

Multimodal Identification of Alzheimer's Disease: A Review

Guian Fang, Mengsha Liu, Yi Zhong, Zhuolin Zhang,
Jiehui Huang[†], Zhenchao Tang[†], Calvin Yu-Chian Chen^{*}

arXiv:2311.12842v1 [eess.IV] 6 Oct 2023

Abstract— Alzheimer's disease is a progressive neurological disorder characterized by cognitive impairment and memory loss. With the increasing aging population, the incidence of AD is continuously rising, making early diagnosis and intervention an urgent need. In recent years, a considerable number of teams have applied computer-aided diagnostic techniques to early classification research of AD. Most studies have utilized imaging modalities such as magnetic resonance imaging (MRI), positron emission tomography (PET), and electroencephalogram (EEG). However, there have also been studies that attempted to use other modalities as input features for the models, such as sound, posture, biomarkers, cognitive assessment scores, and their fusion. Experimental results have shown that the combination of multiple modalities often leads to better performance compared to a single modality. Therefore, this paper will focus on different modalities and their fusion, thoroughly elucidate the mechanisms of various modalities, explore which methods should be combined to better harness their utility, analyze and summarize the literature in the field of early classification of AD in recent years, in order to explore more possibilities of modality combinations.

Index Terms— Alzheimer's disease, Multimodality , Machine learning, Mild cognitive impairment, Early detection

I. INTRODUCTION

As one of the most common neurodegenerative disorder, Alzheimer's disease (AD) is affecting the elderly around the world. According to the latest data from the World Health Organization (WHO), it is expected that the number of dementia patients will reach 55 million in 2019, and this number will increase to 139 million in the coming 2050. Among them, Alzheimer's disease is the most common cause of dementia, accounting for approximately 60-80% of dementia cases. Alzheimer's disease is a degenerative neurological disease characterized by progressive loss of cognition and memory. Currently, the academic community usually believes that AD is related to the neurofibrillary tangles (NFT) and the extracellular Amyloid- β ($A\beta$) deposition, which cause neurons and synapses loss or damage, inflammation and brain tissue atrophy are other changes [1]. Alzheimer's disease can cause changes in brain structure and function, affecting patients from multiple aspects such as speech, emotion, and behavior. As the condition worsens, patients often become

disconnected from society, lose their ability to take care of themselves, and burden their families and society.

There is still no way to completely cure and reverse the progression of dementia. However, early, accurate, and comprehensive diagnosis of Alzheimer's disease can provide timely intervention and slow down the progression of the disease. Experts are increasingly recognizing the importance of early diagnosis of Alzheimer's disease. At present, the clinical examination methods of Alzheimer's disease mainly include: Cognitive Assessment, Non-neuroimaging Biomarkers, Voice and Speech examination, Posture examination, Neuroimaging examination etc.

With the rapid development of large language models (LLMs) like ChatGPT, there has been an emergence of conversational systems based on natural language processing techniques. These systems, including HuatuoGPT [2], BenTsao [3], and DoctorGLM [4], have shown promising performance in the field of medical diagnosis and consultation. Interestingly, we have found that this chatbot-based diagnostic model, utilizing ChatGPT, has exhibited relatively high intelligence. However, due to the complexity and multifactorial nature of Alzheimer's disease, the reliability of its inferences is often questionable, as they are solely based on textual input provided by the patient. The characteristics of a single modality may not be sufficient to support accurate early diagnosis. The changes caused by AD may also have similar manifestations in other diseases. A better approach would be to incorporate multiple modalities from the patient, including text, voice, images, and more, as diagnostic evidence. Multimodal diagnostic methods emerge as the times require. Based on the success of vision and language-based large models, multimodal diagnostic systems appear to be more robust and reliable. Multimodal diagnosis of AD poses a challenging problem with significant implications for the future. Therefore, this review will discuss the methods of multimodal AD diagnosis. Below, we will briefly introduce the diagnostic methods for each modality up to now, and finally discuss the methods for multimodal diagnosis of Alzheimer's disease (AD). More detailed content will be described in the main text.

A. Neuroimaging

Clinical trials have shown that AD will bring key changes to the patient's brain, such as the accumulation of the protein fragment beta-amyloid into clumps (called beta-amyloid plaques) outside neurons and the accumulation of an abnormal

*Corresponding author, [†]Project leader

form of the protein tau (called tau tangles) inside neurons [1]. Brain atrophy is another change, which is due to cell loss and decreased ability of cells to metabolize glucose (glucose is the main fuel for the brain). Thus, AD is associated with pathological amyloid deposition, structural brain atrophy, and altered brain metabolism [5]. In recent decades, major advances in neuroimaging techniques have made these techniques one of the most important biomarkers in the diagnosis of AD. Neuroimaging techniques offer valuable insights into the human brain. Structural magnetic resonance imaging (MRI) enables the detection of brain atrophy, while functional imaging modalities like positron emission tomography (PET) and functional MRI (fMRI) are capable of identifying hypometabolism [6]. Furthermore, metrics such as mean diffusivity (MD) and fractional anisotropy (FA) measured by diffusion tensor imaging (DTI) provide indications of a person's cognitive status. Additionally, electroencephalography (EEG) allows for the assessment of communication activity between nerve cells, while magnetoencephalography (MEG) measures the magnetic fields generated by currents flowing within neurons, providing insights into brain activity [7]. By employing these diverse techniques, researchers gain a comprehensive understanding of the brain's structure, function, and cognitive processes which can help them develop diagnostic methods.

