 3D CNNs outperformed (0–4%) the 2D counterparts. The 3-way classifier achieved an accuracy of 89.47% and their binary classifiers (AD vs. NC 95.39%, AD vs. MCI 86.84% and MCI vs. NC 92.11%) also achieved a better performance (More details with Deep Learning techniques are presented in Table 7).

The winner of the ILSVRC, an annual AI challenge that classify and detect objects and scenes, in 2014, was GoogLeNet [107] from Google. The architecture consists of 22 layers and drastically diminished parameter numbers. The top-5 error rate was 6.67%, close to human level performance. Residual Neural Network (ResNet) [108], the winner of ILSVRC 2015, introduced a novel architecture with skip connections while bypassing the convolutional layers of regular feed-forward network. It beat human level performance, with a top-5 error rate of 3.57%. Naturally, Farooq et al. [109] subsequently provided a framework which included both GoogLeNet and ResNet based on CNN. A 4-way classifier was provided, which differentiated AD, MCI, LMCI (Later stage MCI) and NC. GoogLeNet achieved the highest classification accuracy of 98.88%. Korolev et al. [110] used two network architectures, VoxCNN (similar to VGGNet, the runner-up of ILSVRC 2014) and ResNet to examine the effectiveness of the convolutional network to extract features for 3D image classification. They developed Deep Learning based algorithms that had no need for the handcrafted feature generation, with potentials to form end-to-end models that simplify the MRI classification process.

Hosseini-Asl et al. [111] proposed a framework that uses a source domain-trained 3D-CAE to extract features from MRI brain images, with transfer learning techniques for AD diagnosis. It was also able to perform task-specific classification, such as in classifying MCI vs NC, with a supervised target-domain-adaptable CNN, which is an improvement based on [112]. The AD vs MCI vs NC classifier achieved an accuracy of 94.8%, and AD + MCI vs. NC, AD vs. NC, AD vs. MCI, MCI vs. NC achieved good accuracies from 94.2% to 100% (Table 7).

Islam et al. [82] provided a framework that ensembles three deep CNNs, each of them had different configurations, with OASIS as the dataset (refer to Table 2). In order to reduce overfitting, data augmentation was used to enlarge the dataset. The multi-classfier achieved good results for normal controls, very mild, mild and moderate AD, with average precision of 94%. Wang et al. [113] proposed an eight-Layer CNN with leaky rectified linear unit and max pooling that classify AD and NC, which achieved a sensitivity of 97.96%, a specificity of 97.35%, and an accuracy of 97.65%. They had also compared the proposed method with 8 other existing popular methods, which displayed an increased accuracy of approximately 5% in comparison.

One interesting point we would like to mention is that, past studies have indicated that to achieve a high performance, the number of hidden layers in the Deep Learning models should be neither too small nor too large. The research work we studied confirm this point, for example, Liu et al. [30] had the conclusion that widening the model brings consistent gains while increasing the depth does not. Jain et al. [31], their VGG-16 (16 layers) based CNN network achieved an 3-way classification accuracy of 95.73, whereas Wang et al. [113] eight-layer (8 layers) CNN network achieved a close classification accuracy.

From Table 7 (as compared to Tables 4–6), we can draw some conclusions when comparing Deep Learning and conventional Machine Learning. Firstly, Deep Learning models generally achieve higher performance than conventional Machine Learning techniques. For example, in classifying NC vs AD cases, Deep Learning models achieve an average accuracy of over 95%, while accuracies of traditional Machine Learning range mostly within 75% to 93.83%. For MCI vs NC, Deep Learning has a performance accuracy of 4–20% over conventional methods. However, for cMCI and ncMCI cases, most of the Deep Learning techniques perform better, with exceptions of conventional methods excelling.

Secondly, it is noted that datasets for studies, either Deep or conventional Machine Learning models, are similar, most of them use either ADNI or OASIS. However, the difference lies in that Deep Learning uses more samples for training, such as in Gupta [103], where there were 232 NC with 1278 scans, 200 AD individuals with 755 scans and 411 MCI individuals with 2282 scans from the ADNI dataset. Also, Basaia et al. [117] had 407 NC, 418 AD, 533 sMCI, 280 cMCI from the ADNI databse, augmenting the dataset of any group to 1000. Similarly, Islam et al. [82] used participants from OASIS1: 416 in all, 100 mild to moderate AD, 20 non-demented. The dataset was then augmented and used for the process. Cho et al. [119] provided insights with regards to the needed amounts of image data to achieve certain accuracy level when training a Deep Learning system. According to the research, with CNN model used to classify images into six classes. With a size of 1000 training data per class, the learning curve predicted a 98% classification accuracy (observed accuracy at 97.25%). Based on the learning curve, it predicted that with a training data size of 4092 per class, the Deep Learning classifier can achieve the desired accuracy of 99.5%. Hence, the reasonable size of the training samples should start from about 1000, with other studies confirming this number [117,111,103,106]. If training samples are sufficient, 5000 or more can be chosen to train the deep model to achieve a much higher system accuracy. Some research groups also made use of transfer learning [120–122] to build their deep learning algorithms in attempt to solve the issue of lack of samples. Transfer learning is the fact that the previous learned knowledge can be used to solve a new problem faster and better.

Thirdly, we see that from Table 7, most of the Deep Learning research in AD and its stages classification use CNNs. CNNs enables data-driven learning, finds highly representative, layered features effectively [123] and can tune parameters quickly. A typical CNN usually has alternatively series of convolutional layers and pooling layers, and a few fully connected layers combined together, as shown in Fig. 4.

Table 7 – Summary of Deep Learning Methods in AD Research with MRI.

