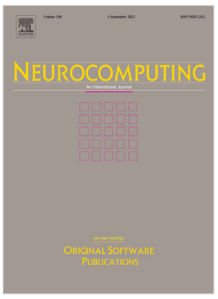
Journal Pre-proof



Deep learning methods for early detection of Alzheimer's disease using structural MR images: A survey

Sonia Ben Hassen, Mohamed Neji, Zain Hussain, Amir Hussain, Adel M. Alimi, Mondher Frikha

PII: S0925-2312(24)00096-1
DOI: https://doi.org/10.1016/j.neucom.2024.127325
Reference: NEUCOM 127325

To appear in: *Neurocomputing*

Received date : 6 October 2023
Revised date : 18 December 2023
Accepted date : 21 January 2024

Please cite this article as: S.B. Hassen, M. Neji, Z. Hussain et al., Deep learning methods for early detection of Alzheimer's disease using structural MR images: A survey, *Neurocomputing* (2024), doi: https://doi.org/10.1016/j.neucom.2024.127325.

Deep learning methods for early detection of Alzheimer's disease using structural MR images: a survey⋆

Sonia Ben Hassen[a,<,1], Mohamed Neji[b], Zain Hussain[c], Amir Hussain[d], Adel M. Alimi[e] and Mondher Frikha[f]

[a]*Advanced Technologies for Image and Signal Processing Unit, National School of Electronics and Telecommunications of Sfax, Tunisia , National School of Electronics and Telecommunications of Sfax, Sfax, Tunis road km 10, 3018, Tunisia*

[b]*National School of Electronics and Telecommunications of Sfax Technopark, Tunis road km 10, 3018, Sfax, Tunisia*

[c]*University of Glasgow, United Kingdom*

[d]*School of Computing, Edinburgh Napier University, Edinburgh, United Kingdom*

[e]*Department of Electrical and Electronic Engineering Science, Faculty of Engineering and the Built Environment, University of Johannesburg, South Africa*

[a]*Advanced Technologies for Image and Signal Processing Unit, National School of Electronics and Telecommunications of Sfax, Tunisia , Tunisia*

ARTICLE INFO

Keywords:
Alzheimer's disease
Structural MRI
Deep Learning
Hippocampus
Progressive MCI
Stable MCI

ABSTRACT

In this paper, we present an extensive review of the most recent works for Alzheimer's disease (AD) prediction, particularly Moderate Cognitive Impairment (MCI) conversion prediction. We aimed to identify the most useful brain-magnetic resonance imaging (MRI) biomarkers as well as the most successful deep learning frameworks used for prediction. To achieve this, we analysed more than 130 works and 7 review articles. A closer look revealed that the hippocampus is an important region of interest (ROI) is affected early by AD and that many related features help detect the disease in its early stages. However, considered alone, this ROI is not sufficient enough to ensure high prediction performance. Therefore, many other brain regions can also provide additional information to improve prediction accuracy. In relation to state-of-the-art deep neural networks, the U-Net represents the most efficient architecture for hippocampus segmentation. The best Dice Similarity Coefficient (DSC) value, equal to 94%, was achieved by the RESU-Net architecture. The best results for MCI conversion prediction were obtained for two models that identify significant landmarks from the whole brain for classification. The multi-stream convolutional neural network achieved the best AUC and specificity of 94.39% and 99.70%, respectively. Finally, a region ensemble model delivered the best accuracy of 85.90%, demonstrating the need for further research to address this challenging problem.

1. Introduction

Dementia, also known as major neurocognitive disorder, is a group of symptoms that can be caused by a number of diseases which over time leads to a loss of cognitive functioning. It typically affects older people, with symptoms worsening over time. Signs include memory problems, difficulty concentrating, language and communication problems and changes in personality and/or behaviour [1]. According to the World Health Organization (WHO), approximately 90 million people are living with dementia worldwide, and this number is expected to increase to 150 million in 2030 and 300 million by 2050, due to population aging [1]. The underlying causes of dementia can involve neurodegeneration, which is the slow and progressive loss of neuronal cells in specified regions of the brain which results in impaired decision making and memory amongst others. Non-neurodegenerative mechanisms include vascular dementia [1]. Alzheimer's disease (AD)

[<]Corresponding author

sonia.benhassen@ieee.org (S.B. Hassen); mohamed.neji@ieee.org (M. Neji); zain.hussain@glasgow.ac.uk (Z. Hussain); hussain.doctor@gmail.com (A. Hussain); adel.alimi@regim.usf.tn (A.M. Alimi); mondher.frikha@enetcom.usf.tn (M. Frikha)

ORCID(s): 0000-0003-4997-9398 (S.B. Hassen)

[1]

is the most common cause of dementia, affecting over 50 million people worldwide [1],[2] and constituting 60-70% of total cases of dementia [3]. AD is characterized by progressive and irreversible neurodegeneration, leading to cognitive and memory impairments and eventually death. A transitional state between normal cognition and the preclinical phase of dementia is referred to as mild cognitive impairment (MCI), where patients show objective deficits and moderate cognitive problems. MCI can represent the prodromal stage of AD. Some previous studies have found that individuals with MCI progress to AD at a rate of approximately 5–25% per year [4]. MCI can be classified, according to its various clinical outcomes, into: progressive MCI (pMCI) and stable MCI (sMCI). An open research challenge is the early and accurate diagnosis of AD, and the detection of early stages, including MCI. Clinical symptoms of AD predominantly appear in advanced stages, i.e., many years after the onset of the processes causing this disease. Therefore, identifying biomarkers for differentiating between individuals with MCI at high risk of conversion to AD dementia versus stable MCI patients is of great importance. Early diagnosis allows for optimal patient care, and such patients can help us develop and trial novel treatments to slow disease progression and one day reverse it. Currently, there are different measurements used for the diagnosis of AD. In addition to clinical assessment and cognitive evaluation using neuropsychological tests, diagnosis relies on Electroencephalographic (EEG) signals [5] and various modalities of neuroimaging. These modalities include T1-weighted anatomical magnetic resonance imaging (MRI) [6], functional MRI [7], positron emission tomography (PET) scans (especially 18F 2-fluoro-2-deoxy- D-glucose (FDG), Amyloid- and tau) [8]-[9], single-photon emission computed tomography (SPECT) [1],[10], diffusion tensor imaging (DTI) [11] and cerebrospinal fluid (CSF) or plasma amyloid- /tau. These different neuroimaging modalities were widely used to identify heterogeneous functional and structural changes at brain level for AD patients including the amyloid plaques, hypo-metabolism [12], and neurological atrophy. In this context, diverse structural biomarkers have been explored to early detect and differentiate between sMCI and pMCI. These biomarkers include cortical thinning, high-dimensional features based on grey matter (GM) and white matter (WM) volume distribution of whole brain [13], as well as atrophy and shape deformity of individual structures such as carpus callosum (CC), hippocampus (HC) and ventricle [14]. In fact, in AD, the tissue loss starts from the GM and it extends over to WM, CC and HC. The rapid damage of tissues leads to structural changes, such as ventricle enlargement, reduction in size of HC and GM, in the brain [14]. Numerous studies have been published on diagnosing AD using the different neuroimaging modalities. Some of them have exploited one modality (particularly PET and MRI) while some others have combined multi modalities in order to provide different information about AD and then increase classification performances [15]. PET imaging has been shown to be efficient for the early diagnosis of AD as functional brain abnormalities are readily detectable in prodromal stage of AD using this technique. In fact, blood flow, glucose metabolism and oxygen are easily visible in the working brain tissues, thus reflecting brain activity and providing insight into its function. Despite the efficiency of PET imaging, it is in general more invasive and more expensive than other modalities. Moreover, it involves injection of a radioactive tracer. Therefore, many researchers were resorted to study MRI imaging. Indeed, MRI is characterized by its high spatial resolution and its sensitivity to the shape and volume of the brain. Therefore, it is able to identify the structural atrophy in early stages of AD. Moreover, unlike PET and SPECT techniques, this neuroimaging modality does not involve radioactive pharmaceutical injection during acquisition process and thereby it is safer for repetitive use in tracking development of the disease. In this context, several MRI-based approaches have been proposed for the early AD prediction.

Otherwise, over the past few decades, artificial intelligence (AI) has been successfully applied in many domains such as robotics, education, and health care [16]. In particular, for the healthcare domain, AI techniques ranging from machine learning to deep learning are mainly used for the purpose of disease diagnosis by providing and intelligent health systems for both common and rare diseases [16]. For example, P.Wu, et al. proposed, in [17], a novel attention-based glioma grading network (AGGN). This method included three meticulously designed modules which are a dual-domain attention module, a multi-scale feature extraction and a multi-modal information fusion one [17]. Thus, it enhanced the ability to distinguish intra-class variability and inter-class similarity existing at different grades of glioma by efficiently capturing potential correlations and key information from scattered features in different imaging modalities [17]. Therefore, it succeeded to achieve good performances even without the manual labeled tumor masks. Moreover, H. Li, et al., developed, in [18], a Cov-Net-based model for accurate recognition of COVID-19 from chest X-ray images. This model is able to build robust and highly discriminative features by implementing

a meticulously designed attention mechanism and a set of skip-connected dilated convolution kernels with different dilation rates [18]. In particular, different machine learning (ML) methods have been progressively developed for the diagnosis and prognosis of AD using neuroimaging modalities. Many types of classifiers such as support vector machines (SVM), linear discriminant analysis (LDA), K-nearest neighbours (KNN), and random forests have been frequently used to identify patterns in data that differentiate between several classes in dementia [19]-[22]. SVM has been shown to be the most efficient for AD as well as MCI classification [23]. Moreover, N. Zeng, et al. introduced, in [24], a new switching delayed particle swarm optimization (SDPSO) algorithm to optimize the SVM model. It used principal component analysis (PCA) for feature extraction and then applied the SDPSO-SVM method to ensure the classification [Y4]. The proposed approach achieved excellent classification accuracies and that outperformed several SVM algorithms [24]. Recently, motivated by the rapid development of computer science and the accumulation of clinical data, deep learning (DL), a computationally powerful tool, has become very popular in the AD research. Particularly, DL techniques have been extensively exploited to extract high-level features and then performing the AD diagnosis [23]. Some of DL studies have limited to differentiate between AD patients and healthy controls and they have shown high performances. While others have focused on the most challenging problem which is to differentiate between sMCI and pMCI of the prodromal stage of AD. For example, N. Zeng, et al., proposed in, [25], a novel deep belief network (DBN) based multi-task learning algorithm for AD prediction. This new algorithm applied, in addition to the PCA, a multi-task feature selection approach to reduce the feature dimension [25]. Then, it exploited the DBN to achieve the classification task. Moreover, to get over the overfitting problem and also enhance the generalization ability and robustness of the model, it used many techniques such as the dropout technology and zero-masking strategy [25]. The proposed approach achieved satisfactory results in multiple binary classification tasks.

There are a number of reviews in the literature on diagnosing AD using DL [[26]-[31]]. They have evaluated the main DL models developed for the early prediction of AD using one or more neuroimaging modalities. Some reviews have focused more on the main efficient DL architectures [26]-[29] whilst others have been interested in the neuroimaging techniques and features used to ensure a good prediction [30],[31].

In this paper, we provide a detailed overview of the most important recent papers, from 2018 to 2022, that investigate and analyze main methods used to for the early prediction of AD. We limited our survey on the works based on structural MR images because many interesting results and conclusions, about the contribution of this neuro-image modality on the early AD prediction, have been explored. Moreover, it is the most used modality in our country for the diagnosis of AD disease in addition to the other advantages mentioned above.

In summary, the main objectives of this review are to:

- Provide the important biomarkers for structural MR images used to detect whether an MCI subject will convert to AD or not;
- Discuss the main DL architectures developed to achieve this challenging task;
- Provide the most recent architectures proposed in the literature for the segmentation of the full hippocampus as well as its subfields;
- Present the important features extracted from the hippocampus for the early AD prediction;
- Show the limited contribution of the hippocampus considered alone and the importance of other zones identified by significant landmarks in the MRI images to improve the prediction performances.

This review is organized as follows: Section II presents an overview of neuroimaging modalities exploited for AD diagnosis. In Section III, we define the datasets used in the literature. In Section IV, we outline the general pipeline of an AD diagnosis system. Section V is dedicated to presenting and analyzing different methods developed for the early prediction of AD patients, including those based on the segmentation of the hippocampus and others using significant zones identified from the whole brain. Section VI is devoted to discussion and future directions. The paper is closed by a conclusion.

2. Overview of neuroimaging modalities

With the increasing of MCI and its potential conversion to AD, the structure and metabolic rate of the brain change accordingly. The degenerative histological changes include cerebral glucose hypo-metabolism, amyloid plaques, and shrinkage of cerebral cortices (transentorhinal cortex and neocortex). These changes could be visualized with the help of biomedical imaging techniques such as MRI, PET and SPECT [3],[32]. These data modalities were shown to be relevant for identifying MCI and AD. Images can be divided into structural imaging including sMRI, CT, functional imaging containing fMRI, PET and SPECT. With these different imaging modalities, the ability to automatically quantify brain morphology, function, connectivity, and pathology has improved considerably.

2.1. Anatomical MRI

MRI is a non-invasive widely accessible tool that can be used in vivo to identify sensitive and specific markers for early AD diagnosis. Particularly, 3D T1-weighted volumetric-MRI gives high-resolution information about the brain structure. In fact, it provides accurate measurements of structural metrics like thickness, volume, surface area and shape for the whole brain or for ROIs (especially the hippocampus (HC) and the entorhinal cortex [33]) to identify brain atrophy and evaluate its degree based on various visual rating scales. Therefore, this allows not only to separate healthy subjects from AD patients with high precision but also to differentiate pMCI and sMCI for early AD diagnosis. Indeed, based on the Braak study for AD staging [34], the disease begins from the transentorhinal cortex (stages I and II), progressing then to the hippocampus (stages III and IV), and finally involving the neocortex (stage V) [35]. These stages were defined based on the changes of accumulation of the neurofibrillary tangles as well as the progression of brain atrophy according to diverse MRI studies [35]. Consequently, several works have focused on HC as a very important ROI and many efforts have been directed to perform accurate segmentation of the HC for subjects with MCI to extract significant HC subfields and then properly assess relevant geometrical features helping the good prediction of the AD. In section V, we will discuss the most recent works that use this modality for the early detection of AD.