However, multimodal imaging studies may offer various advantages over unimodal imaging studies. Multimodal imaging studies are able to study the temporal and topographical relations between many pathological variables, thus improving our understanding of pathophysiological interactions in the body. This approach allows direct comparison of the diagnostic capabilities of different imaging modalities in the same patient sample [8]. Lu et al. [9] found that a network classifier constructed using a combination of FDG-PET and structural MRI images outperformed a network constructed using structural MRI or FDG-PET alone. Liu et al. [10] used a zero-masking strategy for data fusion to extract complementary information from MR and PET to classify AD patients into four AD stages.

Multimodality imaging studies are equally challenging because of the large number of imaging markers that are potential candidates for predicting disease transformation. This situation leads to two related problems. First, as the number of candidate features increases, the risk of data overfitting increases, and second, as the number of candidate features increases, covariance in the predictor variables becomes more severe [8]. Whether these problems can be solved is the key to multimodal imaging research.

B. Cognitive Assessment

Based on the patient's cognitive abilities, several tests are available to assess the level of AD (Alzheimer's disease) and MCI (Mild cognitive impairment). These include: the MMSE (Mini-Mental State) [11], a simplified cognitive mental state test in the form of a score, which provides a quantitative assessment of cognitive state; The MoCA (The Montreal Cognitive Assessment) [12], which provides a rapid assessment of different levels of cognitive impairment; and the ADAS-Cog (Alzheimer's Disease Assessment Scale - Cognitive) [13],

another commonly used clinical and experimental cognitive assessment tool; SCIP (Severe Cognitive Impairment Profile) [14], SIB (Severe Impairment Battery) [15] and so on. Roalf et al. compared the MMSE with the MoCA and concluded that the MoCA as a global assessment tool is superior to the MMSE and provides a reliable and simple conversion method from MoCA to MMSE scores [16].

Most of these cognitive assessment methods are lengthy and complex and do not apply to all patients in all stages of dementia and do not perform well enough in terms of sensitivity [7], [17]. Although these cognitive assessment methods can provide a quantitative evaluation of cognitive status and help doctors understand the cognitive state of patients, they have limitations in terms of sensitivity. They may not capture subtle changes in certain cognitive domains. Some tests also require professional personnel for evaluation and interpretation, and the test results may be influenced by factors such as education and cultural background.

C. Non-neuroimaging Biomarkers

Biomarkers are objective measurements of biological or pathogenic processes aimed at assessing disease risk or prognosis, guiding clinical diagnosis or monitoring therapeutic interventions [18]. Changes in biomarkers can be obtained from neuroimaging on the one hand, and from changes in the composition of biofluids on the other. Combined with clinical approaches and cognitive tests, biomarkers will be more useful to accomplish an accurate assessment of cognitive impairment and its causes, even allowing clinicians to identify and detect pathology caused by AD before it occurs [19]. The section regarding neuroimaging-based biomarkers will be discussed in detail in the neuroimaging section.

Among the Non-neuroimaging Biomarkers and changes that have proven useful so far are hippocampal atrophy on decreased $A\beta_{42}$ in cerebrospinal fluid (CSF) [20], and increased tau protein and phosphorylated tau protein [21]. In addition, blood-based biomarkers have likewise attracted the attention of experts, blood samples can be obtained in a less invasive and cheaper way [22]. Similar to CSF, blood-based biomarkers can measure $A\beta_{42}$ and other forms of $A\beta$ proteins, as well as different forms of tau. Additionally, Neurofilament light chain (NfL) shows promise as a blood-based biomarker for AD. Moreover, in the process of diagnosing AD, genetic factors should not be overlooked. An example of a gene-based biomarker used to detect AD is the $\epsilon 4$ allele of the APOE gene. This genetic variation can be detected through blood samples or buccal swab samples.

These biomarkers can provide direct information about pathological processes such as abnormal protein deposition, inflammatory reactions, and neuronal damage, which help understand the development and progression of AD. However, they lack sufficient specificity, meaning that they may also be present in other neurological disorders or normal aging processes. Some biomarkers require the collection of specific samples, such as cerebrospinal fluid, which may involve a more specialized and technically demanding process, making it relatively less accessible for widespread use.

D. Posture

In recent years, some progress has been made in research on posture (mainly face and gait) in the diagnosis of Alzheimer's disease (AD). Although postures are not currently the primary diagnostic criteria for AD, they play an important role as an aid in early diagnosis and monitoring.

In terms of faces, studies have found that AD patients have problems with facial emotion expression and facial emotion comprehension. AD patients are impaired in facial expression recognition [23] and have significant emotion recognition deficits [24], Fiona et al. [25] found that in Alzheimer's disease, emotion processing deficits were only found in complex and cognitively demanding emotion recognition tasks, while behavioral performance in simple face processing and emotion matching tasks was within the normal range. This offers the potential to use facial expression analysis as a tool for early AD diagnosis and monitoring.