Ref	Application	Method	Datasets	Performance (%)
Gupta et al. 2013 [103]	Classification: NC vs. AD vs. MCI and AD vs. NC, AD vs. MCI and MCI vs. NC.	CNN; using cross–domain features to represent MRI data. Sparse auto-encoder to learn bases from natural images and used convolution to extract features, 2D convolutions on slices.	NC:232, AD:200, MCI:411. 755 scans for each group, 2265 scans in all	NC vs. AD: Acc 94.74, Sen 95.24, Spe 94.26; NC vs. MCI: Acc 86.35, Sen 92.23, Spe 81.45; AD vs. MCI: Acc 88.10, Sen 84.07, Spe 92.11. HC vs. AD vs. MCI: Acc 85.00
Suk et al. 2014 [114]	Three binary classification: AD vs. NC, MCI vs. NC, and cMCI vs. ncMCI	DBM for features and multimodal fusion	cMCI:76, ncMCI:128, NC:101, AD: 93	AD vs. NC: Acc 92.38, Sen 91.54, Spe 94.56; MCI vs. NC: Acc 84.24, Sen 99.58, Spe 53.79; cMCI vs. ncMCI: Acc 72.42, Sen 36.7, Spe 90.98
Payan et al. 2015 [106]	Classification: NC vs. AD vs. MCI and AD vs. NC, AD vs. MCI and MCI vs. NC.	CNN; CNN combined with sparse auto-encoders. Used 3D convolutions on the whole MRI image and 2D convolutions on slices.	NC:232, AD:200, MCI:411. 755 scans for each group, 2265 scans in all	NC vs. AD: Acc 95.39 (2D), 95.39 (3D); AD vs. MCI: Acc 82.24 (2D), 86.64 (3D); NC vs. MCI: Acc 90.13 (2D), 92.11 (3D). HC vs. AD vs. MCI: Acc 85.53 (2D), 89.47 (3D)
Sarraf et al. 2016 [115]	Classification: AD vs NC	CNN architectures of LeNet-5 and GoogleNet	NC: 97, AD: 211	AD vs. NC: Acc 98.84
Farooq et al. 2017 [109]	Classification: HC vs. AD vs. MCI vs. LMCI.	CNN based architecture GoogLeNet and ResNet used	NC:45, AD:33, LMCI:22, MCI:49. MRI Volumes: 355. Images: 38024	GoogLeNet Acc: 98.88, ResNet Acc:98.01, ResNet-152 Acc:98.14
Korolev et al. 2017 [110]	Six binary classification: AD vs. NC, AD vs. EMCI, AD vs. LMCI, LMCI vs. NC, LMCI vs. EMCI, EMCI vs. NC.	ResNet and VoxCNN	NC: 61, AD: 50, LMCI: 43, EMCI: 77, 231 images in all	AD vs. NC: VoxCNN Acc-Auc 79–88, ResNet Acc-Auc 80–87; AD vs. EMCI: VoxCNN 64–66, ResNet 63–67; AD vs. LMCI: VoxCNN 62–61, ResNet 59–62; LMCI vs. NC: VoxCNN 63–67, ResNet 61–65; LMCI vs EMCI: VoxCNN 56–47, ResNet 52–52; EMCI vs. NC: VoxCNN 54–57, ResNet 56–58
Hosseini et al. 2018 [111]	Binary classification: AD vs. NC, AD & MCI vs. NC, AD vs. MCI and MCI vs. NC. Classification: AD vs. MCI vs. NC	3D CNN model based, and modified to suit different domain	NC:70, AD:70, MCI:70. 755 scans for each group,2265 scans in all	AD vs. NC Acc: 99.3; AD & MCI vs. NC Acc: 95.73; AD vs. MCI Acc: 100; and MCI vs. NC Acc: 94.22. for AD vs. MCI vs. NC Acc:94.8
Islam et al. 2018 [82]	Classification: non-demented vs. mild AD vs. moderate AD vs. very mild AD.	2D CNN based architecture, an ensemble of three deep CNN	OASIS-1: 416 in all, 100 mild to moderate AD, 20 non-demented	Non-demented (precision-recall-f1-score): 97–100-99; very mild AD: 100–33-75; mild: 67–86-75; moderate 50–50-50
Lian et al. 2018 [116]	classification: AD vs NC, cMCI vs NC, sMCI vs NC, AD vs cMCI, AD vs sMCI, cMCI vs sMCI.	Hierarchical fully convolutional network	ADNI-1 (NC: 229, AD: 199, sMCI: 226, cMCI: 167) ADNI-2 (NC: 200, AD: 159, sMCI: 239, cMCI: 38), augmented the dataset	AD vs. NC: Acc 90.3, Sen 82.4, Spe 96.5; cMCI vs. sMCI: Acc 80.9, Sen 52.6, Spe 85.4
Wang et al. 2018 [113]	classification: AD vs. NC	8-layer CNN with optimal structure obtained by experiences	NC: 98, AD: 98, augmented the dataset	AD vs. NC: Acc 97.65, Sen 97.96, Spe 97.35
Jain et al. 2019 [31]	3-way classification: AD vs. MCI vs. NC	2D CNN, with VGG-16 and transfer learning	NC: 50, AD: 50, MCI: 50, 4800 images in all	three-way: 95.73; AD vs. NC: 99.14; AD vs. MCI: 99.30; NC vs. MCI: 99.22
Basaia et al. 2019 [117]	classification: NC vs.AD, cMCI vs. NC, sMCI vs. NC, AD vs. cMCI, AD vs. sMCI, cMCI vs. sMCI.	3D CNN	NC: 407, AD: 418, sMCI: 533, cMCI: 280, augmented the dataset of any group to 1000	AD vs. NC: Acc 99.2, Sen 98.9, Spe 99.5; cMCI vs. NC: Acc 87.1, Sen 87.8, Spe 86.5; cMCI vs. AD: Acc 75.4, Sen 74.5, Spe 76.4; AD vs. sMCI: Acc 85.9, Sen 83.6, Spe 88.3; cMCI vs. sMCI: Acc 75.1, Sen 74.8, Spe 75.3
Oh et al. 2019 [118]	Classification: AD vs. NC, pMCI vs. NC, sMCI vs. NC, pMCI vs. AD, sMCI vs. AD	CNN, CAE based unsupervised learning, and transfer learning,	NC: 230, AD: 198, pMCI:166, sMCI:101	AD vs. NC: Acc 86.60, Sen 88.55, Spe 84.54; pMCI vs. NC: Acc 77.37, Sen 81.03, Spe 74.07; sMCI vs. NC Acc 63.04, Sen 59.02, Spe 67.11; pMCI vs. AD: Acc 60.97, Sen 64.53, Spe 56.13; sMCI vs. AD: Acc 75.06, Sen 76.55, Spe 73.39
Liu et al. 2020 [30]	classification: AD, NC and MCI.	3D CNN with instance normalization, small-sized kernels and wider network.	NC: 197, AD: 279, MCI: 330, 3000 scans in all	Acc 66.9, balanced Acc 67.9, Micro-AUC 82.0, Macro-AUC 80.0

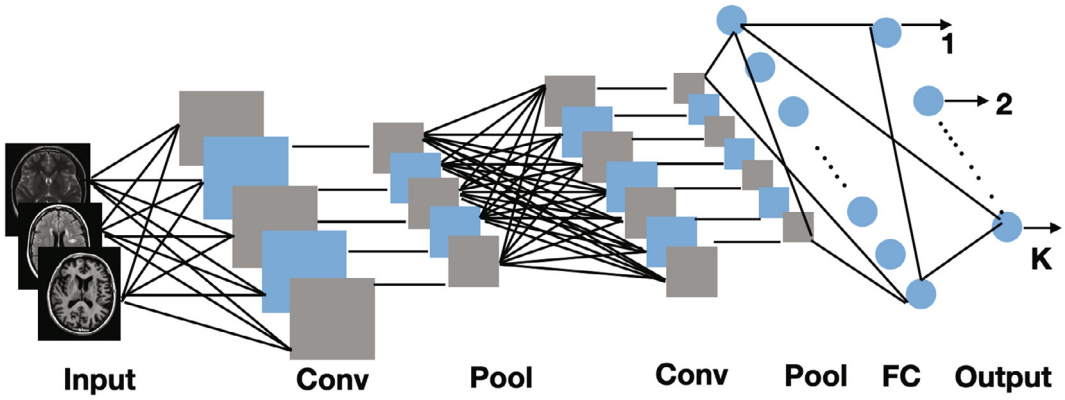

Fig. 4 – A Typical CNN Architecture.