2.2. Functional MRI

Functional MRI (fMRI) represents a non-invasive method to analyze subtle functional connectivity between different brain regions. In fact, it provides a measure of changes in blood oxygen levels that is an indicator of neuronal activity. fMRI is usually acquired during cognitive tasks or also during a resting state (so called rs-fMRI or task-free fMRI (TF- fMRI)). Also, fMRI sequence can be obtained easily during routine clinical structural MRI sessions. Rs-fMRI is emerging as an interesting biomarker that becomes increasingly popular for the AD diagnosis. Several studies have assessed the effects of AD onto brain resting-state activity using rs-fMRI. These studies have proved that the posterior cingulate, precuneus, medial prefrontal cortex and the hippocampus are the main regions of the Default Mode Network (DMN), a distributed network of cortical regions that characterizes the resting state of the human brain, involved in AD [36]. In fact, they have shown that AD causes the decrease of DMN connectivity in addition to the decrease functional synchrony in the hippocampus and its connectivity to the rest of the DMN [36]. Several ML/DL approaches have been developed for the diagnosis of AD from rs-fMRI [36]. They have achieved high accuracies especially for the binary classification of AD vs CN and MCI vs CN. Despite their good performances, they were performed in very small samples (around 20 participants per group). Therefore, there is a need to evaluate these techniques across large, multicenter, datasets.

2.3. Diffusion MRI/DTI

Diffusion MRI or diffusion tensor imaging (DTI) is a non- invasive technology that represents the white matter (WM) fiber bundles in the brain. It allows the analysis of water diffusion at the microstructural level of the brain. This imaging modality is sensitive to the microstructural damage that may be present in the white matter bundles and offers different measures of its integrity [11],[37]. It can also be used to measure the anatomical connectivity between brain regions and detect injuries and signs of impairment that are not detectable with standard anatomical MRI [11],[37]. Therefore, it has become one of the most popular MRI techniques in brain research, as well as in clinical practice [38]-[40]. Particularly, DTI has offered a valuable tool for both the detection of AD and its early stage. In fact, it has contributed to the search for early

biomarkers of the disease [37]. Several ML-based studies have been proposed for the diagnosis of AD and its early-stage MCI using DTI [41]-[52]. They have shown that AD and MCI patients exhibit aberrant fractional anisotropy (FA) and mean diffusivity (MD) values (low FA and high MD) in the white matter of specific cerebral regions, especially the hippocampus and the amygdala, with less severity for MCI [11]. They also showed that measures of diffusivity extracted from the hippocampus are better predictors of MCI conversion to AD than its volume [44].

2.4. PET tracers

Positron Emission Tomography (PET) is a functional imaging technique used to explore brain functional activity associated with different cognitive states. It provides a representation of a specific metabolic process through the detection of a positron-emitting isotope tied to a biologically active molecule [53]. The analysis of the resulted phenomenon allows the detection of cognitive abnormalities in the brain. The well-known radioactive PET tracers used for diagnosing dementia are [18F] FDG PET, Amyloid- PET, Tau PET and SV2A PET.

PET represents the second most-widely used modality after anatomical MRI in AD-related ML/DL studies. The main benefit of this technique is its ability to detect the earliest pathological changes for AD (for example amyloid (A) deposition and tau hyperphosphorylation). In fact, the first stages of the disease are characterized by the loss of synapses and the depositions of certain lesions such as extracellular A amyloid plaques and intracellular tau neurofibrillary tangles in many regions of the brain [35],[54]. Moreover, blood flow, oxygen and glucose metabolism in the working brain tissues are easily visible, thus reflecting brain activity and providing insight into its function [1]. More details for the FPG PET tracer and the others are given in [55]. Although PET tracers become standard tools for clinicians allowing early diagnosis of many neurodegenerative diseases (especially AD and MCI) by providing key biomarkers, they have many limitations. In fact, hypometabolic patterns, for example, overlap between multiple neurodegenerative diseases. A positivity can also refer to various neurodegenerative diseases [55]. Furthermore, the PET process itself (isotope production, radiochemistry, scan) is expensive compared to MRI [55].

2.5. SPECT tracers

Single-Photon Emission Computed Tomography (SPECT) imaging is a functional nuclear imaging technique that aims to evaluate the regional cerebral perfusion of the brain using a specific radioactive tracer [1]. The analysis of captured two- or three-dimensional images allows the dementia diagnosis. In this context, some available biomarkers for AD in clinical use are the cerebral blood flow (CBF)-SPECT and the dopamine transporter single photon emission CT (DAT- SPECT) [1]. They assess cerebral perfusion and investigate the functional integrity of dopaminergic nigrostriatal, respectively, offering thus promising tools for diagnosis and prognosis of dementia [56]-[58]. Despite the importance of the different modalities in the diagnosis of AD, we will focus, in this survey, only on the anatomical (or structural) MR images. In fact, many significant results, about the contribution of this neuro-image modality on the early detection of the disease, have been discussed. Moreover, it is less expensive compared to other modalities, non-invasive and widely accessible tool. Furthermore, it is the most used modality in our country for the diagnosis of AD.

3. Datasets used in the literature

There exist several publicly available neuroimaging datasets that have been used in the literature for the diagnosis of AD and greatly supported the research in dementia. These datasets are the Alzheimer's Disease Neuroimaging Initiative (ADNI), the Open Access Series of Imaging Studies (OASIS), the Australian Imaging, Biomarker & Lifestyle Flagship Study of Ageing (AIBL), the Minimal Interval Resonance Imaging in Alzheimer's Disease (MIRIAD) and the CADDementia, structural brain MRI scans. The most frequently used dataset is the ADNI. It was initially launched in 2003 as a public private partnership by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies and non-profit organizations [59]. It is led by Principal Investigator Michael W. Weiner, MD [3]. ADNI represents a longitudinal multicentric project that contains studies of over a thousand subjects, recruited from over 50 sites across the U.S. and Canada, in different stages ranging from healthy controls (HC) and MCI to AD patients. The first ADNI

study, ADNI-1, was extended to ADNI-GO in 2009 by a Grand Opportunities grant, and further extended to ADNI- 2, and ADNI-3 in 2011 and 2016 respectively, by competitive renewals [60]. Each baseline ADNI dataset contains images from different neuroimaging modalities: anatomical MRI, FDG PET, amyloid PET, tau PET, diffusion MRI, and functional MRI captured for some subjects at multiple time points[1]. Data also includes extensive clinical and cognitive testing, as well as genetic data. Another frequently used dataset is OASIS. It is a series of MRI datasets stored in different collections. The initial dataset released in 2007 is OASIS-1 [61]. It consists of a cross-sectional collection of 416 subjects, aged from 18 to 96 years, including a group of hundred elderly subjects clinically diagnosed with very mild to moderate AD. The other OASIS dataset (OASIS-2) was released in 2010. It represents a longitudinal study of 150 older adults (including some with mild to moderate AD). The latest release in OASIS is OASIS-3. It represents a longitudinal neuroimaging (MRI and PET), clinical, cognitive, and biomarker dataset for 609 normal aging adults and 489 individuals at various stages of cognitive decline including AD. They were collected across several ongoing projects through the WUSTL Knight ADRC over the course of 30 years. Moreover, the AIBL dataset was launched in 2006 [62]. It includes approximately 1100 participants (AD patients, MCI patients and HC). It contains anatomical MRI and amyloid PET images in addition to clinical and cognitive scores. Particularly, the Australian ADNI, a subset of the AIBL dataset, is publicly available through the same technical infrastructure as the ADNI thanks to a collaborative agreement with, and funding from, the AD association. This subset was added directly to the ADNI cohort enhancing its scientific value. Furthermore, the MIRIAD dataset, a database of volumetric T1 MRI scans of 46 mild–moderate Alzheimer's subjects and 23 healthy elderly people, was established in 2013 [63]. It includes a total of 708 scans conducted by the same radiographer with the same scanner and sequences at intervals from 2 weeks to 2 years from baseline, with accompanying information on gender, age and Mini Mental State Examination (MMSE) scores [63]. Therefore, it is particularly interesting as a common resource for researchers to measure longitudinal volume change in serially acquired MR, understand the progression of the disease, and then provide longitudinal biomarkers [63]. In addition to these publicly available datasets, some researchers have used in their studies other multicenter research datasets. However, these datasets are not publicly available [53].
Table 1 contains a more comprehensive analysis of dataset characteristics and their impacts on model performances.

4. General pipeline of an AD diagnostic system

The AD diagnosis system is a succession of important steps starting from image pre-processing until the classification task.

4.1. Pre-processing

The pre-processing of MR images, coming from a magnetic resonance (MR) scanner, includes different techniques to improve the quality of brain scans. These techniques are mainly brain extraction, bias field correction, noise reduction and image registration. In fact, skull-stripping should be performed to remove 'nonbrain tissue' such as skull, neck, muscles and bones [64], [65]. Moreover, the inhomogeneity of image intensities, due to the strength of the magnetic field, should be corrected using, for example, a nonparametric non-uniform intensity normalization algorithm [64], [65]. After, affine or non-linear image registration should be achieved, by setting the default parameters, to convert the alignment of the structural MR images from spatial to common anatomical spaces and register them to a common template.
We note that these preprocessing steps are implemented in several publicly available neuroimaging software tools such as the Freesurfer tool that impliedly includes the preprocessing techniques with the brain segmentation process. After image preprocessing and registration, each MR image can be cropped into a patch with a bounding cube for the ROI to speed up the learning process. The size of the cube is optimally chosen to keep important context information for ROI segmentation.

[1]Note here that all the modalities are not present for all the subjects. For instance, tau PET is available only in ADNI3. Moreover, DTI and fMRI are not found in ADNI1 and only about half of the subjects in ADNI1 have FDG PET [53]

Table 1
Open Datasets comparaison

Dataset	*Description*	*Characteristics*	*Impact on model performances*
ADNI	ADNI-1 (2003): 200 NC, 400 MCI, and 200 Mild AD **Modalities**: MRI and FDG PET datasets ADNI-GO (2009): 200 early MCI (EMCI) new subjects and around 500 NC and MCI subjects rolled over from ADNI-1 **Modalities**:MRI scans, FDG and F18 Amyloid PET scans, clinical and cognitive information, , and APOE genotyping ADNI-2 (2011): multiple new subjects-150 NC, 150 EMCI, 150 late MCI (LMCI), and 200 mild AD Around 450–500 CN and MCI subjects rolled over from ADNI-1 and 200 EMCI added from ADNI GO **Modalities**: longitudinal clinical, MRI, PET (18F-Florbetapir and FDG), cognitive, blood and CSF biomarkers on new subjects with fMRI, DTI, ASL biomarkers ADNI-3 (2016): new came up 135–500 NC, 150–515 MCI and 85–185 AD subjects from longitudinal study in addition to 295–330 NC, 275–320 MCI and 130–150 AD new patients **Modalities**: longitudinal tau-PET scans with AV1451	Collect multicenter multimodal (sMRI, FDG PET, amyloid PET, tau PET, diffusion MRI, and fMRI) and longitudinal data on AD Provide raw data as well as pre-processed data Consider different classes (NC, AD, MCI, pMCI, sMCI, EMCI, and LMCI) allowing interesting researches to ensure the detection of AD from early stages	enhance model performances as it provides a large amount a large amount of data with different modalities and of good quality - ensure reliability and generalizability of the performance results of a given model since evaluated on multicenter data. allow researchers to develop important models that predict the MCI conversion to AD (sMCI vs pMCI classification) which is the most challenging issue helping clinicians to take effective interventions for the progression of the disease
OASIS	OASIS-1 (2007): cross-sectional data set included MRI data from 416 subjects aged 18 to 96 years (for 100-AD and 316-NC subjects) OASIS-2 (2010): longitudinal MRI data of 150 older adults (72 NC, 64 AD and 14 MCI patients) OASIS-3 (2018): longitudinal neuroimaging (MRI and PET), clinical, cognitive, and biomarker dataset for 609 NC and 489 MCI and AD subjects	Collect MRI and PET images of good quality Provide longitudinal data for patients with different ages Contain only three classes (NC, MCI and AD)	enhance model performances as it provides a large amount of data especially with the addition of the latest version OASIS-3 cannot be used to predict the MCI conversion to AD (sMCI vs pMCI classification)
AIBL	Launched in 2006 Longitudinal study over around 1100 participants having age more than 60 years divided into 211-AD, 768-NC and 133-MCI cohort **Modalities**:clinical and cognitive assessment, MRI, biological samples, PET, and lifestyle assessment	Collect MRI and PET images of good quality Consider only elderly patients Contain only three classes (NC, MCI and AD) Unbalanced data (with a large number of NC patients)	cannot be used to predict the MCI conversion to AD (sMCI vs pMCI classification) the unbalanced data can cause the degradation of the model performances if not well adressed.
MIRIAD	Established in 2013 708 longitudinal T1 MRI scans of 46 mild– moderate AD and 23 NC, with accompanying information on gender, age and MMSE scores	only T1 MRI images conducted by the same radiographer with the same scanner and sequences Only two classes	present some limitations for model performances if considered alone for the experiments: the efficiency (small dataset to train a deep model) and the generalizability (same scanner and a single center)

4.2. ROI segmentation

The segmentation represents a challenging task that may affect classification performances in disease diagnosis. For AD classification, this task was considered differently in many papers. Some papers applied firstly tissue segmentation to obtain GM, WM and CSF tissues and then structural segmentation of one or more tissue to obtain relevant features such as regional cortical thickness. Several works segmented the whole brain on anatomical regions (ROIs) and extracted global features such as volumes, surfaces and shapes

directly from ROIs [76]-[77]. While some recent works applied ROI subfields segmentation to extract relevant local features by focusing on the different subfields instead of considering the ROI as a full block.