Gait is a complex cognitive task requiring coordination between a wide range of brain regions, and even in the milder stages of the disease, gait impairment may reflect dementia-induced neurodegeneration [26]. There is growing evidence that cognitive, sensory, and motor changes may precede the clinical manifestations of AD by several years [27]. Gait disturbances reported in early AD include slower gait, shorter stride length, lower cadence (longer stride time/gait cycle), and greater inter-stride variability [28]. Therefore, gait analysis is also considered to be an important tool for assessing motor function and cognitive status in AD patients. For the assessment of gait characteristics, Rosaria et al. suggested that gait characteristics can be divided into temporal, kinematic and kinetic characteristics [29].

In conclusion, facial and gait analysis has shown some potential in AD diagnosis. As technology continues to evolve and more research work is done, these aids are expected to be a useful addition to early AD diagnosis. However, facial and gait analysis is still in the research phase and more validation and standardization work is needed. Individual differences and other factors may have an impact on facial and gait performance, so further research is needed to determine their accuracy and reliability in AD diagnosis and monitoring.

E. Sound

Voice and Speech problems are considered to be one of the most typical symptoms of AD, which is a direct and unavoidable consequence of cognitive impairment [30]. It has been demonstrated that AD patients perform poorly on different language tests [31], presenting naming and word-finding difficulties (anomia) leading to circumlocution, as well as difficulty accessing semantic information intentionally, leading to a general semantic deterioration [32], and behaving differently from normal in some acoustic and rhythmic features [33]. This demonstrates the feasibility of using voice modality to diagnose Alzheimer's disease and that Voice and Speech can be used as an efficient, inexpensive, and easy-to-use tool to help in the diagnosis of AD.

People with AD differ from normal people in semantics, syntax, and rhythm, and researchers have been able to diagnose

AD by looking at a variety of features. The main conventional features used in Alzheimer's disease research are: Frequential Aspects (including interruptions, Voice periods, Fundamental frequency); Intensity (including Amplitude and Phonatory stability); Voice Quality (including noise); Biomechanical aspects (including Vocal fold body Movement Tongue movement) [34]. Many studies have also identified acoustic measures that are highly correlated with pathological speech features or speech alterations [35]. In recent years, with technological advances some new methods have been gradually invested in AD diagnosis, Fasih et al [36] first used eGeMAPS [37], emobase [38] and ComParE [39] feature sets as Alzheimer's disease empirical attempts to introduce and evaluate a new method to represent these acoustic features ADR (active data representation). Liu et al. [40] divide a person's speech data into multiple segments and use the extracted spectrogram features from the speech data to identify AD.

As a non-invasive and rapid diagnostic method, recognition technology based on patient voice data can effectively reduce medical costs compared to medical images that are difficult to obtain. Especially, natural language processing technology, signal processing technology, and deep learning technology have been developed significantly in recent years, and the technology based on automatic processing of voice signal records is gradually becoming mature.

F. Multimodal

Currently, most studies on Alzheimer's disease (AD) use a single data model for prediction, which may have limitations. Psychological or Cognitive appraisal questionnaires may be too subjective and may lack sensitivity [7]. Changes in both posture and voice may be influenced by factors unrelated to AD, such as normal aging. Neuroimaging also suffers from cost and availability issues (availability of PET and MRI scanning instruments varies widely between countries) and from patient bias (e.g., sensitivity to radiation exposure) [8].

It has been shown that fusing complementary information from multiple modalities can improve the diagnostic performance of AD. Multimodal data contains the fusion of complementary information(e.g., Magnetic Resonance Imaging (MRI), Positron Emission Tomography (PET) and genetic data) [41]. Landau et al. also found complementary information between acquired genetic, cerebrospinal fluid, neuroimaging, and cognitive measures [42].

However, there are still challenges in fusing data from multiple modalities to diagnose AD. To begin with, it is important to recognize that different types of data are inherently diverse. Each modality, such as neuroimaging and genetic data, exhibits distinct data distributions, varying numbers of features, and differing levels of diagnostic discrimination for conditions like Alzheimer's disease (AD) [43]. To fuse multimodal data, traditional approaches usually first perform feature selection for each modality separately, and then cascade the selected features used for diagnosis or prognosis. Nevertheless, this approach ignores the potential connection between different modal data [41].

The second challenge is the high dimensionality problem encountered in the fusion analysis of multimodal imaging for

diagnostic AD. When combining data from multiple modalities, the resulting dataset tends to have a high number of dimensions. For instance, a single neuroimaging scan, such as MR or PET images, contains millions of voxels. Classical methods [44], such as principal component analysis (PCA), independent component analysis (ICA), and linear discriminant analysis (LDA), are used in many studies to solve the high-dimensional problem of multimodal fusion analysis. These methods achieve attribute parsimony, but researchers need to put a lot of effort to analyze some important fusion features separately [44].

Lastly, there is the issue of incomplete data, where not all samples possess complete multimodal data. Typically, researchers opt to discard samples with missing data, thereby increasing the risk of sample loss. Alternatively, one approach involves interpolating the missing data using methods like zero interpolation, k-nearest neighbor (KNN), or Expectation Maximization. However, this interpolation approach may introduce unnecessary noise, subsequently compromising the model's performance.

Many researchers have proposed solutions to the above challenges. Zhang et al. [45] proposed a general framework based on kernel methods, which can effectively combine MRI, PET, and CSF features and naturally embed them into traditional support vector machines to effectively solve, achieving high accuracy in AD classification. Zhou et al. [43] proposed a three-stage deep feature learning and fusion framework, which utilizes multimodal neural image data (i.e. MRI and PET) and genetic data (i.e. SNP) to learn potential representations for each individual modality and joint potential representations for each pair of modalities in the first two stages. In the third stage, the classification model is learned using joint potential representations from all modality pairs. Janani et al. [46] used the stack denoising automatic encoder to process EHR and SNP data, used 3D Convolutional neural network (CNNs) to train MRI imaging data, cascaded these intermediate features and transferred them to the classification, indicating that the multi-modal data analysis using DL is better than the single-mode DL model.