Note that in this review paper, we mainly focused on one modality, namely the structural MRI. However, if there are more available data and/or with other modalities, the combined efforts, such as multi-modal fusion [124], can improve the classification performance for both Deep Learning [125,116] and traditional Machine Learning [50,74,77].

3. Discussion and future framework

The objective of future works would be to help with the early detection and diagnosis for individuals at risk of developing AD. However, the challenge is that investigations are only performed when individuals appear to have noticeable cognitive changes, either to themselves or by people around them. Some public databases have started to include MRI scans from participants with self-reported significant memory concerns. The study of ADNI2 (refer to Table 2 and http://adni.loni.usc.edu/about), for example, included 107 of such subjects. There is limitation faced in the recruitment of such individuals. As a remedy, public awareness can be raised via the news, advertisements and pamphlets, to encourage these individuals with memory concerns or early MCI to seek

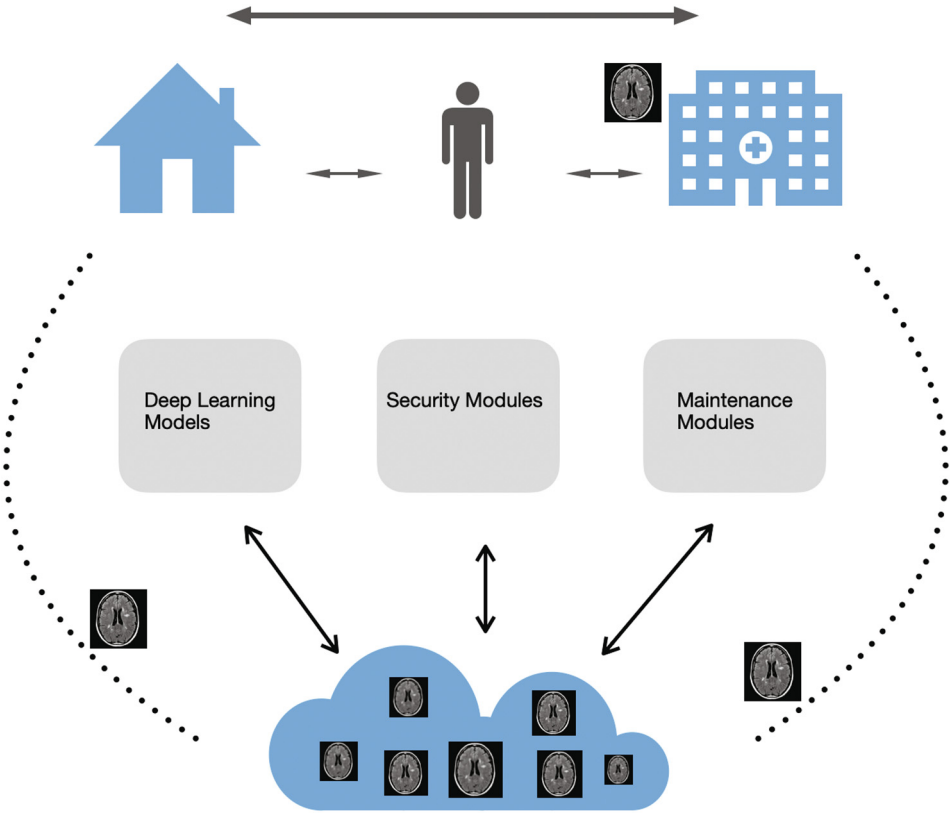

Fig. 5 – A future framework of the Machine Learning based diagnostic approach.

medical attention for further work up. However, with mild symptoms, it may be easy for experts to overlook the slightest structural changes such as atrophy in scans, because human instincts cannot reliably detect such changes, meaning that individuals may not be diagnosed effectively even if they have concerns. There may also be stigma associated with the diagnosis of AD, hence limiting individuals in stepping forward to be investigated.

With the advancement of smart home appliances (sensors) and cloud computing technologies, we can combine traditional diagnosis with modern technologies to form a framework which can provide more reliable diagnosis and prediction of dementia and its stages. For instance, Akl et al. [126] proposed non-obtrusive home application of sensing technologies detection, with the combination of Machine Learning techniques. These MCI subjects can be detected with an area under the ROC curve of 0.97 using a time window of 24 weeks.

Such home applications can well be included into the framework (as shown in Fig. 5). With smart appliances, individuals may be prompted to seek medical attention for further work up. With traditional diagnosis, individuals may be equipped with advice in living with AD early in diagnosis. If neuroimaging is performed, medical professionals may also get opinions from other specialists by uploading the medical images via cloud computing. Patients can also be more involved in their care with copies of these scans. With availability of more samples on-site, and developement of more advanced Deep Learning algorithms from research, we aim to achieve greater accuracies in the diagnosis and prediction of AD.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Xinxing Zhao: Software, Methodology, Validation, Formal analysis, Investigation, Writing - original draft, Writing - review & editing, Visualization. **Candice Ke En Ang:** Methodology, Validation, Formal analysis, Investigation, Writing - original draft, Writing - review & editing, Visualization. **U. Rajendra Acharya:** Methodology, Formal analysis, Writing - review & editing, Visualization. **Kang Hao Cheong:** Conceptualization, Methodology, Validation, Formal analysis, Investigation, Resources, Data curation, Writing - original draft, Writing - review & editing, Visualization, Supervision, Project administration, Funding acquisition.

REFERENCES

[1] National Institute On Aging, What Is Dementia? Symptoms, Types, and Diagnosis, https://www.nia.nih.gov/health/what-dementia-symptoms-types-and-diagnosis, note = [Online] Accessed: 2020-09-19.

[2] National Institute On Aging, What Is Alzheimer's Disease?, https://www.nia.nih.gov/health/what-alzheimers-disease#::text=Alzheimer's note = [Online] Accessed: 2020-09-19.

[3] Maslow K et al. 2010 alzheimer's disease facts and figures. Alzheimer's Dementia J Alzheimer's Assoc 2010;6(2):158–94. https://doi.org/10.1016/j.jalz.2010.01.009.

[4] Alzheimer's Association, 2019 alzheimer's disease facts and figures report, https://www.alz.org/media/documents/alzheimers-facts-and-figures-2019-r.pdf, accessed: 2020-01-15.

[5] Vounou M, Janousova E, Wolz R, Stein JL, Thompson PM, Rueckert D, Montana G, Initiative ADN, et al. Sparse reduced-rank regression detects genetic associations with voxel-wise longitudinal phenotypes in alzheimer's disease. Neuroimage 2012;60(1):700–16. https://doi.org/10.1016/j.neuroimage.2011.12.029.

[6] Alzhemer's Disease International, World Alzheimer Report 2019, https://www.alz.co.uk/research/WorldAlzheimerReport2019.pdf, [Online] Accessed: 2020-03-9 .