4.3. Feature extraction

The features extraction step aims to extract discriminant features from brain images and then obtain relevant information for classification. This step was considered differently in the literature. In fact, in the past few decades, the extraction of the features has been independent of the classification model. However, recently, it has been included in the classification process to avoid sub-optimal learning performance for AD diagnosis due to the possible heterogeneous nature of features and classifiers. Many types of local and global features have been extracted from structural MRI. These features can be divided into six different categories: voxel-level, vertex-level, sliced-level, ROI-level, patch-level, and whole-image-level [31],[78], [79]. In [31], authors reviewed these different approaches and discussed the advantages and drawbacks of each group. Despite the fact that the patch-level representation was proved to be the most effective in characterizing the early AD- related structural changes of the brain, the selection of discriminative patches from tens of thousands of patches in each MR image represents a delicate and challenging task [31]. Consequently, in recent years, many efforts have been directed to perform an automatic discriminative landmarks detection from MR images and then extract patches from these landmarks to learn features [80]. While some papers have considered the preselection of the most informative patches based on anatomical landmarks isolated to feature extraction model and disease diagnosis, the most recent ones have performed these tasks in a unified deep-learning framework [80]. Otherwise, according to the manner of AD alterations detection, the features can be classified into inter-subject similarity and intra-subject variability [81]. The first class detects similarities between persons from different groups showing specific disease severities [22]-[82]-[87]. Although the efficiency of the first type of features, they have some limitations due to the fact that structural atrophies leading to cognitive decline are not homogeneous within a given subject [81]. To deal with this drawback, intra-subject variability can provide relevant information by focusing on the relationships between different inter-related regions for a subject [81]-[88]-[91]. The most used features in the literature include the shape, surface area and volume features from HC and entorhinal cortex [92]. Moreover, they integrate the normalized hippocampal volume (NHV), the surface roughness (SR), the local surface roughness (LSR), the asymmetries between right and left hippocampus [60], the mean curvature of the cortical surface [93]-[94] and recently the average of principal curvature ratio within hippocampal subfields [95] and the entorhinal cortex texture [79], [96].

4.4. Classification

This step allows the diagnosis of AD patients. Most papers considered binary classification: AD vs. NC, MCI vs. NC, early MCI (eMCI) vs. late MCI (lMCI) and pMCI vs. sMCI. Simply classifying AD patients from NC is a relatively easy task. However, it is not very significant. In fact, distinguishing AD from NC can be achieved easily by focusing on the behaviors of the patients and many indicators in clinical diagnosis. Therefore, the most challenging classification is the prediction of pMCI from sMCI and NC. Indeed, as pMCI would progress to AD, identifying pMCI allows to ensure early and accurate diagnosis of AD and then detect the disease from its early stages.

5. Main methods for the classification of AD and its prodromal stage

In this section, we will survey the most recent papers (from 2018 to 2022) that present important works for AD prediction from its prodromal stage. We particularly focus on the biomarkers that allows to detect whether an MCI subject will convert to AD or not and discuss the main DL architectures developed to achieve this challenging task.

As previously stated, the hippocampus is one of the important regions which is early affected by AD. The segmentation of this ROI represents a challenging task that has attracted much attention over the last few decades. Therefore, before surveying the MCI converter prediction models, we will expose, in the following, the most recent architectures proposed in the literature for the segmentation of the hippocampus.

5.1. Models for hippocampus segmentation

Many software tools have been developed over the years by different scientific groups applying various algorithms for automatic brain segmentation (multi-atlas segmentation, model-based segmentation). The most used tools for neuroimaging studies are FreeSurfer 5.0, 5.3.0, 6.0, 7.0 [68], First of FSL 5.0 [69], statistical parametric mapping (SPM) [97], and Automatic segmentation of hippocampus subfield (ASHS) (segmentation of the hippocampus subfield using the included atlas) [98]. These tools have been widely used to segment the human brain including the hippocampus. Moreover, the FreeSurfer, SPM, and FSL are not limited to segmentation alone. They ensure also image preprocessing. Despite their efficiency, these tools are time-consuming. Recently, FASTSURF, a semi-automatic contouring-based segmentation model was developed for the hippocampus segmentation [99]. It used a mesh processing technique. Compared to FreeSurfer, and FSL, this model was found superior.

More recently, authors proposed the FastSurfer, a fast and robust method for volumetric segmentation, morphometric estimation of brain structures including cortical thickness, and reconstruction of cortical geometry [100]. This method ensures the segmentation of the whole-brain (cortical and subcortical regions) into 95 classes including the hippocampus, in under only 1 min on the GPU. The FastSurfer outperforms FreeSurfer with respect to runtime, reliability and sensitivity In the last years, many DL-based models have been proposed for HC segmentation. Most of them were limited to full hippocampus segmentation and considered the hippocampus as a single structure due to image resolution restrictions. However, some methods were interested to the HC subfield segmentation.

5.1.1. DL-based methods for entire hippocampus segmentation

In [77], P. Yi, et al. conducted a comprehensive review of hippocampal segmentation in brain MRI images using traditional ML and DL methods. Moreover, they presented the well-known toolkits applied to brain hippocampal segmentation. Many recent hippocampal segmentation methods were reviewed. These methods are based on a Convolutional neural network (CNN) and U-net. Some of them segmented the hippocampus on a 2D slice sequence while others worked directly on a 3D volume. Despite the computational cost of the 3D method, it is more popular as it naturally captures all spatial contexts contrarily to the 2-D techniques that often ignores the context information on the vertical axis. The most recent CNN-based models are those proposed in [101]-[110]. They paid attention particularly to the speed of segmentation and achieved a Dice score up to 87%. Moreover, the U-net network was known to be one of the most successful segmentation methods. In this context, many U-net-based models have been built for the hippocampus segmentation and achieved very interesting results [111]-[115]. Moreover, D. Carmo et al. proposed in [116] an extended 2D Consensus Hippocampus Segmentation method (E2DHipseg). The method was based on a full convolutional network inspired by a 3-plane U-net. Evaluated on HarP dataset, this work achieved a DSC of 90% compared to the Hippodeep [103], the QuickNat [106] and the Freesurfer version 6.0 with DSCs equals to 85%, 80% and 70%, respectively.

Some methods, among those mentioned above, have released their code or software packages and provided open toolkits for hippocampus segmentation [77]. Some toolkits are Hippodeep [103], quickNat [106], deepNat [104], HippMapp3r [115] and E2DHipseg [116]. Recently, S. Nobakht, et al. developed in [117] an automatic segmentation tool (DeepHarp) to segment the hippocampus from T1-weighted MRI datasets according to the ADNI harmonized hippocampal protocol (HarP). It combines a traditional atlas-based segmentation method with an optimized CNN architecture. The new tool was compared to FSL and Freesurfer algorithms and the two DL approaches, QuickNat and HippoDeep. It was shown that DeepHarp achieved the highest Dice score in accuracy (0.893 for left HC and 0.889 for right HC) using 23 test datasets (46 hippocampi segmented from ADNI-Harp data) while HippoDeep fulfilled the highest Dice score in robustness (0.957 for left HC and 0.96 for right HC). More recently, H. A. Helaly et al. proposed, in [118], a DL-AD hippocampus segmentation framework (DL- AHS) based on U-Net for automatic AD left and right hippocampus segmentation. The framework was applied on the baseline coronal T1-weighted structural MRI data obtained from ADNI and neuroimaging tools and resources collaboratory (NITRIC) datasets. Two different architectures were developed. The first, SHPT-Net model, used simple hyper-parameters tuning in the U-Net while the second, denoted by RESU-Net, applied a transfer learning technique in which the ResNet blocks are integrated in the U-Net. It was proved that both proposed architectures are robust and accurate. They achieved high performance with DSC values equals to 93.5% and 94% for SHPT-Net and RESU-Net, respectively. They

also outperformed many state-of-the-art models [102],[116], [117]. An open-source code of this framework is available at *https://github.com/hadeerhelaly/hippocampus- segmentation*.
All the aforementioned methods only achieved the hippocampus segmentation, independently from the prediction task. However, there is some models that combined hippocampal segmentation with disease diagnosis to improve the prediction performances. In fact, integrating hippocampus segmentation into AD diagnosis is often beneficial to each other. In [119], the hippocampus segmentation was achieved with the clinical score regression by a multi-task deep learning (MDL) method. The proposed network applied an U-Net architecture with a Dice-like loss function for hippocampus segmentation. It was shown that the model reaches the best Dice similarity coefficient (DSC) of 0.893 and positive predicted value (PPV) of 0.890 in hippocampus segmentation compared with the Random Forest (RF) regression-based method and the Multi-atlas (MA) based method. This proves that the MDL method succeeded to locate the boundary of hippocampus accurately. After, in [59], the hippocampal segmentation and the disease status classification tasks were jointly performed by a multi-model DL framework based on CNN. The deep CNN model used ResNet blocks to generate a volumetric HC segmentation mask. The proposed framework was compared to other HC segmentation methods including FSL, the U-Net architecture proposed in [119] and the HippoDeep [103]. It was proved that it achieves the best dice similarity coefficient of 87.0% on the ADNI dataset versus 72.2%, 73.8% and 84.6% for the other methods, respectively. Moreover, it achieves the best positive predicted value (84.6% on the ADNI) and segmentation sensitivity (89.7% on the ADNI), which indicate that it can effectively detect hippocampal regions from the background. Moreover, in [120], J. Sun et al. proposed an improved end-to-end dual-functional 3D CNN to simultaneously perform segmentation of the bilateral hippocampi from 3D brain MRI scans and diagnosis of AD progression states. The proposed network was based on the V-Net and applied to 132 samples from ADNI. It achieved DSCs equals to 0.9035 ± 0.020 and 0.9162 ± 0.023 for left hippocampi and right hippocampi, respectively.
Table 2 provides a more detailed comparison of these architectures highlighting their strengths and their weaknesses.

5.1.2. DL-based methods for HC subfields segmentation

Motivated by the fact that AD affects differently the various HC subfields at different moments during the disease progression, automatic and accurate HC subfield segmentation becomes crucial to obtain early biomarkers of the disease and efficiently detect the conversion of MCI to AD. In addition to Freesurfer that ensures the HC subfield segmentation, some DL-based methods have been recently proposed for hippocampus subfield segmentation. In [121], Y. Shi, et al. proposed the UGNET method using an adversarial training approach. Moreover, in [122], H. Zhu, et al. integrated U-net and dilated dense network to develop a new fully convolutional network (FCN) for HC subfield segmentation of infants and adults. They proposed two versions, the Dilated Dense U-Net (DUnet) and the Residual DUnet (ResDUnet). Although their good performances, these methods are unable to well generalize on unseen data. After, in [123], Z. Yang, et al. developed a multi-scale CNN based Automated hippocampal subfield Segmentation Toolbox (CAST) to quickly and accuratelysegment the different hippocampal subfields in less than 1 min for a subject. CAST showed good segmentation accuracy in terms of dice coefficient (DSC) score. It significantly improved the reliability of segmenting small subfields, such as CA2, CA3, and the entorhinal cortex (ERC). Moreover, it is flexible and capable to segment raw images taken from different MRI scanner vendors with different acquisition parameters without revising any settings and network architecture. It is also applicable for multiple imaging modalities. Both the scripts and trained model are publicly available at *https://github.com/pipiyang/CAST*.
More recently, in [124], J. Manjón, et al. developed the DeepHIPS, a novel 3D UNET-based segmentation method using deep supervision and low-resolution feedback to facilitate the training process. Evaluated on two publically available datasets Kulaga-Yoskovitz and Winterburn, it was proved that the new method outperforms the classic UNET (notably for the Winterburn dataset). In fact, it achieved mean DICEs equals to 0.9116 and 0.9083 for Kulaga-Yoskovitz and Winterburn, respectively, compared to 0.9083 and 0.6148 for classic UNET. It was also shown that DeepHIPS succeeded to segment all hippocampus subfields better than HIPS [125] and ResDUnet [122] methods with an average of 0.9037 over Kulaga-Yoskovitz dataset compared to 0.8879 and 0.8960 for HIPS and ResDUnet, respectively.
Table 3 provides a more detailed comparison of the different discussed HC subfields segmantation models,

Table 2
DL-based methods for entire hippocampus segmentation

Reference	Method	Methodology	Strengths	Weaknesses
[103]	Hippdeep	CNN-Based model	- Short running time (<2min for the segmentation of raw brain T1 images) - Quite robust to changes in the subject and MRI contrast - Open toolkit - DSC equal to 85%	- Independent from the prediction task - Segment the hippocampus as a single structure
[106]	QuickNAT	A fully CNN: it follows a consensus of multiple 2D U-Net like architectures	- Short running time (complete hippocampal segmentation within 20s) - It provides a trained model that can segment multiple brain regions - Open toolkit - DSC equal to 80%	- Independent from the prediction task - Segment the hippocampus as a single structure
[116]	E2DHipseg, an extended 2D consensus hippocampal segmentation method	Full CNN inspired by a 3-plane U-net (merging sagittal, coronal and axial modified U-Nets into an activation consensus), including residual connection and VGG weight transfer	- Efficient DL segmentation model (DSC equal to 90%) thus using 3-plane U-Net can enhance the performances of the segmentation task - Open toolkit	- Independent from the prediction task - Segment the hippocampus as a single structure
[117]	DeepHarp	It combines a traditional atlas-based segmentation method with an optimized CNN architecture	- A simple method, as it only requires an atlas registration and subsequent application of the trained CNN - Robust and accurate method (DSC equal to 89.3% for left HC and 96% for right HC) - Validated using 114 multi-center datasets of patients with AD, MCI, HC and cerebrovascular disease	- Independent from the prediction task - Segment the hippocampus as a single structure
[118]	▪ SHPT-Net ▪ RESU-Net	▪ SHPT-Net: use of simple hyper-parameters tuning in the U-Net ▪ RESU-Net: integration of ResNet blocks in the U-Net	-Very robust and accurate methods (DSC=93.5% for SHPT-Net and 94% for RESU-Net) - Open source code	- Independent from the prediction task - Segment the hippocampus as a single structure
[119]	Multi-task DL method (MDL)	U-Net architecture with a Dice-like loss function	- It combines hippocampal segmentation with disease diagnosis to improve the prediction performances - Smooth and accurate method as it can remove some confounding voxels very well - DSC equal to 89.3%	- Limited to full hippocampus segmentation - Tested only on the ADNI-1 database
[59]	Multi-model	CNN model based on Resnet blocks	- It enhances AD prediction performances by integrating the hippocampus segmentation task with AD classification - Tested on the ADNI dataset and the HarP dataset - DSC equal to 87% on the ADNI dataset and DSC of 86.7% on the HarP dataset	- Limited to full hippocampus segmentation
[120]	Dual-functional neural network	3D-CNN based on V-Net	- It enhances AD prediction performances by using V-Net and jointly learning hippocampus segmentation and AD classification - Robust for both left and right HC segmentation with DSC equal to 90.5% for left HC and 91.6% for right HC	- The study included only 132 samples: 100 for the training phase, and the remaining 32 for the test -It does not segment the HC subfields

their strengths and their weaknesses.