G. The main highlights of this literature survey

(1) We have conducted a comprehensive review of the current mainstream AD diagnostic methods based on various modalities, summarizing the research progress and recent advancements in these modalities over the past five years.

(2) We have analyzed the latest research on the application of multimodal techniques in AD diagnosis and discussed the current challenges in multimodal fusion. We also present different solutions proposed by researchers in this field.

(3) We provide possible directions and suggestions for future multimodal AD diagnostic technologies.

II. NEUROIMAGING

Currently, the consensus regarding Alzheimer's disease (AD) is that individuals first enter the preclinical stage (such as being asymptomatic), then progress to mild cognitive impairment (MCI), and eventually develop AD dementia. MCI

patients are further classified into early-stage MCI (EMCI) or late-stage MCI (LMCI) based on their performance on cognitive screening tools, with LMCI patients having the highest risk of AD conversion, while EMCI has a relatively lower conversion rate [47]. However, relying solely on cognitive deficits to determine whether a patient is in the early or late stage of MCI is quite challenging, as traditional cognitive assessment tools often fail to provide accurate judgments. This is where neuroimaging methods prove valuable, as they can detect substantial changes within the brain and provide crucial clues by visualizing and measuring gross alterations in the brains of AD subjects [48]. Neuroimaging is not only applicable to individuals already diagnosed with AD dementia but also to MCI patients and even cognitively normal individuals in the preclinical stage, allowing for accurate assessment of the extent of the disease in patients. The Alzheimer's Disease Neuroimaging Initiative (ADNI) database (<http://adni.loni.usc.edu>) collects and analyzes a large-scale dataset consisting of neuroimaging, biomarkers, cognitive assessments, and clinical data, including clinical evaluations, brain imaging modalities (such as MRI, PET), and biological samples (such as cerebrospinal fluid, plasma). These data are utilized to investigate the pathological mechanisms of AD, biomarkers for disease progression, the interplay between genetics and environmental factors, and the evaluation of novel treatment approaches.

In this chapter, we extensively discussed several mainstream neuroimaging techniques (Table I) and explored their inter-relationships. Additionally, we analyzed computer-based approaches for diagnosing AD, aiming to provide more accurate prediction and diagnostic tools to anticipate cognitive decline and the progression from MCI to AD dementia in patients.

A. Magnetic Resonance Imaging (MRI)

In AD research, amyloid-beta (A β) is the most commonly studied biomarker, and there is substantial evidence indicating the presence of A β deposition in a large proportion of individuals at the preclinical stage [49]. β -amyloid protein (A β) is generated through the cleavage of the amyloid precursor protein by the β -secretase and γ -secretase complex. Once released as monomers, A β can form oligomers with cellular toxicity or neuroregulatory properties [50], leading to neurodegeneration and cognitive impairment. Therefore, it is crucial to detect the presence of A β before its deposition reaches a significant level.

Magnetic Resonance Imaging (MRI) can effectively measure the levels of A β in various brain regions. Changes associated with A β deposition are closely related to increased gray matter atrophy, particularly in the hippocampus. Thus, the hippocampus is one of the key regions affected in AD dementia. By scanning brain regions such as the hippocampus, entorhinal cortex, amygdala, medial and lateral temporal lobes, lateral ventricles, medial temporal gyri, and cortical gray matter, MRI can detect brain atrophy caused by severe neuronal loss [51].

Currently, MRI techniques are well-established. For example, the team at the University of California, San Francisco (UCSF) Medical Center utilizes FreeSurfer for cortical reconstruction and volume segmentation [52]. Detailed information

TABLE I
NEUROIMAGING TECHNIQUES IN AD DIAGNOSIS

Technique	Function	Advantages	Disadvantages
MRI	Provides detailed brain structural images, including brain volume, atrophy, etc.	Non-invasive, high resolution	Cannot directly observe metabolism and functional activity
DTI	Assesses integrity and connectivity of white matter fiber bundles	Provides white matter microstructure information, detects fiber bundle damage	Limited in fiber crossing and challenging areas
sMRI	Analyzes changes in brain tissue morphology and density	Non-invasive, provides structural information	Does not provide functional and metabolic information
fMRI	Detects brain activity and functional connectivity	Observes brain functional activity, provides functional connectivity maps	Limited sensitivity, affected by motion and noise
PET	Evaluates brain metabolism and function	Observes brain metabolism and functional activity, widely used in biomarker research	Requires radioactive tracers, expensive
FDG-PET	Assesses brain glucose metabolism	Detects abnormal brain glucose metabolism, associated with AD	Requires radioactive tracers, expensive, cannot directly observe other metabolic activities

regarding MRI data acquisition, preprocessing, and quality control can be found at (<https://ida.loni.usc.edu>).

B. Diffusion Tensor Imaging (DTI)

DTI is a magnetic resonance imaging (MRI) technique used to assess microstructural changes within the white matter fiber tracts of the brain [53]. It provides information by measuring the diffusion of water molecules in tissues. The most commonly used metric in DTI is Fractional Anisotropy (FA), which reflects the degree of diffusion anisotropy in different directions. A decrease in FA is often considered an indicator of axonal degradation and demyelination in the white matter of the brain [54].