[7] Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, Iwatsubo T, Jack Jr CR, Kaye J, Montine TJ, et al. Toward defining the preclinical stages of alzheimer's disease: Recommendations from the national institute on aging-alzheimer's association workgroups on diagnostic guidelines for alzheimer's disease. Alzheimer's Dementia 2011;7(3):280–92. https://doi.org/10.1016/j.jalz.2011.03.003.

[8] Gauthier S, Reisberg B, Zaudig M, Petersen RC, Ritchie K, Broich K, Belleville S, Brodaty H, Bennett D, Chertkow H, et al. Mild cognitive impairment. The Lancet 2006;367(9518):1262–70. https://doi.org/10.1016/S0140-6736(06)68542-5.

[9] Michaud TL, Su D, Siahpush M, Murman DL. The risk of incident mild cognitive impairment and progression to dementia considering mild cognitive impairment subtypes. Dementia Geriatric Cogni Disorders Extra 2017;7(1):15–29. https://doi.org/10.1159/000452486.

[10] Mitchell AJ, Shiri-Feshki M. Rate of progression of mild cognitive impairment to dementia–meta-analysis of 41 robust inception cohort studies. Acta Psychiatr Scand 2009;119(4):252–65. https://doi.org/10.1111/j.1600-0447.2008.01326.x.

[11] Folstein MF, Folstein SE, McHugh PR. "mini-mental state": a practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975;12(3):189–98. https://doi.org/10.1016/0022-3956(75)90026-6.

[12] Morris JC. The Clinical Dementia Rating (CDR). Neurology 1993;43(11):2412–2412-a. doi:10.1212/WNL.43.11.2412-a.

[13] Kasban H, El-Bendary M, Salama D. A comparative study of medical imaging techniques. Int J Inf Sci Intell Syst 2015;4(2):37–58.

[14] Wright A. Brain scanning techniques (ct, mri, fmri, pet, spect, dti, dot). Cerebra Positively Different 2010:1–14.

[15] Tong T, Gao Q, Guerrero R, Ledig C, Chen L, Rueckert D. A novel grading biomarker for the prediction of conversion from mild cognitive impairment to alzheimer's disease. IEEE Trans Biomed Eng 2017;64:155–65. https://doi.org/10.1109/TBME.2016.2549363.

[16] Ledig C, Schuh A, Guerrero R, Heckemann RA, Rueckert D. Structural brain imaging in alzheimer's disease and mild cognitive impairment: biomarker analysis and shared morphometry database. Sci Rep 2018;8(1):1–16. https://doi.org/10.1038/s41598-018-29295-9.

[17] Jack Jr CR, Knopman DS, Jagust WJ, Petersen RC, Weiner MW, Aisen PS, Shaw LM, Vemuri P, Wiste HJ, Weigand SD, et al. Tracking pathophysiological processes in alzheimer's disease: an updated hypothetical model of dynamic biomarkers. Lancet Neurol 2013;12(2):207–16. https://doi.org/10.1016/S1474-4422(12)70291-0.