5.2. Models for AD classification and MCI converter prediction

In this subsection, we will expose and analyze the most recent methods (from 2018) developed for MCI converter prediction. We will particularly focus on the biomarkers used to ensure the early prediction of AD. We note that the reviewed studies differed by the sample sizes of the dataset. Therefore, it is difficult to quantitatively compare them. However, a qualitative analysis of the results can be performed.

5.2.1. Models based on the hippocampus features

In this subsection, we will present the well-known models developed for AD prediction from early stages, based on features extracted from the hippocampus and its surrounding regions. In [126], a new approach was proposed to predict the MCI-to-AD conversion using MRI images from ADNI database. This approach reduced the amount of input images by slicing the data in different directions. Moreover, it exploited different useful information obtained from both CNN-based features and FreeSurfer- based features for sMCI vs pMCI classification. The CNN was firstly trained with AD/NC patches and then used to extract deep features from 2.5D patches related to hippocampus regions of MCI subjects. Furthermore, FreeSurfer extracted more morphological information from the whole brain such as surface area, cortical volume and cortical thickness average providing thus complementary information for good classification. It was shown that the 2.5D CNN model, when trained by the CN vs. AD training set, particularly achieved the best accuracy in the sMCI vs. pMCI task with an accuracy of 73.04%. It was also proved that the full model achieved an accuracy of 79.9% and an AUC of 86.1% in leave-one-out cross-validations. These results outperformed those obtained with only CNN-based features or FreeSurfer-based features.

In [127], S. Basaia, et al., developed a CNN-based DL algorithm for the prediction of the individual diagnosis of AD, sMCI and pMCI based on a single cross-sectional brain structural T1-weighted MRI images for 1409 subjects from ADNI dataset (294 AD, 253 pMCI, 510 sMCI and 352 NC) and 229 subjects from an independent Milan dataset (124 with probable AD, 27 pMCI, 23 sMCI, and 55 NC). The major contribution of this paper with regard to many previous works is its ability to compare data from different centers, neuroimaging protocols and scanners then resulting in more reliable and reproducible results. It was proved that the proposed method achieved very high rates in the AD vs HC classification with 99% using the ADNI dataset only and 98% using both ADNI and Milan datasets. Moreover, for the most challenging classification sMCI vs pMCI, it reached an accuracy of 75% for the two datasets. In [128], C. Platero et al. exploited the local surface roughness (LSR) as a new marker for discriminating AD progression by assessing hippocampal atrophy. This new marker was derived from the mean curvatures in the most statistically significant local regions of the hippocampal surfaces. It was used to differentiate among four clinical groups: NC, subjective cognitive decline (SCD), MCI (sMCI and pMCI), and AD by detecting significant differences in hippocampal atrophy for them. The performances of the new proposed marker were tested using T1-weighted MRI images of 307 subjects (70 NC, 87 SCD, 137 MCI and 13 AD). Moreover, 97 patients with MCI were classified into 61 sMCI and 36 pMCI after a follow-up of 3-year. The LSR was compared to other traditional markers that are left and right hippocampal volumes (LV and RV), normalized hippocampal volumes (NHV) (volumes averaged and normalized with ICV) and hippocampal SR. It was proved that this marker succeeded in better discriminating between clinical groups at different stages especially between the NC and SCD with an accuracy equal to 68.2% compared to that obtained with NHV which is 57.2%. Moreover, the classification accuracy of the LSR for sMCI and pMCI was 74.3% while it achieved only 68.3% with NHV.

Furthermore, it was shown, in the maps of the mean curvature of each group, that the atrophy patterns extended particularly to the CA1 and subiculum subfields of the hippocampus. Furthermore, to predict AD and understand the evolution of the disease, Y. Liu et al., developed in [129] a novel general system AD using longitudinal data from ADNI dataset. The data corresponds to 3D whole brain DTI images for 151 subjects (51 AD and 100 non-AD) measured at four-time points of baseline, 6 months, 12 months, and 24 months.

On the one hand, the classification task in the new system was ensured by the CNN model Visual Geometry Group (VGG) that contained, in addition to max-pooling layers, a global average pooling (GAP) layer to improve the AD prediction accuracy. It was shown that using the training dataset at baseline, the VGG model achieved accuracies of 0.8675, 0.8452, 0.8335, and 0.7463 in predicting AD in the test datasets at baseline, 6 months, 12 months, and 24 months, respectively. Moreover, it achieved AUCs of 0.8571, 0.8291, 0.8583, and 0.7756 for prediction of AD in the test datasets at baseline, 6 months, 12 months, and 24 months, respectively, using the training data at the same time points.

On the other hand, another important module was used for the discovery of potential causal relationships between the brain neuroimaging regions and AD and then the identification of brain regions underlying AD. This module was based on the conditional generative adversarial network (CGAN) and two-sample tests. It was proved that the ventricular system (the ventricles and enlarged ventricle) represents a useful biomarker for the diagnosis of AD whereas the temporal lobes (including the hippocampus) are crucial in AD

development at the early stages. To improve the AD and MCI diagnosis, R. Cui and M. Liu proposed, in [130], a new hippocampus analysis framework that integrates multi-level and multi-type features. The framework combined local features extracted by a 3D Densely Connected Convolutional Networks (3D DenseNets) and traditional global shape features obtained by applying a SPHARM-PDM (Spherical Harmonic shape description) [131] on the binary hippocampal masks. The 3D DenseNet was built on 3D local patches covering the whole hippocampus and its surrounding regions such as amygdala and parahippocampal gyrus to extract much more useful information. It was proved that the combination of DenseNet and shape analysis performs better than the use of each method alone for all binary classifications: AD vs. NC, pMCI vs. sMCI and MCI vs. NC. Moreover, by comparing DenseNet to shape analysis, it was shown that local visual features provide much more information about AD diagnosis than global features. Furthermore, the new method was compared to previous methods using conventional features such as hippocampal volume, GM volumes of 93 ROIs of the whole brain and voxel-wise features. It was proved that it achieves the best performances for all classifications and particularly for AD vs. NC with an area under the ROC curve of 96.95% and an accuracy of 92.29%, which is higher than the ROI-based-method (86.51%) and voxel- based-method (86.02%). For the most challenging scenario pMCI vs. sMCI, it achieved an AUC ROC of 79.7% and an accuracy of 75.00% that improved those of ROI-based method (AUC of 71.89% and accuracy of 73.23%) and voxel-based method (AUC of 73.83% and accuracy of 72.98%). After, K. Popuri, et al. proposed, in [132], a new framework that not only predicts the current clinical diagnosis of a subject from its associated MRI image at any stage of the disease progression spectrum but also determine if he will convert to AD (dementia of Alzheimer's type (DAT)) (DAT+ trajectory) or it will stay non AD in the future (DAT− trajectory) over different ranges of time to conversion (TTC) before the disease onset. The new developed system applied both tissue and structural segmentation to extract different anatomical ROIs from GM and CSF tissue regions. After, the structural volume-based features in most discriminative ROIs was used to compute a new AD biomarker that is a MRI DAT score defined as a similarity metric between the subject's MRI structural patterns and the DAT patterns by measuring the neurodegeneration in the brain though an ensemble-learning based classifier. Therefore, based on this score, it assesses the presence of AD structural atrophy patterns in NC and MCI stages. The classifier is trained on the two extremes groups of patients: NC and AD and evaluated independently on many images from different databases: ADNI, AIBL, OASIS, and MIRIAD. It was shown that entorhinal, inferior temporal, middle temporal, and fusiform areas represent the top cortical ROIs (left and right) exploited by the classifier to compute the score. Moreover, the new system achieved good performances particularly for short TTC with an AUC of 0.81 for TTC of 6 months and 0.73 for TTC of up to 7 years for sMCI versus pMCI classification. However, it failed to expect the NC to be on the DAT+ trajectory, especially for long TTC. Therefore, additional features other than volumes should be considered to improve the performance of the model particularly during the early stage. In the same context and motivated by the fact that entorhinal cortex is the first region affected by AD, L. S., et al. investigated, in [96], entorhinal cortex texture to identify MCI patients from an early stage and differentiate sMCI from pMCI. They used 608 structural MR images from the ADNI dataset (194 NC, 200 MCI, 84 MCIc or pMCI, and 130 AD subjects). Authors showed that entorhinal cortex texture changes can predict the conversion from MCI to AD better than the hippocampal atrophy. Moreover, combining texture and volume features improved the classification performances to reach AUCs of 0.914, 0.740, 0.756 and 0.780 for NC vs. AD, NC vs MCI, sMCI vs MCIc (pMCI) and MCI vs AD groups. They also showed that the entorhinal cortex degeneration was more pronounced in the pMCI subjects than in the MCI persons whereas hippocampal volume reduction in both pMCI and MCI stable groups was similar. This proves that particularly through entorhinal cortex analysis, we could identify that MCI subjects in the future could develop the disease. Recently, A. Li, et al. developed in [60], a multi-channel cascaded CNNs-based method that exploited, in addition to hippocampal shapes, asymmetries in the atrophy between left and right hippocampus to enhance the classification performances. They used 807 structural MR images from the ADNI dataset (194 AD, 164 pMCI, 233 sMCI, and 216 NC). The proposed system was composed of a multi-channel 3D CNNs to extract features of left and right hippocampal shapes from binary hippocampus segmentation masks as well as asymmetrical features by analyzing internal differences between the two hippocampus masks. Moreover, a 2D CNN was cascaded at the end of the deep 3D CNNs to capture high-level correlation features by combining the features from the two hippocampi. These high-level features are finally used with the asymmetry features for early diagnosis and classification of AD patients. It was

proved that the proposed method achieved accuracies of 85.9%, 73.3% and 71.0% and AUCs of 88.4%, 74.6% and 71.9%. For AD vs. NC, MCI vs. NC and pMCI vs. sMCI classifications, respectively. Moreover, it performed better than the individual 3D CNNs for learning the combined shape and asymmetry features. Furthermore, it was shown that the new proposed method outperformed the traditional SPHARM-PDM [131] illustrating then its capability to well modeling the 3D surface of the hippocampal structure for disease diagnosis. However, it achieved lower performances than the method in [130] that used a DenseNet built on 3D local patches covering the whole hippocampus and its surrounding regions providing then much more visual and appearance information. In [95], unlike most studies which extracted features from the entire hippocampus, T. H. K, et al. focused on the hippocampus subfields to identify relevant features allowing to predict whether an MCI patient will convert to AD using structural MRI images from the ADNI database. The features models variations, including atrophy and structural changes to the cortical surface. In fact, the authors exploited the relationship between the alteration of the hippocampal subfields and the progression of cognitive impairment and extracted strong neuropathological features that are the rate of change in volume and the surface area from the hippocampal sub-regions. Moreover, they noticed that the boundaries of the hippocampal subfields grow increasingly spurious during conversion to AD. Consequently, they proposed the average of ratio of principal curvatures (RPC) as a new important biomarker in addition to the principal curvatures (maximum and minimum curvatures) to detect whether an MCI subject will develop AD according to the change in surface pattern. It was proved that the presubiculum, the subiculum, CA1, CA4, and the molecular layer are the important ROIs from which the most relevant features used in the MCI prediction are extracted. Moreover, based on these features, the proposed work achieved an accuracy of 79.95%.
In table 4, we summarize the different models which exploit HC features to early predict AD.

5.2.2. Models based on whole brain features

In [133], M. Liu et al. developed a CNN-based multi-instance learning model (Deep MIL) for AD classification and MCI conversion prediction. The proposed model first pre-selected the top 40 anatomical landmarks via a data-driven landmark discovery algorithm to generate image patches. Then, it learned local patch-level feature representations of image patches and global representations of multiple landmarks to produce subject-level feature representations for AD diagnosis. The new proposed method was compared to state-of-the-art approaches, the ROI-based method, the voxel based morphometry, the conventional landmark-based morphometry (CLM) and a landmark-based deep single- instance learning (LDSIL) method that is a variant of LDMIL. Trained on the ADNI-1 and evaluated on the ADNI-2 and the MIRIAD datasets, it was proved that the proposed LDMIL method generally outperforms the other methods in both AD vs. NC classification and MCI conversion prediction (pMCI vs. sMCI). On ADNI-2, it achieved an AUC value of 0.959, which is much better than those yielded by ROI, VBM, and CLM methods. Moreover, it obtained the best performances for MCI conversion prediction with an AUC of 0.7764 and an accuracy of 0.769. Furthermore, authors proposed to adopt the knowledge learned from AD and NC subjects to guide the prediction of MCI conversion by training the classifier using pMCI and AD as positive samples, and sMCI and NC subjects as negative ones in ADNI-1. The performances with transferred knowledge were improved to achieve an AUC of 0.7904 and an accuracy of 0.7834. Moreover, C. Lian et al. proposed, in [86], a hierarchical fully convolutional network (H-FCN) to locate both discriminative local patches and regions in the whole brain sMRI and then automatically learn multi-scale (i.e., patch, region-, and subject level) features for AD classification and MCI conversion prediction of subjects from the two independent datasets (ADNI-1 and ADNI-2). The proposed H-FCN network included successively location proposals followed by patch-level sub-networks, then region- level subnetworks, and finally a subject-level sub-network. Therefore, it was fed not only by mono-scale feature representations, but fused complementary multi-scale feature representations to achieve AD classification. To boost the diagnostic performance, authors applied the network pruning strategy to remove uninformative patches and regions and then refine the initial H-FCN. Moreover, they exploited the transfer learning technique to transfer the network parameters learned from AD and NC subjects and then enhancing the pMCI vs. sMCI classification performances. Furthermore, they used data augmentation to reduce the over-fitting risk. The main advantage of this work is that the localization of discriminative brain regions in sMR images is unified with feature extraction and classifier construction contrarily to many other works where these tasks are independent. The unification of these