Using DTI imaging to evaluate white matter fibers, studies have found significant fiber damage in the connections between the hippocampus and the posterior cingulate gyrus in AD patients. This suggests that white matter damage may be associated with gray matter atrophy in the temporo-parietal brain network implicated in AD. A recent meta-analysis of DTI studies in AD patients confirmed the significant role of fiber damage in the posterior cingulate gyrus and the major white matter tracts connecting the prefrontal cortex with the medial temporal or parietal cortex [55].

C. Structural MRI

Structural MRI is a technique used to assess brain changes in AD patients by detecting alterations in the gray matter and white matter. This technique is particularly sensitive to changes in gray matter volume due to neuronal loss and atrophy [56]. Brain atrophy is widely recognized as a hallmark of AD and is associated with the severity of the disease [57]. In Alzheimer's disease, neurons in the medial temporal lobe and hippocampus are particularly vulnerable to loss, which is consistent with the occurrence of significant cognitive impairment in clinical presentations. Studies have found that in the early stages of AD dementia, the volume of the hippocampus has already decreased by 20% [58] (Figure 1). Other brain

regions such as the lateral temporal lobe, parietal lobe, and frontal lobe have also been found to be adversely affected [59].

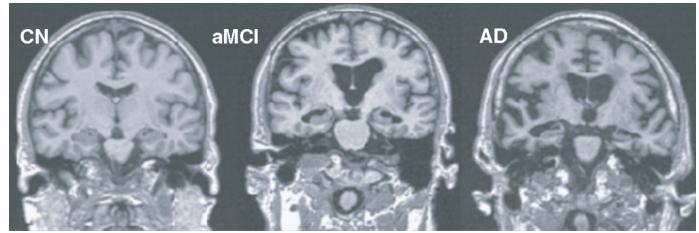


Fig. 1. MRI imaging of hippocampal volume structure in normal control group (CN), MCI and AD patients

D. Functional Magnetic Resonance Imaging (fMRI)

Functional Magnetic Resonance Imaging (fMRI) measures synaptic activity and assesses changes in brain during resting state and cognitive tasks. Studies have found reduced hippocampal activity in individuals with AD during episodic memory tasks, consistent with early memory deficits [58] (Figure 2).

In healthy individuals, the default mode network is active during resting state but deactivated during cognitive tasks. AD patients exhibit impaired intrinsic functional connectivity of the default mode network during resting state [60]. fMRI evaluation has shown that in the early stages of MCI, there is overactivation in the hippocampus during memory tasks, but a reduction in activity is observed in the preclinical stage, potentially indicating a sign of clinical decline. However, fMRI has limitations, such as the inability to directly measure neuronal activity, observe subcortical functional activity, and susceptibility to motion artifacts, which may result in false-negative and false-positive outcomes.

E. Positron Emission Tomography (PET)

In AD, the deposition of β -amyloid ($A\beta$) is a prominent feature. As the disease progresses, $A\beta$ leads to the pathological

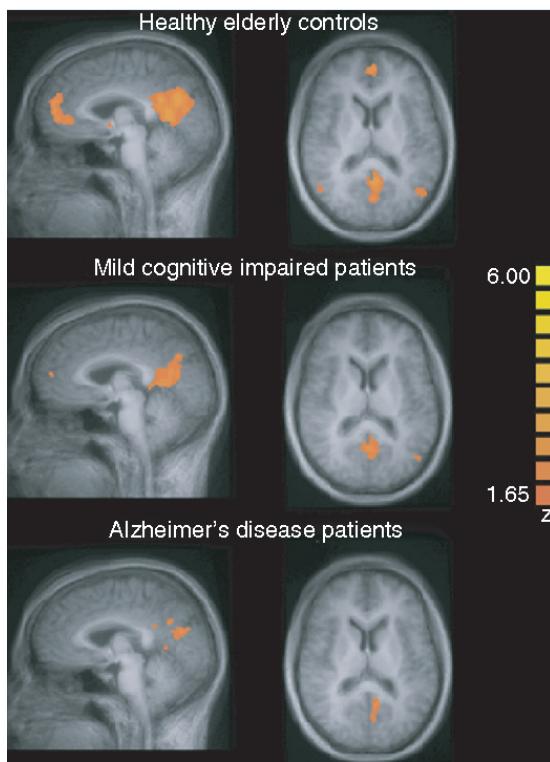


Fig. 2. Schematic diagram of fMRI default network function connections for normal control group, MCI, and AD patients

accumulation of tau protein (Figure 3) [61], resulting in the formation of 'senile' plaques and neurofibrillary tangles [62]. These abnormal accumulations are associated with an increased risk of developing AD [63]. Among the amyloid-binding compounds, ¹¹C-PIB and tau have been widely used, and their uptake can be measured using Positron Emission Tomography (PET).

Longitudinal PET scans have shown the accumulation of tau protein in AD patients. In the early stages, tau primarily accumulates in the medial temporal lobe, and later spreads to the lateral temporal lobe, as well as the superior and medial regions of the parietal lobe, before severe cognitive impairment occurs. Various radiotracers have been developed that selectively bind and highlight pathological structures in the central nervous system, including amyloid plaques, neurofibrillary tangles, activated microglia, and reactive astrocytes. Examples of compounds that enable amyloid plaque imaging include [¹⁸F]FDGDP, ¹⁸F-BAY94-9172, ¹¹C-S B-13, ¹¹C-B F-227, and ¹¹C-PIB.