[18] Yu Q, Mai Y, Ruan Y, Luo Y, Zhao L, Fang W, Cao Z, Li Y, Liao W, Xiao S, et al. An mri-based strategy for differentiation of frontotemporal dementia and alzheimer's disease. Alzheimer's Res Therapy 2021;13(1):1–12. https://doi.org/10.1186/s13195-020-00757-5.
[19] Kok C, Jahmunah V, Oh SL, Zhou X, Gururajan R, Tao X, Cheong KH, Gururajan R, Molinari F, Acharya UR. Automated prediction of sepsis using temporal convolutional network. Comput Biol Med 2020;127 . https://doi.org/10.1016/j.compbiomed.2020.103957 103957.
[20] Jahmunah V, Oh SL, Rajinikanth V, Ciaccio EJ, Cheong KH, Arunkumar N, Acharya UR. Automated detection of schizophrenia using nonlinear signal processing methods. Artif Intell Med 2019;100 . https://doi.org/10.1016/j.artmed.2019.07.006 101698.
[21] Lin AX, Ho AFW, Cheong KH, Li Z, Cai W, Chee ML, Ng YY, Xiao X, Ong MEH. Leveraging machine learning techniques and engineering of multi-nature features for national daily regional ambulance demand prediction. Int J Environ Res Public Health 2020;17(11):4179. https://doi.org/10.3390/ijerph17114179.
[22] Ho AFW, To BZYS, Koh JM, Cheong KH. Forecasting hospital emergency department patient volume using internet search data. IEEE Access 2019;7:93387–95. https://doi.org/10.1109/ACCESS.2019.2928122.
[23] Cheong KH, Ngiam NJ, Morgan GG, Pek PP, Tan BY-Q, Lai JW, Koh JM, Ong MEH, Ho AFW. Acute health impacts of the southeast asian transboundary haze problem—a review. Int J Environ Res Public Health 2019;16(18):3286. https://doi.org/10.3390/ijerph16183286.
[24] Ho AFW, Zheng H, Cheong KH, En WL, Pek PP, Zhao X, Morgan GG, Earnest A, Tan BYQ, Ng YY, et al. The relationship between air pollution and all-cause mortality in singapore. Atmosphere 2020;11(1):9. https://doi.org/10.3390/atmos11010009.
[25] Cheong KH, Poeschmann S, Lai JW, Koh JM, Acharya UR, Yu SCM, Tang KJW. Practical automated video analytics for crowd monitoring and counting. IEEE Access 2019;7:183252–61. https://doi.org/10.1109/ACCESS.2019.2958255.
[26] Tanveer M, Richhariya B, Khan R, Rashid A, Khanna P, Prasad M, Lin C. Machine learning techniques for the diagnosis of alzheimer's disease: a review. ACM Trans Multimed Comput Commun Appl (TOMM) 2020;16(1s):1–35. https://doi.org/10.1145/3344998.
[27] Siuly S, Zhang Y. Medical big data: neurological diseases diagnosis through medical data analysis. Data Sci Eng 2016;1(2):54–64. https://doi.org/10.1007/s41019-016-0011-3.
[28] Leandrou S, Petroudi S, Kyriacou PA, Reyes-Aldasoro CC, Pattichis CS. Quantitative mri brain studies in mild cognitive impairment and alzheimer's disease: a methodological review. IEEE Rev Biomed Eng 2018;11:97–111. https://doi.org/10.1109/RBME.2018.2796598.
[29] Ahmed MR, Zhang Y, Feng Z, Lo B, Inan OT, Liao H. Neuroimaging and machine learning for dementia diagnosis: recent advancements and future prospects. IEEE Rev Biomed Eng 2018;12:19–33. https://doi.org/10.1109/RBME.2018.2886237.
[30] Liu S, Yadav C, Fernandez-Granda C, Razavian N. On the design of convolutional neural networks for automatic detection of alzheimer's disease. In: Machine Learning for Health Workshop, PMLR; 2020. p. 184–201.
[31] Jain R, Jain N, Aggarwal A, Hemanth DJ. Convolutional neural network based alzheimer's disease classification from magnetic resonance brain images. Cogn Syst Res 2019;57:147–59. https://doi.org/10.1016/j.cogsys.2018.12.015.
[32] Nguyen DT, Ryu S, Qureshi MNI, Choi M, Lee KH, Lee B. Hybrid multivariate pattern analysis combined with extreme learning machine for alzheimer's dementia diagnosis using multi-measure rs-fmri spatial patterns. PloS One 2019;14(2) . https://doi.org/10.1371/journal.pone.0212582 e0212582.
[33] Zhang Y, Dong Z, Phillips P, Wang S, Ji G, Yang J, Yuan T-F. Detection of subjects and brain regions related to alzheimer's disease using 3d mri scans based on eigenbrain and machine learning. Front Comput Neurosci 2015;9:66. https://doi.org/10.3389/fncom.2015.00066.
[34] Kishore P, Kumari CU, Kumar M, Pavani T. Detection and analysis of alzheimer's disease using various machine learning algorithms, Mater Tod Proc. doi:10.1016/j.matpr.2020.07.645 .
[35] Casanova R, Hsu F-C, Sink KM, Rapp SR, Williamson JD, Resnick SM, Espeland MA, Initiative ADN, et al. Alzheimer's disease risk assessment using large-scale machine learning methods. PloS One 2013;8(11). https://doi.org/10.1371/journal.pone.0077949.
[36] Coupé P, Eskildsen SF, Manjón JV, Fonov VS, Collins DL. Simultaneous segmentation and grading of anatomical structures for patient's classification: application to alzheimer's disease. NeuroImage 2012;59:3736–47. https://doi.org/10.1016/j.neuroimage.2011.10.080.
[37] Eskildsen SF, Coupé P, García-Lorenzo D, Fonov V, Pruessner JC, Collins DL, Initiative ADN, et al. Prediction of alzheimer's disease in subjects with mild cognitive impairment from the adni cohort using patterns of cortical thinning. Neuroimage 2013;65:511–21. https://doi.org/10.1016/j.neuroimage.2012.09.058.
[38] Diciotti S, Ginestroni A, Bessi V, Giannelli M, Tessa C, Bracco L, Mascalchi M, Toschi N. Identification of mild alzheimer's disease through automated classification of structural mri features. In: 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, IEEE; 2012. p. 428–31. doi:10.1109/EMBC.2012.6345959.
[39] Ramaniharan AK, Manoharan SC, Swaminathan R. Laplace beltrami eigen value based classification of normal and alzheimer mr images using parametric and non-parametric classifiers. Expert Syst Appl 2016;59(C):208–16. https://doi.org/10.1016/j.eswa.2016.04.029.
[40] Gupta Y, Lee KH, Choi KY, Lee JJ, Kim BC, Kwon GR, N.R.C. for Dementia, A.D.N. Initiative, Early diagnosis of alzheimer's disease using combined features from voxel-based morphometry and cortical, subcortical, and hippocampus regions of mri t1 brain images. PloS One 2019;14(10):e0222446. doi:10.1371/journal.pone.0222446 .
[41] Vaithinathan K, Parthiban L, F. the Alzheimer's Disease Neuroimaging Initiative, A novel texture extraction technique with t1 weighted mri for the classification of alzheimer's disease. J Neurosci Methods 2019;318:84–99. doi:10.1016/j.jneumeth.2019.01.011.
[42] Wei JKE, Jahmunah V, Pham T-H, Oh SL, Ciaccio EJ, Acharya UR, Yeong CH, Fabell MKM, Rahmat K, Vijayananthan A, et al. Automated detection of alzheimer's disease using bi-directional empirical model decomposition. Pattern Recognit Lett. doi:10.1016/j.patrec.2020.03.014 .
[43] Gerardin E, Chételat G, Chupin M, Cuingnet R, Desgranges B, Kim H-S, Niethammer M, Dubois B, Lehéricy S, Garnero L, et al. Multidimensional classification of hippocampal shape features discriminates alzheimer's disease and mild cognitive impairment from normal aging. Neuroimage 2009;47(4):1476–86. https://doi.org/10.1016/j.neuroimage.2009.05.036.
[44] Plant C, Teipel SJ, Oswald A, Böhm C, Meindl T, Mourao-Miranda J, Bokde AW, Hampel H, Ewers M. Automated detection of brain atrophy patterns based on mri for the prediction of alzheimer's disease. Neuroimage 2010;50(1):162–74. https://doi.org/10.1016/j.neuroimage.2009.11.046.
[45] Westman E, Simmons A, Zhang Y, Muehlboeck J-S, Tunnard C, Liu Y, Collins L, Evans A, Mecocci P, Vellas B, et al.