three important steps allows them to be more seamlessly coordinated in a task-oriented manner. This improves diagnostic performance. Moreover, contrarily to the two well-known patch-based methods LBM and DMIL that pre-selected the top 40 landmarks as inputs, the H-FCN method considered all anatomical landmarks equally as potential location proposals to extract much more information for classification. In fact, it was proved that the proposed H-FCN method outperforms ROI method [134], VBM method [135], and the two well-known patch-based methods [133], [136], especially on the more challenging classification, i.e., pMCI vs. sMCI. It achieves, indeed, the best performances with accuracy and AUC equals to 0.809 and 0.781, respectively, for an image patch size $25 \times 25 \times 25$. After, M. Liu et al., applied in [137] a data-driven technique to locate discriminative anatomical landmarks from MR images. Then, they proposed a landmark-based deep feature learning (LDFL) framework using CNN models to jointly achieve features extraction from informative patches and AD/MCI classification training. It was proved that the proposed framework boosts the AD/MCI diagnosis performance, compared with several state-of-the-art methods. Recently, Y.-X. Zhao et al. developed in [138] a region ensemble model based on a divide and conquer strategy to better perform the MCI conversion prediction task and overcome the problem of dimensionality. In fact, the model first extracted a global feature map from the whole brain. After, each regional feature map, extracted using a segmentation model, was fed to a region-based diagnosis subnetwork to extract discriminative features. Finally, an ensemble subnetwork was constructed for the weighted fusion of the regional features with an attention mechanism to obtain the final description for the subject. Moreover, to regularize the training process, the authors proposed a new relation regularized loss. Experiments on ADNI 1 and 2 for MCI conversion prediction proved that the proposed method has achieved better performances when compared to state-of-the- art methods with an AUC ROC improved from 79.3% to 85.4%. Moreover, K. Hett et al. proposed in [81] a new graph- based grading framework that exploited both inter-subject similarity features and intra-subject variability features to provide complementary information for accurately predicting conversion of MCI subjects to AD. The inter-subject similarities were estimated with a patch-based grading (PBG) approach while intra-subjects' variability were modeled by a graph-based approach (computed with the difference of the grading value distributions for each structure). The new proposed framework was applied to two different anatomical scales, the whole brain structures and the hippocampal subfields, to finally develop a multiscale approach. This multi- scale graph-based grading (MGG) method efficiently and simultaneously combined relevant information for the MCI conversion prediction. Evaluating the new proposed method using the ADNI-1 dataset, it was proved that it achieved an 80.6% of AUC and 76% of accuracy for the MCI conversion prediction (within three years). Moreover, when combined with cognitive scores, the proposed method obtained 85.5% of AUC and 80.6% of accuracy. Otherwise, experiments have shown that the most frequently selected brain structures are the temporal lobe, the postcentral gyrus, the anterior cingulate gyrus, the hippocampus and the precuneus. Furthermore, the most frequently selected hippocampal subfields are the CA1- SP, the CA1-SRLM, and the subiculum. More recently, C. Lian et al. proposed in [139] an attention-guided DL framework to automatically extract multilevel (local and global, inter-subject consistent and individual specific) discriminative features from the whole-brain sMRI scans for AD diagnosis and MCI conversion prediction. The proposed framework took into account the fact that the locations of the AD-associated brain atrophy as well as the different degrees of severity in certain brain regions differ from one individual to another. Moreover, it exploited, in addition to information conveyed in image patches, global structural information of the brain to boost the diagnostic performance. In fact, it was composed of two stages. In the first stage, a backbone fully convolutional network (FCN) was built to generate an AD attention map (DAM) for the localization of discriminative regions in each individual brain sMRI. In the second stage, a multi-branch hybrid network (HybNet) was developed to learn high multilevel features from DAM-weighted backbone feature maps and image patches extracted from the input sMRI for brain disease diagnosis. Trained on ADNI-1 and evaluated on ADNI-2, the proposed method showed superior performances for the MCI conversion prediction compared to the well-known DL models, the H-FCN [86], and the DMIL [133]. It achieved indeed an accuracy of 0.827 versus 0.809 and 0.769 for the two other models, respectively, and an AUC of 0.793 versus 0.781 and 0.776. Moreover, for the AD classification, the ACC and AUC values were improved from 0.903 (for the H-FCN) to 0.919 and 0.951 (for H-FCN) to 0.965, respectively. Furthermore, M. Ashtari-Majlan et al. developed, in [140], a new approach to identify the most statistically significant anatomical landmarks and then extract the corresponding patches. The novel approach was based on the multivariate T2 Hotelling statistical test. The patches were fed to a multi- stream deep CNN, one

patch to each stream, to ensure the final classification.
This paper presents many strengths. In fact, it eliminated morphological structure deformations by using rigid registration. This avoided possible inherent errors in the classification phase and reduced computational complexity. Moreover, by employing T2 Hotelling test, it reduced the impact of potential errors caused by inter-subject brain shape differences. Finally, it used transfer learning in the training phase for the pMCI versus sMCI classification to overcome the complexity of learning due to the fact that structural changes in MCI brains caused by dementia may be very subtle compared to CN and AD. Trained and evaluated on ADNI-1, the multi-stream CNN method achieved high performances for AD versus CN as well as pMCI versus sMCI classifications. It reached indeed an accuracy of 97.78% versus 91.90% and 90.30% for the well-known HybNet [139] and H-FCN [86] models, respectively, when classifying AD and CN individuals. Moreover, it achieved an AUC of 99.97% versus 96.5% and 95.1% for the two other methods, respectively. Furthermore, it performed very well in distinguishing pMCI and sMCI with an AUC and a specificity of 94.39% and 99.70%, respectively. However, the best accuracy was still achieved by the region ensemble model [138] with a value of 85.90%.
Table 5 focuses more on the models which identify essential landmarks from MR images to achieve the AD prediction. In this table, we define the different methodologies developed in the literature by enumerating their main steps, their strengths and their limits. Moreover, we present their landmarks identification process.

6. Discussion and future directions

Numerous studies have been conducted to early predict AD from MCI stage. The studies differed by several factors including the sample sizes, the different extracted features (global, local, patch-level, region-level, and subject-level), and the classification approaches. Some conclusions can be drawn by analyzing the reviewed studies. The prediction of MCI converters to AD is still a challenging task. This task was considered differently in all the works. Some of them considered the hippocampus as the most crucial ROI that is early affected by the disease and then provided discriminative information for AD prediction from MCI stage. Therefore, efforts were directed to accurately segment this region. In this context, the 3-D U-NET network was shown to be very performant for full HC as well as HC subfields segmentations. It succeeded to capture all spatial aspects contrary to the 2-D techniques that often do not take into account the context information on the vertical axis. Moreover, integrating hippocampus segmentation with AD diagnosis in the same process often improves prediction performances. Furthermore, by analyzing the relationship between the alteration of the hippocampal subfields and the progression of cognitive impairment to AD, the presubiculum, subiculum, CA1-SP, CA1-SRLM, CA4, and molecular layer subfields were proved to be the most significant parts involved in the MCI converters prediction. Although the importance of the hippocampus, features extracted from this ROI can be reinforced by features from surrounding regions, such as parahippocampus and amygdala, providing much more information to improve the prediction of AD from early stages. Otherwise, some other cortical ROIs can be considered in assessing the presence of AD structural atrophy patterns in NC and MCI stages and then helping to perfectly predict AD from these stages. These ROIs are the entorhinal, inferior temporal, middle temporal, and fusiform areas. Particularly, the entorhinal cortex texture changes were shown to be better than the hippocampal atrophy for the prediction of the conversion from MCI to AD. This can be explained by the fact that the memory circuit starts from the entorhinal cortex to the subiculum and then to CA1 and CA3. However, working directly on these predefined regions for all images without taking into account the variability of the locations and the severity of brain atrophy from one individual to another impacts the early prediction of AD. Therefore, many efforts were directed towards identifying the most significant landmarks from the input MRI images, to be considered in the classification process, ensuring accurate prediction of MCI conversion. Some research has shown that there are differences between females and males in the structure of specific brain regions. They affirmed that a woman's hippocampus, associated with learning and memorization, is larger than a man's and works differently. Conversely, a man's amygdala, related to the experience of emotions and the recollection of such experiences, is larger than that of a woman. Furthermore, in [141], Wierenga LM et al., showed that males have greater variability than females in all brain measures, including subcortical volumes and regional cortical surface area and thickness. In addition, L. V. Eijk, et al., examined, in [142], sex differences in the

hippocampal subfield volumes adjusted for brain size. They showed that males have a larger parasubiculum, fimbria, hippocampal fissure, and presubiculum while females have larger volumes for the hippocampal tail and no sex differences were found in the CA2/3, CA4, HATA, or GCDG subfields. Therefore, considering these sex differences in the prediction process can improve its performance.
Moreover, it was shown in [143] that patients with cognitive impairment showed abnormal emotion patterns. In [144], a novel DL-based system was proposed to detect cognitive impairment, by analyzing facial expressions and emotions while participants watched designed video stimuli. This system achieved an accuracy of 73.3%. Consequently, new DL models that combine MCI facial emotion with brain MR images can be tested to show the impact of this combination on the prediction performances.
Finally, it was observed that there exists a compromised link between semantic organization and brain oscillations in several neurological conditions such as AD or semantic dementia [145]. This can open new perspectives in quantitatively analyzing neurological disorders connected with distorted semantics and exploit this aspect in recent neuro-computational models [145].

7. Conclusion

In this paper, we provide a detailed review of the most recent works that have exploited structural MR images for the early detection and prediction of AD in patients from their stage of MCI. We discussed the main DL architectures developed to achieve this challenging task and highlighted decisive brain structures for prediction. An important class of works were based on the hippocampus and its surrounding zones to achieve an early AD diagnosis. Therefore, we presented, in this paper, the most recent architectures proposed in the literature for the segmentation of the full hippocampus as well as its subfields. We showed that the U-Net architecture is very efficient for the hippocampus segmentation. In particular, the RESU-Net was shown to be most successful with a DSC value equal to 94%.
Despite the efficiency of the models based on the hippocampus, another class of works that identified the most significant and informative landmarks from MRI images and extracted the corresponding patches as inputs to the DL system, have shown very interesting results. The best results for the MCI conversion prediction were then achieved by the multi-stream CNN with an AUC and a specificity of 94.39% and 99.70%, respectively, in addition to a region ensemble model with the best accuracy of 85.90%.

Acknowledgments

We deeply acknowledge Organism for Supporting this study through IA4MANECEVA Project number PAQ-Collabora - C3-1, Tunisia. The research leading to these results has received funding from the Ministry of Higher Education and Scientific Research of Tunisia under grant agreement number LR11ES48.
Prof Amir Hussain would like to acknowledge the support of the UK Engineering and Physical Sciences Research Council (EPSRC) - Grants Ref. EP/M026981/1, EP/T021063/1, EP/T024917/1.

Conflict of interest

The authors declare that they have no conflict of interest.

References

[1] Md. R. Ahmed, et al., "Neuroimaging and Machine Learning for Dementia Diagnosis: Recent Advancements and Future Prospects," IEEE Reviews Biomed Eng., vol. 12, pp. 19 - 33, Dec. 2019.
[2] R. Ganotra, S. Dora, and S. Gupta, "Identifying brain regions contributing to Alzheimer's disease using self regulating particle swarm optimization," Int J Imaging Syst Technol., no. 1, pp. 106 - 117, June. 2020.
[3] Y. Huang, et al., "Diagnosis of Alzheimer's Disease via Multi-Modality 3D Convolutional Neural Network," Front. Neurosci., vol. 13, no. 509, pp. 1 - 12, May. 2019.
[4] J. S. Lee, et al., "Machine Learning-based Individual Assessment of Cortical Atrophy Pattern in Alzheimer's Disease Spectrum: Development of the Classifier and Longitudinal Evaluation," Scientific Reports, pp. 1 - 10, Mar. 2018.
[5] H. Jia, et al. "Assessing the Potential of Data Augmentation in EEG Functional Connectivity for Early Detection of Alzheimer's Disease," Cognitive Computation, Sep. 2023.
[6] GB. Frisoni, et al. "The clinical use of structural MRI in Alzheimer disease," Nat. Rev. Neurol., vol. 6, no. 2, pp. 67 - 77, Feb. 2010.