Using different PET tracers, the deposition of A_β can be visualized and quantitatively measured throughout the brain. Current imaging agents include ¹¹C-labeled radioactive tracer PiB or ¹⁸F-labeled tracers such as [¹⁸F]florbetaben, [¹⁸]florbetapir (also called [¹⁸]AV-45), and [¹⁸]flutemetamol.

F. ¹⁸F-Fluorodeoxyglucose PET (FDG-PET)

In dementia patients, ¹⁸F-FDG-PET (¹⁸F-labeled fluorodeoxyglucose positron emission tomography) can detect the reduction of brain metabolism, which is a hallmark of neurodegeneration. ¹⁸F-FDG-PET measures the regional glucose

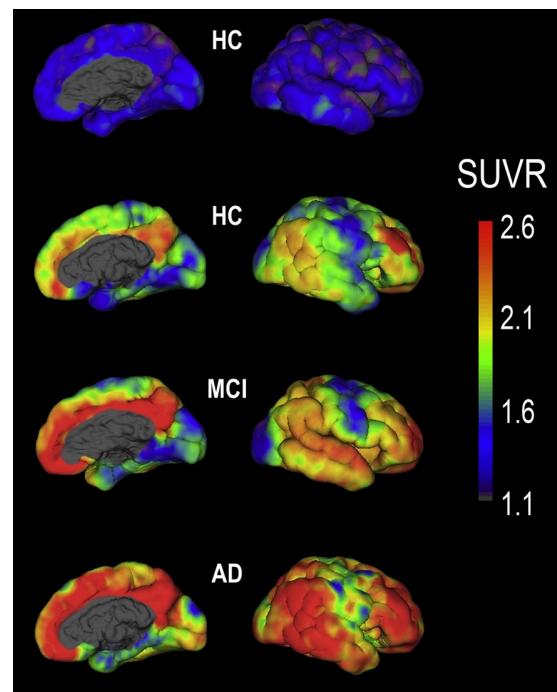


Fig. 3. PET images of AD patients and healthy controls. It can be observed that AD patients have significant A_β deposition in the brain.

consumption directly associated with the local intensity of glutamatergic synaptic and astrocytic activity in the brain. It can assess the degree and location of decreased brain metabolism, reflecting impaired neuronal function [64] (Figure 4).

¹⁸F-FDG-PET is particularly useful in early diagnosis as it can reveal the characteristic pattern of neurodegeneration in individuals with mild cognitive impairment associated with Alzheimer's disease earlier than MRI. It detects early changes related to AD pathology more sensitively and can be used to predict the transition from cognitive normality to mild cognitive impairment.

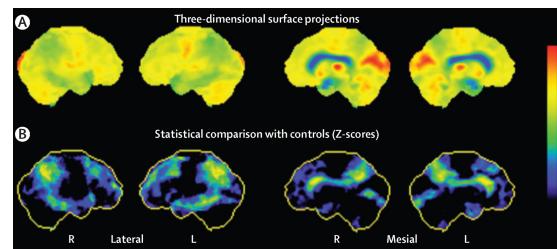


Fig. 4. Figure A shows an FDG-PET image of an AD patient, with yellow or red indicating hyperglycemia and green indicating hypometabolism. Figure B shows the comparison between the FDG-PET scan of this patient and the healthy population, with green or yellow indicating high deviation, and black or blue indicating no deviation or low deviation.

G. Neuroimaging Analysis Methods

Several classical machine learning techniques have been used to explore Alzheimer's disease (AD), ranging from image decomposition techniques like principal component analysis to more complex nonlinear decomposition algorithms. With

the advent of deep learning paradigms, it is now possible to directly extract high-level abstract features from MRI images [65].

The methods for diagnosing AD based on neuroimaging can be broadly categorized into four types [65]: slice-based, patch-based, voxel-based, and ROI-based. Slice-based methods assume that useful features are contained within 2D image slices, reducing the number of learnable parameters. Voxel-based methods are the most direct, using voxel intensity values from the entire 3D brain scan. Patch-based methods extract features related to AD patterns by extracting small 3D cubes (called "patches") from the brain. ROI (Region of Interest) methods emphasize specific regions in the brain that are related to AD rather than the entire brain. The diagnostic process often requires prior knowledge of abnormal regions associated with AD, such as the hippocampus.

Slice-based methods simplify certain important regions into 2D images, reducing the number of hyperparameters. Farooq et al. [66] performed slice-based axial scanning of GM volume in the images, discarding slices that did not contain informative information. Convolutional neural networks (CNNs) have been shown to have a revolutionary impact in the field of neuroimaging. Recent studies have utilized 2D CNNs for Alzheimer's disease (AD) detection on magnetic resonance imaging (MRI) scans, where each MRI scan is divided into multiple 2D image slices. The CNN is trained by computing the loss function between the labels of each subject and the predicted outputs of each image slice. Wang et al. [67] propose a VGG-inspired network that integrates convolutional block attention modules into the VIN backbone. However, a major drawback of slice-based methods is that 2D CNNs cannot understand the dependencies between voxels in the images. Converting 3D MRI scans into 2D image slices results in data loss. This is because brain regions span multiple 2D slices of the MRI scan, and the features related to the size and shape of the brain region will be lost after slicing. Ebrahimi et al. [68] introduced Temporal Convolutional Networks (TCNs) to address this issue. TCNs model the MRI feature sequences generated by CNNs to capture time-based dependencies, which are used for AD detection. TCNs can understand the temporal dependencies between 3D image slices within the 2D MRI volume.