Multivariate analysis of mri data for alzheimer's disease, mild cognitive impairment and healthy controls. Neuroimage 2011;54(2):1178–87. https://doi.org/10.1016/j.neuroimage.2010.08.044.
[46] Pachauri D, Hinrichs C, Chung MK, Johnson SC, Singh V. Topology-based kernels with application to inference problems in alzheimer's disease. IEEE Trans Med Imag 2011;30:1760–70. https://doi.org/10.1109/TMI.2011.2147327.
[47] Wolz R, Julkunen V, Koikkalainen J, Niskanen E, Zhang DP, Rueckert D, Soininen H, Lötjönen J, Initiative ADN, et al. Multi-method analysis of mri images in early diagnostics of alzheimer's disease. PloS One 2011;6(10) . https://doi.org/10.1371/journal.pone.0025446 e25446.
[48] Cuingnet R, Gerardin E, Tessieras J, Auzias G, Lehéricy S, Habert M-O, Chupin M, Benali H, Colliot O. Automatic classification of patients with Alzheimer's disease from structural MRI: a comparison of ten methods using the ADNI database. NeuroImage 2011;56(2):766–81. https://doi.org/10.1016/j.neuroimage.2010.06.013.
[49] Cho Y, Seong J-K, Jeong Y, Shin SY. Individual subject classification for alzheimer's disease based on incremental learning using a spatial frequency representation of cortical thickness data. NeuroImage 2012;59(3):2217–30. https://doi.org/10.1016/j.neuroimage.2011.09.085.
[50] Westman E, Muehlboeck J-S, Simmons A. Combining mri and csf measures for classification of alzheimer's disease and prediction of mild cognitive impairment conversion. Neuroimage 2012;62(1):229–38. https://doi.org/10.1016/j.neuroimage.2012.04.056.
[51] Aguilar C, Westman E, Muehlboeck J-S, Mecocci P, Vellas B, Tsolaki M, Kloszewska I, Soininen H, Lovestone S, Spenger C, et al. Different multivariate techniques for automated classification of mri data in alzheimer's disease and mild cognitive impairment. Psychiatr Res Neuroimag 2013;212(2):89–98. https://doi.org/10.1016/j.pscychresns.2012.11.005.
[52] Wee C-Y, Yap P-T, Shen D, Initiative ADN. Prediction of alzheimer's disease and mild cognitive impairment using cortical morphological patterns. Human Brain Map 2013;34(12):3411–25. https://doi.org/10.1002/hbm.22156.
[53] Ahmed OB, Benois-Pineau J, Allard M, Amar CB, Catheline G. Classification of alzheimer's disease subjects from mri using hippocampal visual features. Multimed Tools Appl 2014;74:1249–66. https://doi.org/10.1007/s11042-014-2123-y.
[54] Zhou Q, Goryawala M, Cabrerizo M, Wang J, Barker W, Loewenstein DA, Duara R, Adjouadi M. An optimal decisional space for the classification of alzheimer's disease and mild cognitive impairment. IEEE Trans Biomed Eng 2014;61(8):2245–53. https://doi.org/10.1109/TBME.2014.2310709.
[55] Farhan S, Fahiem MA, Tauseef H. An ensemble-of-classifiers based approach for early diagnosis of alzheimer's disease: classification using structural features of brain images. Hindawi 2014. https://doi.org/10.1155/2014/862307.
[56] Min R, Cheng J, Price T, Wu G, Shen D. Maximum-margin based representation learning from multiple atlases for alzheimer's disease classification. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer; 2014. p. 212–9.
[57] Tong T, Wolz R, Gao Q, Guerrero R, Hajnal JV, Rueckert D, Initiative ADN, et al. Multiple instance learning for classification of dementia in brain mri. Med Image Anal 2014;18(5):808–18. https://doi.org/10.1016/j.media.2014.04.006.
[58] Chyzhyk D, Savio A, Graña M. Evolutionary elm wrapper feature selection for alzheimer's disease cad on anatomical brain mri. Neurocomputing 2014;128:73–80. https://doi.org/10.1016/j.neucom.2013.01.065.
[59] Bron EE, Smits M, Niessen WJ, Klein S. Feature selection based on the svm weight vector for classification of dementia. IEEE J Biomed Health Inf 2015;19(5):1617–26. https://doi.org/10.1109/JBHI.2015.2432832.
[60] Moradi E, Pepe A, Gaser C, Huttunen H, Tohka J. Machine learning framework for early mri-based alzheimer's conversion prediction in mci subjects. NeuroImage 2015;104:398–412. https://doi.org/10.1016/j.neuroimage.2014.10.002.
[61] Cheng B, Liu M, Zhang D, Munsell BC, Shen D. Domain transfer learning for mci conversion prediction. IEEE Trans Bio-med Eng 2015;62(7):1805–17. https://doi.org/10.1109/TBME.2015.2404809.
[62] Zhang Y, Wang S, Phillips P, Dong Z, Ji G, Yang J. Detection of alzheimer's disease and mild cognitive impairment based on structural volumetric mr images using 3d-dwt and wta-ksvm trained by psotvac. Biomed Signal Process Control 2015;21:58–73. https://doi.org/10.1016/j.bspc.2015.05.014.
[63] Liu M, Zhang D, Shen D. Relationship induced multi-template learning for diagnosis of alzheimer's disease and mild cognitive impairment. IEEE Trans Med Imag 2016;35(6):1463–74. https://doi.org/10.1109/TMI.2016.2515021.
[64] Tohka J, Moradi E, Huttunen H, Initiative ADN, et al. Comparison of feature selection techniques in machine learning for anatomical brain mri in dementia. Neuroinformatics 2016;14(3):279–96. https://doi.org/10.1007/s12021-015-9292-3.
[65] Beheshti I, Demirel H, Initiative ADN, et al. Feature-ranking-based alzheimer's disease classification from structural mri. Magn Reson Imag 2016;34(3):252–63. https://doi.org/10.1016/j.mri.2015.11.009.
[66] Hu K, Wang Y, Chen K, Hou L, Zhang X. Multi-scale features extraction from baseline structure mri for mci patient classification and ad early diagnosis. Neurocomputing 2016;175:132–45. https://doi.org/10.1016/j.neucom.2015.10.043.
[67] Liu M, Zhang D, Adeli E, Shen D. Inherent structure-based multiview learning with multitemplate feature representation for alzheimer's disease diagnosis. IEEE Trans Biomed Eng 2015;63(7):1473–82. https://doi.org/10.1109/TBME.2015.2496233.
[68] Kim J, Lee B. Automated discrimination of dementia spectrum disorders using extreme learning machine and structural t1 mri features. In: 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE; 2017. p. 1990–3. doi:10.1109/embc.2017.8037241.
[69] Beheshti I, Demirel H, Matsuda H, Initiative ADN, et al. Classification of alzheimer's disease and prediction of mild cognitive impairment-to-alzheimer's conversion from structural magnetic resource imaging using feature ranking and a genetic algorithm. Comput Biol Med 2017;83:109–19. https://doi.org/10.1016/j.compbiomed.2017.02.011.
[70] Liu J, Wang J, Hu B, Wu F-X, Pan Y. Alzheimer's disease classification based on individual hierarchical networks constructed with 3-d texture features. IEEE Trans Nanobiosci 2017;16(6):428–37. https://doi.org/10.1109/TNB.2017.2707139.
[71] Lama RK, Gwak J, Park J-S, Lee S-W. Diagnosis of alzheimer's disease based on structural mri images using a regularized extreme learning machine and pca features. J Healthcare Eng 2017. https://doi.org/10.1155/2017/5485080.
[72] Nanni L, Lumini A, Zaffonato N. Ensemble based on static classifier selection for automated diagnosis of mild cognitive impairment. J Neurosci Methods 2018;302:42–6. https://doi.org/10.1016/j.jneumeth.2017.11.002.
[73] Sørensen L, Nielsen M, Initiative ADN, et al. Ensemble support vector machine classification of dementia using