[7] MD. Greicius, et al., "Default-mode network activity distinguishes Alzheimer's disease from healthy aging: Evidence from functional MRI," Proc. Natl. Acad. Sci., vol. 101, no. 13, pp. 4637 - 4642, Mar. 2004.
[8] J. Fripp, et al., "Appearance modeling of 11C PiB PET images: Characterizing amyloid deposition in Alzheimer's disease, mild cognitive impairment and healthy aging," NeuroImage., vol. 43, no. 3, pp. 430 - 439, Nov. 2008.
[9] C. Cabral, and M. Silveira, "Classification of Alzheimer's disease from FDG-PET images using favourite class ensembles," In Proc. IEEE 35th Int. Conf. Eng in Med and Bio Soc (EMBC), 2013, pp. 2477 - 2480.
[10] M. Liu, D. Zhang, and D. Shen, "Relationship induced multi-template learning for diagnosis of Alzheimer's disease and mild cognitive impairment," IEEE Trans. Med. Imag., vol. 35, no. 6, pp. 1463 - 1474, Jun. 2016.
[11] L. Billeci, et al., "Machine Learning for the Classification of Alzheimer's Disease and Its Prodromal Stage Using Brain Diffusion Tensor Imaging Data: A Systematic Review," Processes., vol. 8, no. 9, pp. 1 - 28, Sep. 2020.
[12] G. D. Rabinovici, et al., "A beta Amyloid and Glucose Metabolism in Three Variants of Primary Progressive Aphasia," Annals of Neurology., vol. 64, no. 4, pp. 388 - 401, Oct. 2008.
[13] Y. Fan, et al., "Structural and functional biomarkers of prodromal Alzheimer's disease: a high-dimensional pattern classification study," Neuroimage., vol. 41, no. 2, pp. 277 - 285, Jun. 2008.
[14] D. Chitradevi, S. Prabha, and A. D. Prabhu, "Diagnosis of Alzheimer disease in MR brain images using optimization techniques," Neural Computing and Applications., vol. 33, pp. 223 - 237, May. 2020.
[15] U. Morar, et al. "Prediction of Cognitive Test Scores from Variable Length Multimodal Data in Alzheimer's Disease," Cognitive computation, Jul. 2023.
[16] Y. Kumar, et al., "Artificial intelligence in disease diagnosis: a systematic literature review, synthesizing framework and future research agenda," Journal of Ambient Intelligence and Humanized Comput., vol. 14, pp. 8459–8486, 2023.
[17] P.Wu, et al., "AGGN: Attention-based glioma grading network with multi-scale feature extraction and multi-modal information fusion," Computers in Biology and Medicine, vol. 152, Jan. 2023.
[18] H. Li, et al., "Cov-Net: A computer-aided diagnosis method for recognizing COVID-19 from chest X-ray images via machine vision" , Expert Sys. With Applic., vol. 207, Nov. 2022.
[19] S. I. Dimitriadis, D. Liparas, and M. N. Tsolaki, "Random Forest feature selection, fusion and ensemble strategy: Combining multiple morphological MRI measures to discriminate among healhy elderly, MCI, cMCI and alzheimer's disease patients: From the alzheimer's disease neuroimaging initiative (ADNI) database," J. Neurosci. Methods, vol. 302, pp. 14 - 23, May 2018.
[20] J. Ramirez, et al., "Ensemble of random forests One vs. Rest classifiers for MCI and AD prediction using ANOVA cortical and subcortical feature selection and partial least squares", J Neurosci Methods, vol. 302, pp. 47 - 57, May 2018.
[21] I. Beheshti, H. Demirel, and H. Matsuda, "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., vol. 83, pp. 109 - 119, Apr. 2017.
[22] E. Moradi, et al., "Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects", Neuroimage, vol. 104, pp. 398 - 412, Jan. 2015.
[23] M. A. Myszczynska, et al., "Applications of machine learning to diagnosis and treatment of neurodegenerative diseases", Springer Nature Reviews, Neurology, 2020.
[24] N. Zeng, et al., "A new switching-delayed-PSO-based optimized SVM algorithm for diagnosis of Alzheimer's disease" , Neurocomput., vol. 320, pp. 195-202, Dec. 2018.
[25] N. Zeng, H. Li, and Y. Peng, "A new deep belief network-based multi-task learning for diagnosis of Alzheimer's disease", Neural Comput. & Applic., vol. 35, no. 16, pp. 11599-11610, June 2011.
[26] N. Yamanakkanavar, J. Y. Choi, and B. Lee, "MRI Segmentation and Classification of Human Brain Using Deep Learning for Diagnosis of Alzheimer's Disease: A Survey", Sensors, vol. 20, no. 3243, pp. 1 - 28, 2020.
[27] S. AL-Shoukry, T. H. Rassem, and N. M. Makbol, "Alzheimer's Diseases Detection by Using Deep Learning Algorithms: A Mini-Review," IEEE Access, vol. 8, pp. 77131 – 77141, Apr 2020.
[28] H. Sridhar, G. Chamarajan, and Y. Charishma, "Alzheimer's Diseases: A Survey," Int. J. Artificial Intelligence, vol. 8, no. 1, pp. 33 - 39, Jun. 2021.
[29] D. Agarwal, et al., "Transfer Learning for Alzheimer's Disease through Neuroimaging Biomarkers: A Systematic Review," Sensors, vol. 21, no. 21, Oct. 2021.
[30] S. Grueso, and R. Viejo-Sobera, "Machine learning methods for predicting progression from mild cognitive impairment to Alzheimer's disease dementia: a systematic review," Alzheimer's Research & Therapy, vol. 13, no. 162, pp. 1 - 29, Sep. 2021.
[31] S. Gao, and D. Lima, "A review of the application of deep learning in the detection of Alzheimer's disease," In. J. Cognitive Computing in Engineering, vol.3, pp. 1 - 8, Dec. 2021.
[32] N. M. Correa, et al., "Fusion of fMRI, sMRI, and EEG data using canonical correlation analysis," in Proc. IEEE Int Conf on Acoustics, Speech and Signal Processing, (Taipei: IEEE), pp. 385 - 388, 2009.
[33] P. Scheltens, et al., "Structural magnetic resonance imaging in the practical assessment of dementia: beyond exclusion," Lancet Neurol, vol. 1, no. 1, pp. 13 - 21, May. 2022.
[34] H. Braak, and E. Braak, "Neuropathological stageing of Alzheimer- related changes," Acta Neuropathol., vol. 82, no. 4, pp. 239 - 259, Sep. 1991.
[35] I. Brusini, et al., "Shape Information Improves the Cross-Cohort Performance of Deep Learning-Based Segmentation of the Hippocampus," Front. Neurosci., vol. 14, no. 15, pp. 1 - 19, Jan. 2020.
[36] P. Vemuri, D. T. Jones, and C. R. Jack Jr, "Resting state functional MRI in Alzheimer's disease," Alzheimer's research & therapy, vol. 4, no. 1, pp. 1 - 9, Jan. 2012.
[37] P.S. Weston, et al., "Diffusion imaging changes in grey matter in Alzheimer's disease: A potential marker of early neurodegeneration," Alzheimer's Res. Ther., vol. 7, no. 1, Jul. 2015.
[38] L. Billeci, et al., "White matter connectivity in children with autism spectrum disorders: A tract-based spatial statistics study," BMC Neurol., vol. 12, Nov. 2012.

[39] M. Szmuda, et al., "Diffusion tensor tractography imaging in pediatric epilepsy—A systematic review," Neurologia i Neurochirurgia Polska., vol. 50, no. 1, pp. 1 - 6, Jan. 2016.
[40] A. Arab, et al., "Principles of diffusion kurtosis imaging and its role in early diagnosis of neurodegenerative disorders," Brain Res. Bull., vol. 139, pp. 91 - 98, May. 2012.
[41] M.J. Müller, et al., "Functional implications of hippocampal volume and diffusivity in mild cognitive impairment," NeuroImage., vol. 28, no. 4, pp. 1033 - 1042, Dec. 2005.
[42] M.J. Müller, et al., "Diagnostic utility of hippocampal size and mean diffusivity in amnestic MCI," Neurobiol. Aging., vol. 28, no. 3, pp. 398 - 403, Mar. 2007.
[43] A. Fellgiebel, et al., "Color-coded diffusion-tensor-imaging of posterior cingulate fiber tracts in mild cognitive impairment," Neurobiol. Aging., vol. 26, no. 8, pp. 1193 - 1198, Aug. 2005.
[44] I.H. Choo, et al., "Posterior cingulate cortex atrophy and regional cingulum disruption in mild cognitive impairment and Alzheimer's disease," Neurobiol. Aging., vol. 31, no. 5, pp. 772 - 779, May. 2010.
[45] A. Demirhan, et al., "Feature selection improves the accuracy of classifying Alzheimer disease using diffusion tensor images. In Proc. IEEE 12th Int. Sym on Biomedical Imaging (ISBI), 2015, pp. 126 - 130.
[46] G. Prasad, et al., "Brain connectivity and novel network measures for Alzheimer's disease classification," Neurobiol. Aging., vol. 36, Jan. 2015.
[47] A. Ebadi, et al., "Ensemble Classification of Alzheimer's Disease and Mild Cognitive Impairment Based on Complex Graph Measures from Diffusion Tensor Images," Front. Mol. Neurosci, vol. 11, Feb. 2017.
[48] T. Maggipinto, et al., "DTI measurements for Alzheimer's classification," Phys. Med. Boil., vol. 62, no. 6, pp. 2361 - 2375, Feb. 2017.
[49] G.W. Eldeeb, N. Zayed, and I.A. Yassine, "Alzheimer'S Disease Classification Using Bag-Of-Words Based on Visual Pattern of Diffusion Anisotropy for DTI Imaging," In Proc. IEEE 40th Annu Int Conf Eng Med Biol Soc (EMBC), 2018, pp. 57 - 60.
[50] C. Ye, et al., "Connectome-wide network analysis of white matter connectivity in Alzheimer's disease," NeuroImage Clin, vol. 22, Feb. 2019.
[51] J.L.D. Da Rocha, "Fractional Anisotropy changes in Parahippocampal Cingulum due to Alzheimer's Disease," Scientific Reports, pp. 1 - 8, Feb. 2020.
[52] Dou, X, et al., "Characterizing white matter connectivity in Alzheimer's disease and mild cognitive impairment: An automated fiber quantification analysis with two independent datasets," Cortex, vol. 129, pp. 390 - 405, Aug. 2020.
[53] J. A. Samper Gonzalez, Learning from multimodal data for classification and prediction of Alzheimer's disease, doctoral thesis, Sorbonne university, Feb. 2021.
[54] H. V. Vinters, "Emerging concepts in Alzheimer's disease," Annu. Rev. Pathol, vol. 10, pp. 291 - 319, Oct. 2014.
[55] P. N. E. Young, et al., "Imaging biomarkers in neurodegeneration: current and future practices," Alzheimer's Research & Therapy, vol. 12, no. 49, pp. 1 - 17, Apr. 2020.
[56] S. Y. Park et al., "Analysis of cerebral blood flow with single photon emission computed tomography in mild subcortical ischemic vascular dementia," Nucl. Medicine Mol. Imag., vol. 48, no. 4, pp. 272 - 277, Dec. 2014.
[57] S. Shimizu et al., "Utility of the combination of DAT SPECT and MIBG myocardial scintigraphy in differentiating dementia with lewy bodies from Alzheimer's disease," Eur. J. Nucl. Medicine Mol. Imag., vol. 43, no. 1, pp. 184 - 192, Jan. 2016.
[58] E. Imabayashi et al., "Validation of the cingulate island sign with optimized ratios for discriminating dementia with lewy bodies from Alzheimer's disease using brain perfusion SPECT," Ann. Nucl. Medicine, vol. 31, no. 7, pp. 536 - 543, Aug. 2017.
[59] Liu. Manhua, et al., "A multi-model deep convolutional neural network for automatic hippocampus segmentation and classification in Alzheimer's disease," NeuroImage., vol 208, Mar. 2020.
[60] Li. Aojie, et al., "Hippocampal shape and asymmetry analysis by cascaded convolutional neural networks for Alzheimer's disease diagnosis," Brain Imaging and Behavior., vol. 15, no. 5, Jan. 2021.
[61] D.S. Marcus et al., "Open Access Series of Imaging Studies (OASIS): cross-sectional MRI data in young, middle aged, nondemented, and demented older adults," Journal of Cognitive Neuroscience, vol. 19, no. 9, pp. 1498 - 1507, 2007.
[62] D. Cárdenas-Peña, D. Collazos-Huertas, and G. Castellanos-Dominguez, "Enhanced data representation by kernel metric learning for dementia diagnosis", Front. Neurosci., vol. 11, no. 413, Jul. 2017.
[63] I. B. Malone, et al., "MIRIAD—Public release of a multiple time point Alzheimer's MR imaging dataset," NeuroImage., vol. 70, pp. 33 - 36, Apr. 2013.
[64] M. K. Singh, and K. K. Singh, "A Review of Publicly Available Automatic Brain Segmentation Methodologies, Machine Learning Models, Recent Advancements, and Their Comparison', Ann Neurosci., vol. 28, no. 1-2, pp. 82 - 93, Mar. 2021.
[65] N. Yamanakkanavar, J. Y. Choi and B. Lee, "MRI Segmentation and Classification of Human Brain Using Deep Learning for Diagnosis of Alzheimer's Disease: A Survey", Sensors 2020, vol. 20, no. 3243, pp. 1 - 28, June 2020.
[66] J. Zhang, et al., "Automatic craniomaxillofacial landmark digitization via segmentation-guided partially-joint regression forest model and multiscale statistical features," IEEE Trans. Biomed. Eng., vol. 63, no. 9, pp. 1820 - 1829, Dec. 2015.
[67] J. Zhang, Liu M, and D. Shen, "Detecting anatomical landmarks from limited medical imaging data using two-stage task-oriented Deep Neural Networks," IEEE Trans. Img. Proc., vol. 26, no. 10, pp. 4753 - 4764, Jun. 2018.
[68] B. Fischl, "FreeSurfer, " NeuroImage, vol. 62, no. 2, pp. 774 - 781, Aug. 2012.
[69] M. Jenkinson, et al., "FSL, " NeuroImage, vol. 62, no. 2, pp. 782 - 790, Aug. 2012.
[70] P. Coupé, et al., "Patch-based segmentation using expert priors: Application to hippocampus and ventricle segmentation," NeuroImage., vol. 50, no. 2, pp. 940 - 954, Jan. 2011.
[71] J.E. Iglesias, and M.R. Sabuncu, "Multi-atlas segmentation of biomedical images: A survey," Med Image Anal., vol. 24, no. 1, pp. 205 - 219, Aug. 2015.
[72] M. Liu, D. Zhang, and D. Shen, "View-centralized multi-atlas classification for Alzheimer's disease diagnosis," Hum Brain Mapp., vol. 36, no. 5, pp.1847 - 1865, May. 2015.