Voxel-based methods capture all 3D information within a single brain scan, but they often treat all brain regions uniformly without considering specific anatomical structures. At the voxel level, tissue density (e.g., GM/WM) is commonly used as a feature for classification algorithms. 3D CNNs are typically employed in voxel-based methods [69]. Murcia et al. [70] modeled the dataset by applying Convolutional Autoencoders (CAEs). CAEs can directly extract data-driven features from three-dimensional images without the need for feature extraction beforehand. However, voxel-based methods are structurally complex and require a large number of training parameters, which can lead to overfitting. Additionally, voxel-level feature representations have high dimensionality [71]. Therefore, dimensionality reduction of features becomes a primary challenge in improving the classification performance of voxel-based methods.

Patch-based methods offer an intermediate scale between voxel-level and slice-level approaches, allowing for more effective capturing of local structural changes in images. By extracting features from patches (also known as patches) sampled randomly from MR images, many weak classifiers can be built and combined to make the final decision for AD diagnosis [72]. This approach involves training a fully convolutional network (FCN) using randomly sampled patches from the entire MRI volume. With the trained FCN, high-risk voxels can be selected and fed into a multilayer perceptron (MLP) for individual-level AD classification. Lian et al. [73] proposed a hierarchical fully convolutional network (H-FCN) that learns multiscale feature representations (e.g., patch-level, region-level, and subject-level) from sMRI scans, constructing a hierarchical classification model.

In recent years, patch-based methods have gained wide application in deep learning. In AD diagnosis, gray matter (GM) images of the brain are typically divided into 3D patches based on regions defined by the Automated Anatomical Labeling (AAL) atlas. These patches are then used to train various deep learning models. Through this approach, the models can learn to extract high-level features and patterns from the patches for AD classification and diagnosis. Dense networks [70], 2D and 3D convolutional neural networks (CNNs) [74] [75] [76] [77], and residual networks [78] have been used to learn features from sMRI for AD diagnosis. Zhu et al. [79] proposed a dual-attention multiple instance deep learning network (DA-MIDL), which consists of patch networks with spatial attention blocks to extract discriminative features from each sMRI patch. The method also employs attention-based multiple instance learning (MIL) pooling operations to balance the relative contributions of each patch. Patch-based methods capture disease-related patterns in the brain by extracting features from small image patches. The main challenge in this approach is selecting image patches with the most informative content to capture both local (patch-level) and global (image-level) features.

The ROI-based methods, which rely on pre-segmented regions of interest (ROIs), have also been widely used in AD diagnosis. This process typically involves three main components: 1) predefining the ROIs of interest, 2) extracting imaging features, and 3) constructing a classification model. With the emergence of machine learning (ML) methods, automatic feature extraction techniques have been employed [80] [81] [82]. Heckemann et al. [83] proposed a multi-atlas propagation with enhanced registration (MAPER), the first automatic whole-brain multi-region segmentation method that enables automatic segmentation of acquired MRI images. Liu et al. [84] used the MAPER method for feature extraction and selection, segmenting 15 brain structures in each MRI image and calculating the gray matter (GM) volume of brain regions as candidate predictive factors. Finally, by utilizing classifiers such as support vector machines (SVM), the impact of predictive factors on the AD progression can be evaluated. Li et al. [85] employed principal component analysis (PCA) for dimensionality reduction to obtain highly discriminative features by removing redundant features. However, ROI-based methods still have some limitations. The definition of ROIs

requires researchers to accumulate extensive experience, and the segmentation of ROIs can be influenced by individual differences and subjective factors among experts. Moreover, morphological abnormalities caused by brain diseases may involve multiple ROIs or partially predefined ROIs, which may result in unstable performance.

Lee et al. [86] proposed a feature representation method that combines voxel-based, region-based, and patch-based approaches, known as a hybrid method. Specifically, they employed a deep neural network (DNN) to learn complex voxel relationships within each region. In this approach, the brain was divided into predefined regions (region-based method), and the intricate nonlinear relationships between voxels (voxel-based method) were determined based on the anatomical shape of each region (patch-based method). By integrating these three approaches, the proposed method aimed to capture comprehensive information from different spatial scales, improving the representation of brain features for AD diagnosis.

III. CLINICAL DATA OF BIOMARKERS

Although neuroimaging plays a crucial role in the diagnosis and research of Alzheimer's disease (AD), it may not provide sufficient sensitivity and specificity, especially in early diagnosis, due to the complex nature of AD. Therefore, there is a need to identify more specific biomarkers to detect pathological changes in AD at an earlier stage. In the field of AD research, the identification and detection of biomarkers play a vital role in early diagnosis, monitoring disease progression, and evaluating treatment efficacy. Various methods have been developed for biomarker detection, providing valuable insights into the underlying pathological processes of AD. This section provides an overview of four commonly used biomarker detection methods: cerebrospinal fluid (CSF) analysis, plasma analysis, APOE gene variants, and assessment of oxidative stress and inflammation (Table II). To facilitate AD research and early diagnosis, the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (<http://adni.loni.usc.edu>) was launched in 2003, aiming to develop and validate biomarkers for early detection and treatment of AD.