structural mri and mini-mental state examination. J Neurosci Methods 2018;302:66–74. https://doi.org/10.1016/j.jneumeth.2018.01.003.
[74] Hojjati SH, Ebrahimzadeh A, Khazaee A, Babajani-Feremi A, Initiative ADN, et al. Predicting conversion from mci to ad by integrating rs-fmri and structural mri. Comput Biol Med 2018;102:30–9. https://doi.org/10.1016/j.compbiomed.2018.09.004.
[75] Shaikh TA, Ali R. Automated atrophy assessment for alzheimer's disease diagnosis from brain mri images. Magn Reson Imag 2019;62:167–73. https://doi.org/10.1016/j.mri.2019.06.019.
[76] Popuri K, Ma D, Wang L, Beg MF. Using machine learning to quantify structural mri neurodegeneration patterns of alzheimer's disease into dementia score: Independent validation on 8,834 images from adni, aibl, oasis, and miriad databases. Human Brain Map 2020;41(14):4127–47. https://doi.org/10.1002/hbm.25115.
[77] Zhang Y, Wang S, Xia K, Jiang Y, Qian P, Initiative ADN, et al. Alzheimer's disease multiclass diagnosis via multimodal neuroimaging embedding feature selection and fusion. Inf Fusion 2021;66:170–83. https://doi.org/10.1016/J.INFFUS.2020.09.002.
[78] Bi X, Jiang Q, Sun Q, Shu Q, Liu Y. Analysis of alzheimer's disease based on the random neural network cluster in fmri. Front Neuroinf 2018;12:60. https://doi.org/10.3389/fninf.2018.00060.
[79] Li Q, Wu X, Xu L, Chen K, Yao L, Initiative ADN, et al. Classification of alzheimer's disease, mild cognitive impairment, and cognitively unimpaired individuals using multi-feature kernel discriminant dictionary learning. Front Comput Neurosci 2018;11:117. https://doi.org/10.3389/fncom.2017.00117.
[80] Challis E, Hurley P, Serra L, Bozzali M, Oliver S, Cercignani M. Gaussian process classification of alzheimer's disease and mild cognitive impairment from resting-state fmri. NeuroImage 2015;112:232–43. https://doi.org/10.1016/j.neuroimage.2015.02.037.
[81] Acharya UR, Fernandes SL, WeiKoh JE, Ciaccio EJ, Fabell MKM, Tanik UJ, Rajinikanth V, Yeong CH. Automated detection of alzheimer's disease using brain mri images–a study with various feature extraction techniques. J Med Syst 2019;43(9):302. https://doi.org/10.1007/s10916-019-1428-9.
[82] Islam J, Zhang Y. Brain mri analysis for alzheimer's disease diagnosis using an ensemble system of deep convolutional neural networks. Brain Inf 2018;5(2):2. https://doi.org/10.1186/s40708-018-0080-3.
[83] Ardekani BA, Bachman AH, Figarsky K, Sidtis JJ. Corpus callosum shape changes in early alzheimer's disease: an mri study using the oasis brain database. Brain Struct Funct 2014;219(1):343–52. https://doi.org/10.3233/JAD-131526.
[84] Zhang J, Yu C, Jiang G, Liu W, Tong L. 3d texture analysis on mri images of alzheimer's disease. Brain Imag Behav 2012;6(1):61–9. https://doi.org/10.1007/s11682-011-9142-3.
[85] Lee S, Kim KW, A.D.N. Initiative, Associations between texture of t1-weighted magnetic resonance imaging and radiographic pathologies in alzheimer's disease. Eur J Neurol doi:10.1111/ene.14609 .
[86] Schuff N, Woerner N, Boreta L, Kornfield T, Shaw L, Trojanowski J, Thompson P, Jack Jr C, Weiner M, Initiative ADN. Mri of hippocampal volume loss in early alzheimer's disease in relation to apoe genotype and biomarkers. Brain 2009;132(4):1067–77. https://doi.org/10.1093/brain/awp007.
[87] Chupin M, Chetelat G, Lemieux L, Dubois B, Garnero L, Benali H, Eustache F, Lehericy S, Desgranges B, Colliot O. Fully automatic hippocampus segmentation discriminates between early alzheimer's disease and normal aging. In: 2008 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro. IEEE; 2008. p. 97–100. doi:10.1109/ISBI.2008.4540941.
[88] Morra JH, Tu Z, Apostolova LG, Green AE, Avedissian C, Madsen SK, Parikshak N, Toga AW, Jack Jr CR, Schuff N, et al. Automated mapping of hippocampal atrophy in 1-year repeat mri data from 490 subjects with alzheimer's disease, mild cognitive impairment, and elderly controls. Neuroimage 2009;45(1):S3–S15. https://doi.org/10.1016/j.neuroimage.2008.10.043.
[89] Costafreda SG, Dinov ID, Tu Z, Shi Y, Liu C-Y, Kloszewska I, Mecocci P, Soininen H, Tsolaki M, Vellas B, et al. Automated hippocampal shape analysis predicts the onset of dementia in mild cognitive impairment. Neuroimage 2011;56(1):212–9. https://doi.org/10.1016/j.neuroimage.2011.01.050.
[90] Li Y, Liu Y, Wang P, Wang J, Xu S, Qiu M, A.D.N.I. (ADNI), et al., Dependency criterion based brain pathological age estimation of alzheimer's disease patients with mr scans. Biomed Eng Online 2017;16(1):50. doi:10.1186/s12938-017-0342-y.
[91] Liu Y, Julkunen V, Paajanen T, Westman E, Wahlund L-O, Aitken A, Sobow T, Mecocci P, Tsolaki M, Vellas B, et al. Education increases reserve against alzheimer's disease—evidence from structural mri analysis. Neuroradiology 2012;54(9):929–38. https://doi.org/10.1007/s00234-012-1005-0.
[92] Apostolova LG, Hwang KS, Kohannim O, Avila D, Elashoff D, Jack Jr CR, Shaw L, Trojanowski JQ, Weiner MW, Thompson PM, et al. Apoe4 effects on automated diagnostic classifiers for mild cognitive impairment and alzheimer's disease. NeuroImage: Clinical 2014;4:461–72. https://doi.org/10.1016/j.nicl.2013.12.012.
[93] Ahmed OB, Mizotin M, Benois-Pineau J, Allard M, Catheline G, Amar CB, Initiative ADN, et al. Alzheimer's disease diagnosis on structural mr images using circular harmonic functions descriptors on hippocampus and posterior cingulate cortex. Comput Med Imag Graph 2015;44:13–25. https://doi.org/10.1016/j.compmedimag.2015.04.007.
[94] Huang L, Pan Z, Lu H, et al., Automated diagnosis of alzheimer's disease with degenerate svm-based adaboost. In: 2013 5th International Conference on Intelligent Human-Machine Systems and Cybernetics, vol. 2. IEEE; 2013. p. 298–301. doi:10.1109/IHMSC.2013.219.
[95] Kuncheva LI, Whitaker CJ. Measures of diversity in classifier ensembles and their relationship with the ensemble accuracy. Mach Learn 2003;51(2):181–207. https://doi.org/10.1023/A:1022859003006.
[96] Rokach L. Ensemble-based classifiers. Artif Intell Rev 2010;33(1–2):1–39. https://doi.org/10.1007/s10462-009-9124-7.
[97] Bron EE, Smits M, Niessen WJ, Klein S. Feature selection based on the svm weight vector for classification of dementia. IEEE J Biomed Health Inf 2015;19(5):1617–26. https://doi.org/10.1109/JBHI.2015.2432832.
[98] Jiang W, Zavesky E, Chang S-F, Loui A. Cross-domain learning methods for high-level visual concept classification. In: 2008 15th IEEE International Conference on Image Processing. IEEE; 2008. p. 161–4. doi:10.1109/ICIP.2008.4711716 .
[99] Ahmed OB, Benois-Pineau J, Allard M, Catheline G, Amar CB, Initiative ADN, et al. Recognition of alzheimer's disease and mild cognitive impairment with multimodal image-derived biomarkers and multiple kernel learning. Neurocomputing 2017;220:98–110. https://doi.org/10.1016/j.neucom.2016.08.041.
[100] Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, Van Der Laak JA, Van Ginneken B, Sánchez CI. A survey on deep learning in medical image analysis. Med Image Anal 2017;42:60–88. https://doi.org/10.1016/j.media.2017.07.005.