[73] D. Zarpalas, et al., "Accurate and fully automatic hippocampus segmentation using subject-specific 3D optimal local maps into a hybrid active contour model," IEEE J Trans Eng Health Med., vol. 2, no.1 - 16, Jan. 2014.
[74] K.M. Pohl, et al., "A hierarchical algorithm for MR brain image parcellation,"IEEE Trans Med Imaging., vol. 26, no. 9, pp. 1201 - 1212, Sep. 2007.
[75] Y. Hao, et al., "Local label learning (LLL) for subcortical structure segmentation: Application to hippocampus segmentation," Hum Brain Mapp., vol. 35, no. 6, pp. 2674 - 2697, Jun. 2014.
[76] A-P. Fard et al., "Sagittal Cervical Spine Landmark Point Detection in X-Ray Using Deep Convolutional Neural Networks," IEEE Access, vol. 10, pp. 59413 - 59427, Jun. 2022.
[77] P. Yi, et al., "Hippocampal Segmentation in Brain MRI Images Using Machine Learning Methods: A Survey", Chinese Journal of Electronics, vol. 30, no. 5, pp. 793 - 814, Sept. 2021.
[78] R. Cuingnet et al., "Automatic classification of patients with Alzheimer's disease from structural MRI: A comparison of ten methods using the ADNI database," NeuroImage, vol. 56, no. 2, pp. 766 - 781, May 2011.
[79] S. Leandrou, et al., "Quantitative MRI Brain Studies in Mild Cognitive Impairment and Alzheimer's disease: A Methodological Review", IEEE Reviews in Biomedical Engineering, vol. 11, pp. 97 - 111, Jan. 2018.
[80] M. Liu, et al.,, "Anatomical landmark based deep feature representation for MR images in brain disease diagnosis," IEEE J. Biomed. Health Informat., vol. 22, no. 5, pp. 1476 - 1485, Sep. 2018.
[81] K. Hett, et al., "Multi-scale graph-based grading for alzheimer's disease prediction," Medical Image Analysis, vol. 67, pp. 1 - 13, Jan. 2021.
[82] C. Ledig, et al., "Structural brain imaging in alzheimer's disease and mild cognitive impairment: biomarker analysis and shared morphometry database," Scientific reports., vol. 8, Jul. 2018.
[83] C-Y. Wee, P-T. Yap, D. Shen, "Prediction of Alzheimer's disease and mild cognitive impairment using cortical morphological patterns," Human brain mapping., vol. 34, no. 12, pp. 3411 - 3425, Dec. 2013.
[84] K. Hett, et al., "Graph of hippocampal subfields grading for alzheimer's disease prediction," In Int Work Mac Lear Med Imag., pp. 259 - 266, Sep. 2018.
[85] T. Tong, et al., "A novel grading biomarker for the prediction of conversion from mild cognitive impairment to Alzheimer's disease," IEEE Tran on Biom Engin., vol. 64, no. 1, pp. 155 - 165, Jan. 2017.
[86] C. Lian, et al., "Hierarchical Fully Convolutional Network for Joint Atrophy Localization and Alzheimer's Disease Diagnosis Using Structural MRI," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 42, no. 4, pp. 880 – 893, Dec. 2018.
[87] S. Basaia, et al., "Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks," NeuroImage: Clinical, vol. 21, Feb. 2019.
[88] L. Zhou, et al., "Hierarchical anatomical brain networks for MCI prediction : revisiting volumetric measures," PloS one., vol. 6, no. 7, Jul. 2011.
[89] R. Cuingnet, et al., "Spatial and anatomical regularization of svm: a general framework for neuroimaging data," IEEE Trans. on Patt Anal and Machine Int, vol. 35, no. 3, pp. 682 - 696, Jun. 2012.
[90] H-I. Suk, S-W. Lee, and D. Shen, "Deep ensemble learning of sparse regression models for brain disease diagnosis," Med Image Anal, vol. 37, pp. 101 - 113, Apr. 2017.
[91] C-Y. Wee, et al., "Cortical graph neural network for AD and MCI diagnosis and transfer learning across populations," NeuroImage : Clinical, vol. 23, Jul. 2019.
[92] K. Moodley, "Diagnostic differentiation of mild cognitive impairment due to Alzheimer's disease using a hippocampus-dependent test of spatial memory," Hippocampus, vol. 25, no. 8, pp. 939 - 951, Mar. 2015.
[93] X. Long, et al., "Healthy aging: an automatic analysis of global and regional morphological alterations of human brain," Acad. Radiol, vol. 19, no. 7, pp. 785 - 793, Jul. 2012.
[94] X. Long, et al., "Distinct laterality alterations distinguish mild cognitive impairment and Alzheimer's disease from healthy aging: statistical parametric mapping with high resolution MRI," Hum. Brain Mapp, vol. 34, no. 12, pp. 3400 – 3410, Dec. 2013.
[95] T-H. Kung, et al., "Neuroimage Biomarker Identification of the Conversion of Mild Cognitive Impairment to Alzheimer's Disease", Front Neurosci, vol. 15, no. 584641, pp. 1 - 16, Feb. 2021.
[96] S. Leandrou, et al., "Assessment of Alzheimer's disease based on texture analysis of the entorhinal cortex," Front. Aging Neurosci, vol. 12, no. 176, Jul. 2020.
[97] J. Ashburner, "SPM: A history," NeuroImage, vol. 62, no. 2, pp. 791 - 800, Aug. 2012.
[98] P-A. Yushkevich, et al., "Automated volumetry and regional thickness analysis of hippocampal subfields and medial temporal cortical structures in mild cognitive impairment," Hum Brain Mapp, vol. 36, no. 1, pp. 258 - 287, Jan. 2015.
[99] F. Bartel, et al. "Fast segmentation through surface fairing (FastSURF): A novel semi-automatic hippocampus segmentation method," PLoS One, vol. 14, no. 1, Jan. 2019.
[100] L. Henschel et al., "FastSurfer - A fast and accurate deep learning based neuroimaging pipeline," NeuroImage, vol. 219, Oct. 2020.
[101] Z. Xie and D. Gillies, "Near real-time hippocampus segmentation using patch-based canonical neural network", arXiv preprint, arXiv:1807.05482, 2018.
[102] J. Dolz, et al., "3D fully convolutional networks for subcortical segmentation in MRI: a large-scale study," Neuroimage, vol. 170, pp. 456 - 470, Apr. 2018.
[103] B. Thyreau, et al., "Segmentation of the hippocampus by transferring algorithmic knowledge for large cohort processing," Med Image Anal, vol. 43, pp.214 - 228, Jan. 2018.
[104] C. Wachinger, M. Reuter and T. Klein, "Deepnat: Deep convolutional neural network for segmenting neuroanatomy," NeuroImage, vol.170, pp.434 - 445, Apr. 2018.
[105] D. Chen, et al., "Enhancement mask for hippocampus detection and segmentation," IEEE Int Conf on Info and Automation, Fujian, China, pp.455 - 460, Feb. 2019.

[106] A-G. Roy, et al., "Quicknat: A fully convolutional network for quick and accurate segmentation of neuroanatomy," NeuroImage, vol.186, pp.713 - 727, Feb. 2019.
[107] N-K. Dinsdale, M. Jenkinson and A-I. Namburete, "Spatial warping network for 3D segmentation of the hippocampus in MR images," in Int Conf on Med Imag Comp and Computer-Assisted Intervention, Shenzhen, China, pp.284 - 291, Oct. 2019.
[108] D. Ataloglou, et al., "Fast and precise hippocampus segmentation through deep convolutional neural network ensembles and transfer learning," Neuroinformatics, vol. 17, no. 4, pp.563 - 582, Mar. 2019.
[109] S. Pang, et al., "Hippocampus segmentation based on iterative local linear mapping with representative and local structure-preserved feature embedding," IEEE Trans on Med Imag, vol. 38, no.10, pp. 2271 - 2280, Mar. 2019.
[110] Y. Liu and Z. Yan, "A combined deep-learning and lattice boltzmann model for segmentation of the hippocampus in MRI", Sensors, vol. 20, no. 13, Jun. 2020.
[111] C. Jia, et al., "Three-dimensional segmentation of hippocampus in brain MRI images based on 3CN-net," In Proc of the 3rd Int Conf on Innov in Art Intelligence, Suzhou, China, pp.17 - 20, Mar. 2019.
[112] B. Hou, et al., "Multi-target interactive neural network for automated segmentation of the hippocampus in magnetic resonance imaging," Cognitive Computation, vol. 11, no. 5, pp. 630 - 643, Jul. 2019.
[113] W. Yao, S. Wang, and H. Fu, "Hippocampus segmentation in MRI using side U-net model," Int Conf on Neural Inf Proc, Sydney, NSW, Australia, pp.143 - 150, Dec. 2019.
[114] I. Brusini, et al., "Shape information improves the cross-cohort performance of deep learning-based segmentation of the hippocampus," Front Neurosci, vol. 14, no. 15, Jan. 2020.
[115] M. Goubran, et al., "Hippocampal segmentation for brains with extensive atrophy using three-dimensional convolutional neural networks," Human Brain Mapping, vol. 41, no. 2, pp. 291 - 308, Feb. 2020.
[116] D. Carmo, et al., "Hippocampus segmentation on epilepsy and Alzheimer's disease studies with multiple convolutional neural networks," Heliyon, vol. 7, no. 2, Feb. 2021.
[117] S. Nobakht, et al., "Combined Atlas and Convolutional Neural Network-Based Segmentation of the Hippocampus from MRI According to the ADNI Harmonized Protocol," Sensors, vol. 21, no. 7, pp. 1 - 15, Apr. 2021.
[118] H. A. Helaly, M. Badawy and A. Y. Haikal, "Toward deep MRI segmentation for Alzheimer's disease detection," Neural Computing and App, vol. 34, pp. 1047 - 1063, Aug. 2021.
[119] L. Cao, et al., "Multi-task Neural Networks for Joint hippocampus Segmentation and Clinical Score Regression," Mul Tools and App, vol. 77, pp. 29669 - 29686, Jan. 2018.
[120] J. Sun, et al., "Dual-functional neural network for bilateral hippocampi segmentation and diagnosis of Alzheimer's disease," Int J Comput Assist Radiol Surg, Z. Liu, vol. 15, n. 3, pp. 445 - 455, Dec. 2019.
[121] Y. Shi, K. Cheng and Z. Liu, "Hippocampal subfields segmentation in brain MR images using generative adversarial networks," BioMed Eng OnLine, vol. 18, no. 5, Jan. 2019.
[122] H. Zhu, et al., "Dilated Dense U-Net for Infant Hippocampus Subfield Segmentation," Front Neuroinform, vol. 13, no. 30, pp. 1 - 12, Apr. 2019.
[123] Z. Yang et al. "CAST: A multi-scale convolutional neural network based automated hippocampal subfield segmentation toolbox," NeuroImage, vol. 218, May. 2020.
[124] J. Manjón, J. E. Romero and P. Coupe, "A novel deep learning based hippocampus subfield segmentation method," Scientific Reports, vol. 12, no. 1, Jan. 2022.
[125] J. E. Romero, P. Coupé and J. V. Manjón, "HIPS: a new hippocampus subfield segmentation method," Neuroimage, vol. 163, pp. 286 - 295, Dec. 2017.
[126] W. Lin, et al. "Convolutional Neural Networks-Based MRI Image Analysis for the Alzheimer's Disease Prediction from Mild Cognitive Impairment," Front. Neurosci, vol. 12, Nov. 2018.
[127] S. Basaia, et al., "Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks", Elsev. NeuroImage Clin. Vol. 21, 2019.
[128] C. Platero et al., "Discriminating Alzheimer's disease progression using a new hippocampal marker from T1-weighted MRI: the local surface roughness," Hum. Brain Mapp. vol 40, no. 5, pp. 1666- 1676, Apr. 2019.
[129] Y. Liu et al., "Deep Feature Selection and Causal Analysis of Alzheimer's Disease," Front Neurosci., vol. 13, Nov. 2019.
[130] R. Cui, and M. Liu, "Hippocampus analysis by combination of 3D DenseNet and shapes for Alzheimer's disease diagnosis," IEEE J of Bio & Health Infor, vol. 23, no.5, pp. 2099 - 2107, Sep. 2019.
[131] O. Lindberg et al., "Hippocampal shape analysis in Alzheimer's disease and frontotemporal lobar degeneration subtypes," J. Alzheimers Dis., vol. 30, no. 2, pp. 355 - 365, May. 2012.
[132] K. Popuri, et al., "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 Mapp, vol. 41, no. 14, pp. 4127 - 4147, Jun. 2020.
[133] M. Liu, et al., "Landmark-based deep multi-instance learning for brain disease diagnosis," Medical Image Analysis, vol. 43, pp. 157 - 168, Jan. 2018.
[134] D. Zhang et al., "Multimodal classification of Alzheimer's disease and mild cognitive impairment," NeuroImage, vol. 55, no. 3, pp. 856 - 867, Apr. 2011.
[135] J. Ashburner and K-J. Friston, "Voxel-based morphometry: The methods," NeuroImage, vol. 11, no. 6, pp. 805 - 821, Jun. 2000.
[136] J. Zhang, et al., "Detecting anatomical landmarks for fast Alzheimer's disease diagnosis," IEEE Tran on Med Imag, vol. 35, no. 12, pp. 2524 - 2533, Jun. 2020.
[137] M. Liu, C. Lian, and D. Shen, "Anatomical-landmark-based deep learning for Alzheimer's disease diagnosis with structural magnetic resonance imaging," in Deep Learning in Healthcare: Paradigms and Applications. Springer International Publishing, vol. 171, pp. 127 - 147, Nov. 2019.

[138] Y.-X. Zhao et al., "Region Ensemble Network for MCI Conversion Prediction with a Relation Regularized Loss," in Int Conf on Med Imag Comp and Comp-Assisted Interv. Springer, pp. 185 - 194, Sep. 2021.
[139] C. Lian et al., "Attention-guided hybrid network for dementia diagnosis with structural MR images," IEEE Trans on Cybernetics, vol. 52, no. 4, Apr. 2022.
[140] M. Ashtari-Majlan, A. Seifi and M. M. Dehshibi, "A multi-stream convolutional neural network for classification of progressive MCI in Alzheimer's disease using structural MRI images", IEEE Journal of Biom and Health Informatics, vol. 26, no. 8, pp. 3918 - 3926, Mar. 2022.
[141] LM. Wierenga, et al. "Greater male than female variability in regional brain structure across the lifespan", Wiley, Hum Brain Mapp., vol. 43, pp. 470 - 499, Sep. 2020.
[142] L. V. Eijk, et al., "Region-specific sex differences in the hippocampus," NeuroImage, vol. 215, pp. 1 - 11, July 2020.
[143] Z. Fei, et al., "A Novel deep neural network-based emotion analysis system for automatic detection of mild cognitive impairment in the elderly," Neurocomputing, vol. 468, pp. 306 - 316, Jan. 2022
[144] Z. Jiang, et al., "Automated analysis of facial emotions in subjects with cognitive impairment," PLOS One, vol. 17, no. 1, Jan 2022,
[145] M. Ursino, and G. Pirrazini, "Construction of a Hierarchical Organization in Semantic Memory: A Model Based on Neural Masses and Gamma-Band Synchronization," Cognitive computation, Sep. 2023.