[101] Shen D, Wu G, Suk H-I. Deep learning in medical image analysis. Annu Rev Biomed Eng 2017;19:221–48. https://doi.org/10.1146/annurev-bioeng-071516-044442.
[102] Jo T, Nho K, Saykin AJ. Deep learning in alzheimer's disease: diagnostic classification and prognostic prediction using neuroimaging data. Front Aging Neurosci 2019;11:220. https://doi.org/10.3389/fnagi.2019.00220.
[103] Gupta A, Ayhan M, Maida A. Natural image bases to represent neuroimaging data. In: International Conference on Machine Learning; 2013. p. 987–94 .
[104] Brosch T, Tam R, A.D.N. Initiative, et al. Manifold learning of brain mris by deep learning. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer; 2013. p. 633–40. doi:10.1007/978-3-642-40763-5_78.
[105] Hinton GE, Osindero S, Teh Y-W. A fast learning algorithm for deep belief nets. Neural Comput 2006;18(7):1527–54. https://doi.org/10.1162/neco.2006.18.7.1527.
[106] Payan A, Montana G. Predicting alzheimer's disease: a neuroimaging study with 3d convolutional neural networks. arXiv preprint arXiv:1502.02506 .
[107] Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. Going deeper with convolutions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. p. 1–9. https://doi.org/10.1109/CVPR.2015.7298594.
[108] Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, Huang Z, Karpathy A, Khosla A, Bernstein M, et al. Imagenet large scale visual recognition challenge. Int J Comput Vis 2015;115(3):211–52.
[109] Farooq A, Anwar S, Awais M, Rehman S. A deep cnn based multi-class classification of alzheimer's disease using mri. In: 2017 IEEE International Conference on Imaging Systems and Techniques (IST). IEEE; 2017. p. 1–6. https://doi.org/10.1109/IST.2017.8261460.
[110] Korolev S, Safiullin A, Belyaev M, Dodonova Y. Residual and plain convolutional neural networks for 3d brain mri classification. In: 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017). IEEE; 2017. p. 835–8. doi:10.1109/ISBI.2017.7950647.
[111] Hosseini-Asl E, Ghazal M, Mahmoud A, Aslantas A, Shalaby AM, Casanova MF, Barnes G, Gimel'farb GL, Keynton RS, El-Baz AS. Alzheimer's disease diagnostics by a 3d deeply supervised adaptable convolutional network. Front Biosci 2018;23:584–96. https://doi.org/10.2741/4606.
[112] Hosseini-Asl E, Keynton R, El-Baz A. Alzheimer's disease diagnostics by adaptation of 3d convolutional network. In: 2016 IEEE International Conference on Image Processing (ICIP). IEEE; 2016. p. 126–30. https://doi.org/10.1109/ICIP.2016.7532332.
[113] Wang S-H, Phillips P, Sui Y, Liu B, Yang M, Cheng H. Classification of alzheimer's disease based on eight-layer convolutional neural network with leaky rectified linear unit and max pooling. J Med Syst 2018;42(5):85. https://doi.org/10.1007/s10916-018-0932-7.
[114] Suk H-I, Lee S-W, Shen D. Hierarchical feature representation and multimodal fusion with deep learning for ad/mci diagnosis. NeuroImage 2014;101:569–82. https://doi.org/10.1016/j.neuroimage.2014.06.077.
[115] Sarraf S, Tofighi G. Classification of alzheimer's disease structural mri data by deep learning convolutional neural networks. arXiv preprint arXiv:1607.06583 .
[116] Lian C, Liu M, Zhang J, Shen D. Hierarchical fully convolutional network for joint atrophy localization and alzheimer's disease diagnosis using structural mri. IEEE Trans Pattern Anal Mach Intell. doi:10.1109/TPAMI.2018.2889096 .
[117] Basaia S, Agosta F, Wagner L, Canu E, Magnani G, Santangelo R, Filippi M, A.D.N. Initiative, et al., Automated classification of alzheimer's disease and mild cognitive impairment using a single mri and deep neural networks. NeuroImage: Clinical 2019;21:101645. doi:10.1016/j.nicl.2018.101645.
[118] Oh K, Chung Y-C, Kim KW, Kim W-S, Oh I-S. Classification and visualization of alzheimer's disease using volumetric convolutional neural network and transfer learning. Sci Rep 2019;9(1):1–16. https://doi.org/10.1038/s41598-019-54548-6.
[119] Cho J, Lee K, Shin E, Choy G, Do S. How much data is needed to train a medical image deep learning system to achieve necessary high accuracy?. arXiv preprint arXiv:1511.06348 .
[120] Castro AP, Fernandez-Blanco E, Pazos A, Munteanu CR. Automatic assessment of alzheimer's disease diagnosis based on deep learning techniques. Comput Biol Med 2020 . https://doi.org/10.1016/j.compbiomed.2020.103764 103764.
[121] Ebrahimi-Ghahnavieh A, Luo S, Chiong R. Transfer learning for alzheimer's disease detection on mri images. In: 2019 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology (IAICT). IEEE; 2019. p. 133–8. doi:10.1109/ICIAICT.2019.8784845 .
[122] Ramzan F, Khan MUG, Rehmat A, Iqbal S, Saba T, Rehman A, Mehmood Z. A deep learning approach for automated diagnosis and multi-class classification of alzheimer's disease stages using resting-state fmri and residual neural networks. J Med Syst 2020;44(2):37. https://doi.org/10.1007/s10916-019-1475-2.
[123] Shin H-C, Roth HR, Gao M, Lu L, Xu Z, Nogues I, Yao J, Mollura D, Summers RM. Deep convolutional neural networks for computer-aided detection: Cnn architectures, dataset characteristics and transfer learning. IEEE Trans Med Imag 2016;35(5):1285–98. https://doi.org/10.1109/TMI.2016.2528162.
[124] Lazli L, Boukadoum M, Mohamed OA. A survey on computer-aided diagnosis of brain disorders through mri based on machine learning and data mining methodologies with an emphasis on alzheimer disease diagnosis and the contribution of the multimodal fusion. Appl Sci 2020;10(5):1894. https://doi.org/10.3390/app10051894.
[125] Spasov S, Passamonti L, Duggento A, Liò P, Toschi N, Initiative ADN, et al. A parameter-efficient deep learning approach to predict conversion from mild cognitive impairment to alzheimer's disease. Neuroimage 2019;189:276–87. https://doi.org/10.1016/j.neuroimage.2019.01.031.
[126] Akl A, Taati B, Mihailidis A. Autonomous unobtrusive detection of mild cognitive impairment in older adults. IEEE Trans Biomed Eng 2015;62(5):1383–94. https://doi.org/10.1109/TBME.2015.2389149.