Table 3
DL-based methods for HC subfields segmentation

Reference	Method	Methodology	Strengths	Weaknesses
[121]	UG-Net model and an adversarial model	A generative network based on the modified U-Net to conduct the 2D segmentation of the slices extracted from brain nuclear magnetic image +an adversarial network with CNN to discriminate the expert annotation and the segmentation images generated by generative network	- UG-net can retain more local information required for accurate segmentation of small targets. It performs on the dataset obtained from center for imaging of neurodegenerative diseases (CIND) for CA1, CA2, DG, CA3, Head, Tail, SUB, ERC and PHG in hippocampal subfields with DSCs of 0.919, 0.648, 0.903, 0.673, 0.929, 0.913, 0.906, 0.884 and 0.889 respectively	It achieves higher accuracy in the segmentation of the hippocampal and the larger subfields but it doesn't perform well enough in the smaller hippocampal subfields
[122]	▪ Dilated Dense UNet (DUnet)	- It integrates dilated dense network in the U-net to develop DUnet	- The integration of the dilated dense network allows to effectively fuse the low-level features in the contracting path and the high-level features in the expanding path which is not well achieved with the Unet structure and then obtain optimal results	- Enable to generalize on unseen data
	▪ Residual DUnet (ResDUnet)	- It uses residual connections to group every pair of convolutional layers in DUnet and obtain the ResDUnet for HC subfield segmentation	The use of residual connections ensures the promotion of information propagation and the acceleration the convergence -Improvement of the average Dice coefficient by 2.1 and 2.5% for infant hippocampal subfield segmentation, and 0.5 and 0.6% for adult hippocampal subfield segmentation, compared to the classic 3D U-net	- Especially useful for segmenting infant hippocampal subfields (small structures)
[123]	CAST	- It is based on a 3D multi-scale deep CNN	-Quickly and accurately segment the different hippocampal subfields in less than 1 min for a new subject - Reliable in segmenting small subfields such as CA2, CA3, and the entorhinal cortex (ERC) - Flexible and capable to segment raw images taken from two different MRI scanner vendors with different acquisition parameters without revising any settings and network architecture - Applicable for multiple imaging modalities	The training of CAST requires approximately three days on a single workstation with a high-quality GPU card
[124]	DeepHIPS	Novel 3D UNET-based segmentation method using deep supervision and low-resolution feedback to facilitate the training process	- Reliable for segmenting raw images - Robust with mean DICEs equals to 0.9116 and 0.9083 for Kulaga-Yoskovitz and Winterburn datasets	- Tested only on Kulaga-Yoskovitz and Winterburn datasets - The training datasets of the method are quite small to ensure a good generalization (especially in the case of Winterburn)

Table 4
Models based on HC features

Reference	Method or Biomarker	Key points	pMCI vs sMCI classification performances
[126]	2.5 D CNN model	- Reduce the amount input data by slicing data in different direction - Use CNN model pretrained with AD/NC to classify pMCI versus sMCI - Combine features related to HC and global features from the whole brain to improve the AD prediction	Accuracy of 73.04%
[127]	CNN-based method	- Reliable and general model, able to compare multi center data taken from different protocols scanners	Accuracy of 75%
[128]	The local surface roughness (LSR)	- Develop the LSR as a new biomarker for early AD detection - Prove that the CA1 and subiculum subfields of the HC are the most affected by the atrophy **Limit**: use unbalanced data (only 13 AD)	Accuracy of 74.3%
[129]	VGG model +CGAN	- Use the CGAN to discover the potential causal relationship between brain regions and AD - Prove that the ventricular system is a useful biomarker for AD diagnosis while the temporal lobes are crucial for early stages	Prediction accuracies of the model using the training dataset at baseline to predict AD in the test datasets at baseline, 6 months, 12 months, and 24 months are 0.8675, 0.8452, 0.8335, and 0.7463, respectively
[130]	3D-DenseNet +SHARPM-PDM	- Combine multi-level and multi-type features to enhance classification performances - Use 3D-DenseNet to extract local features and SHARPM-PDM for global shape analysis - Prove that local features are more relevant than global features for MCI converter prediction - **Limited on binary classification**	Accuracy of 75%
[132]	An ensemble-learning- based MRDATS classifier	-Propose a MRI DAT score based on volume features to measure the similarity between the subject's MRI patterns and the predefined AD patterns and assess the presence of AD structural atrophy patters in NC and MCI stages - Tested on different public datasets - Not very efficient particularly for early detection of AD as it is based only on volume features	- AUC of 81% for TTC of 6 months - AUC of 73% for TTC of up to 7 years
[96]	Entorhinal cortex texture and hippocampal volume features	- Investigate entorhinal cortex texture and hippocampal volume features for the early prediction of AD - Prove that entorhinal cortex changes are more relevant than hippocampal atrophy to identify MCI to AD converters	AUC of 78%
[60]	Multi-channels cascaded CNNs (3D-CNN+2D-CNN)	- Use 3D-CNN to extract global features from hippocampal shapes and then 2D-CNN to refine them and capture high level ones - Exploit the asymmetry between left HC and right HC - Better than the SHARPM-PDM for the extraction of global HC shape features - Use only global features and this can reduce the model performances	- Accuracy of 71% - AUC of 71.9%
[95]	- Curvature Analysis+ multi-layer perceptron classifier (MLP) - RCP biomarker	- Propose the RCP as a new important biomarker for MCI to AD converter - Focus on HC subfields instead of the entire HC - Prove that the presubiculum, the subiculum, CA1, CA4, and the molecular layer are the most important ROIs used for early prediction of AD	Accuracy of 79.95%

Table 5
Models based on the whole brain

Reference	Methodology	Dataset	Landmarks identification process	Strengths	Limits
[133]	A CNN-deep MIL: 1) MR pre-processing 2) Pre-selection of 40 anatomical landmarks 3) Landmark-based instance extraction 4) Multi-instance CNN classification	ADNI-1, ADNI-2 and MIRIAD	Data-driven landmark discovery algorithm via a group comparison between AD and NC subjects in the training set (the landmarks is defined according to their discriminative power in distinguishing between AD and NC)	- Adopt the knowledge learned from AD and NC subjects to guide the prediction of MCI conversion - AUC with transferred knowledge equal to 79.04% and accuracy equal to 78.34%	- Do not consider some confounding factors such as age and gender in the discriminative landmark discovery - Do not use any prior knowledge of previously discovered AD-related brain areas to identify the landmarks
[86]	A hierarchical fully convolutional network (H-FCN): 1) MR pre-processing 2) Location proposals 3) patch-level sub-networks 4) region-level sub-networks 5) subject-level sub-network	ADNI-1 and ADNI-2	Construct the voxel-wise anatomical correspondence across all linearly-aligned sMRIs according to the Colin27 template then use image voxels widely distributed over the whole template brain image as location proposals	- Use prior knowledge to efficiently filter out obviously uninformative voxels from selected location proposals - Unify the localization of discriminative brain regions in sMR images with feature extraction and classifier construction to be more seamlessly coordinated in a task-oriented manner - Use transfer learning - Use the network pruning strategy to remove uninformative patches and regions - Accuracy of 80.9% and AUC of 78.1%	-use fixed size of input image patches while structural changes caused by AD may vary across different locations
[137]	Anatomical-landmark-based DL for AD diagnosis: 1) MR pre-processing 2) A data-driven technique to locate discriminative anatomical landmarks from MR images 3) A landmark-based deep feature learning (LDFL) framework using CNN to jointly achieve features extraction and classification	ADNI-1, ADNI-2 and MIRIAD	An automatic landmark discovery method	Integrating the features extraction with the classification task boosts the AD/MCI diagnosis performances	
[138]	A region ensemble model using a divide and conquer strategy: 1) extract global feature map from the whole brain 2) fed each regional feature map to a region-based diagnosis subnetwork to extract discriminative features 3) construct an ensemble subnetwork for weighted fusion of the regional features with an attention mechanism to obtain the final description for the subject	ADNI-1 and ADNI-2	Design an ensemble sub-network to automatically identify discriminative regions in the whole brain via an attention module	- Propose a novel regularized loss criterion to regularize the training process - Achieve very good results (AUC ROC equal to 85.4% and accuracy of 85.9%)	use hand-crafted hyper-parameters which it is a cumbersome process instead of running proper cross-validation strategy
[81]	A multiscale graph-based grading method: 1) Preprocessing 2) Brain structure segmentation 3) Patch-based grading approach 4) Graph-based approach 5) Multiscale approach to analyze alterations of whole brain structures and hippocampus subfields at the same time	ADNI-1	It is not based on anatomical landmarks	- Exploit the complementarity of hippocampal subfields and whole brain analysis, and the complementarity of inter-subject similarity and intra-subject variability - Achieve an AUC of 85.5% and an accuracy of 80.6% – Prove that the most frequently selected brain structures are the temporal lobe, the postcentral gyrus, the anterior cingulate gyrus, the hippocampus and the precuneus and the most frequently selected hippocampal subfields are the CA1-SP, the CA1-SRLM, and the subiculum	Dependence of the method to the quality of segmentation maps
[139]	Attention-guided DL framework: 1) Image pre-processing 2) A backbone FCN to generate an attention map containing discriminative regions 3) A multi-branch HybNet to learn multi-level features	ADNI-1, ADNI-2 and AIBL	Use a backbone FCN automatically identify discriminative brain regions from the whole-brain sMRI yielding a DAM	- Take into account the variability in the individuals' brains - Accuracy of 82.7%	Performing discriminative regions localization and AD diagnosis independently
[140]	A multi-stream convolutional neural network: 1) Image pre-processing 2) Identify the most statistically significant anatomical landmarks 3) use multi-stream deep CNN to ensure the classification	ADNI-1	Multivariate T2 Hotelling statistical test	- Use rigid registration to eliminate morphological structure deformations - Reduce potential errors caused by inter-subject brain shape differences - Use transfer learning -AUC of 94.3% and specificity of 99.7%	- Evaluated only on ADNI-1

SONIA BEN HASSEN was born in Sfax, Tunisia, in 1986. She received the Diplôme d'Ingénieur degree (Hons.) in signals and systems and the M.Sc. degree (Hons.) in telecommunication from the Tunisia Polytechnic School, in 2009 and 2010, respectively, and the Ph.D. degree (Hons.) in telecommunication from the National Engineers School of Tunis in 2016. She is currently an Assistant Professor in the National school of Electronic and Telecommunications of Sfax (Department of Telecommunication) and member of the Advanced Technologies for Image and Signal (ATISP) laboratory. Her research interests include statistical signal processing, array processing with an emphasis on direction of arrival estimation for wireless communications, image processing and pattern recognition.

Mohamed Neji (Senior Member, IEEE) was born in Sfax, Tunisia, in 1979. He received the degree in computer science applied to management in 2003, and the PhD. degree in computer systems engineering in 2014. He is a Assistant Professor in the Data Engineering and Decisional Systems department in the National School of Electronics and Communications of Sfax, University Of Sfax. His research interests include the applications of intelligent methods (neural networks) to pattern recognition, and vision systems. He is the IEEE Tunisia Section treasurer (2017-2020) and the Student Activities Committee Chair (2023 and 2024). He is also an Leader auditor (ISO 9001 and 21001). He is a coordinator and member of many research projects.

Adel M. Alimi (Senior Member, IEEE) was born in Sfax, Tunisia, in 1966. He received the degree in electrical engineering, in 1990, and the Ph.D. and H.D.R. degrees in electrical and computer engineering, in 1995 and 2000, respectively. He is currently a Professor in electrical and computer engineering with the University of Sfax. His research interests include the applications of intelligent methods (neural networks, fuzzy logic, and evolutionary algorithms) to pattern recognition, robotic systems, vision systems, industrial processes, intelligent pattern recognition, learning, analysis, and intelligent control of large-scale complex systems.

Dr Hussain obtained his BSc (Hons) in Medicine from the University of St Andrews, UK and MBChB Medicine from the University of Edinburgh, UK, in 2017 and 2020 respectively. He is an Academic Core Psychiatry Trainee in Inverclyde Royal Hospital, an Honorary Clinical Lecturer at the University of Glasgow and an Executive Committee Member of the Royal College of Psychiatrists Digital Psychiatry Special Interest Group.

He has been an Associate Faculty Member of the Clinical Educator Programme since 2020, and undertakes a range of teaching and research activities in Edinburgh Medical School. He is an SSC1 Tutor and an Early Year Guide, on the MBChB programme. He also sits on the steering group of the Undergraduate Certificate in Medical Education (UCME), is the founding Co-ordinator of the FY Clinical Skills Tutor Scheme (CEP & EUMES), and works on medical education research projects.

He was a 2020-22 Foundation Fellow at the Royal College of Psychiatrists, as part of which he received CPD funding of ~£1500. He has previously held appointments as a Visiting Researcher at the Boston University School of Public Health, USA (2020) and the School of Computer Science and Engineering, Nanyang Technological University, Singapore (2018), and as a Visiting Student Researcher at the Nuffield Department of Surgical Sciences, University of Oxford, UK (2015-17). In 2022, he received the NIHR Mental Health TRC and Royal College of Psychiatrists Research Taster Award, as part of which he completed an academic placement at the Department of Psychiatry, University of Oxford.

Amir Hussain (Senior Member, IEEE) received the PhD degree in novel neural network architectures and algorithms for real-world applications from the University of Strathclyde, Glasgow, U.K. He is the director of the Centre of AI and Robotics, Edinburgh Napier University. His research interests include cognitive data science and trustworthy AI and NLP technologies to engineer the next-generation of smart and secure industrial and healthcare systems. He is elected as an Executive Committee member of the U.K. Computing Research Committee (UKRC) - the National Expert Panel of the IET and the BCS for U.K. computing research. He is the founding chief editor of the Cognitive Computation Journal (Springer)

Mondher FRikha is a full professor at National School of Electronic and Telecommunications of Sfax, Tunisia. Her research interests include Speech and audio processing, Signal and image processing and Pattern recognition. He is a leader of ATSIP Research Unit at Enet'Com, University of Sfax.

Declaration of interests

☒ 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.