A. Cerebrospinal Fluid and Blood Plasma Sampling

Sampling of cerebrospinal fluid (CSF) and plasma is one of the most direct and convenient methods for studying biochemical changes occurring in the central nervous system. Compared to brain biopsy or direct insertion of microdialysis probes into the brain, CSF and plasma sampling methods are less invasive and easier to obtain [87]. Through the collection of CSF and plasma samples, valuable information about biomarkers related to the central nervous system can be obtained. These biomarkers include amyloid- β (A β), total tau (Tau), and phosphorylated p-TAU_181p (pTau). CSF samples can be collected through lumbar puncture, and in some cases, from the fourth ventricle or lateral ventricle puncture. Studies have found a significant decrease in the average concentration of A β 42 in the CSF of Alzheimer's disease (AD) patients [88], which is believed to be associated with the deposition

of A β protein in plaques. According to the "amyloid cascade hypothesis," these deposits prevent A β 42 from entering the CSF, leading to a decrease in its concentration. Low levels of CSF A β 42 may also be indicative of amyloid deposition [89].

The advantages of CSF collection include relatively low cost and providing insights into neurodegeneration, tau protein, and amyloid pathology [64]. However, CSF biomarkers alone cannot provide detailed information about the extent and progression of pathology or neurodegeneration over time. Additionally, CSF analysis cannot directly determine the location and severity of pathology. Therefore, when using CSF biomarkers for diagnosis and disease progression monitoring, it is often necessary to combine them with other clinical and imaging examinations to obtain a more comprehensive evaluation.

Compared to cerebrospinal fluid (CSF), plasma sampling is less invasive and more cost-effective. However, the utility of plasma A β 42 as a biomarker for Alzheimer's disease (AD) is still unclear, as it may not accurately reflect A β changes in the brain. Some studies have shown that plasma A β 42 and A β 40 do not reflect A β accumulation in the brains of AD patients [90] [91]. Most studies have found no significant changes in plasma A β levels in sporadic AD patients [92] [93]. However, there are also studies that have found increased [94] or decreased [95] plasma A β 42 levels in AD patients. Furthermore, some studies have suggested that the ratio of A β 40/A β 42 in plasma can predict the transition from normal cognition to mild cognitive impairment and AD [96]. Nakamura et al. [97] efficiently measured plasma amyloid- β biomarkers using immunoprecipitation and mass spectrometry combined methods, demonstrating the potential clinical utility in predicting brain amyloid- β burden at an individual level. While plasma sampling is less invasive and more easily accessible, the use of plasma biomarkers for AD diagnosis and progression monitoring is still under investigation. Further research is needed to better understand the relationship between plasma biomarkers and the underlying pathophysiological changes in AD.

In addition to A β protein, other plasma biomarkers such as pTau forms, Neurofilament light chain, and Glial fibrillary acidic protein have also been found to be associated with AD [98]. Researchers have analyzed the correlation between plasma biomarkers and AD and used statistical methods to assess the strength of the association between plasma biomarkers and A β -PET burden. They found a significant correlation between plasma biomarkers and regions of high A β deposition in the brain [97]. These findings suggest that plasma biomarkers may provide valuable insights into the pathological processes associated with AD and have the potential to serve as non-invasive tools for AD diagnosis and monitoring. However, further research is still needed to fully understand the relationship between plasma biomarkers and AD and to validate their clinical utility.

Hansson et al. [99] conducted statistical analysis using Spearman's correlation coefficient to evaluate the predictive value of plasma A β as a factor for the subsequent development of AD. Verberk et al. [100] compared baseline demographic and clinical characteristics using t-tests, Mann-

TABLE II
BIOMARKER DETECTION METHODS

Biomarker Method	Characteristics	Pros and Cons
Cerebrospinal Fluid (CSF)	<ul style="list-style-type: none"> - Provides direct information about neurological diseases - Allows detection of various biomarkers such as amyloid-beta and tau proteins 	<ul style="list-style-type: none"> - Detection method is relatively complex, requiring lumbar puncture or intrathecal injection for sample collection - Invasive sampling may cause discomfort and risks
Blood Plasma	<ul style="list-style-type: none"> - Non-invasive and easy to collect - Allows detection of certain biomarkers such as proteins and metabolic products 	<ul style="list-style-type: none"> - Biomarker concentrations in plasma are relatively low, limited detection sensitivity - Susceptible to interference from other factors such as diet and medication
APOE Allelic Variants	<ul style="list-style-type: none"> - Genetic biomarker for assessing individual genetic risk - Detected through genetic typing or sequencing 	<ul style="list-style-type: none"> - Genetic variants associated with AD risk exhibit diversity - Genetic factors only partially explain AD risk
- Oxidative Stress and Inflammation	<ul style="list-style-type: none"> - Reflects the activity and extent of cellular oxidative damage and inflammatory response - Evaluation can be done by detecting related biomarkers - Requires further validation and replication of research findings 	<ul style="list-style-type: none"> - Oxidative stress and inflammation levels are influenced by multiple factors - Lack of specificity, potential overlap with other diseases - Consistency and reliability of results need confirmation

Whitney U tests, and chi-square tests. In addition, they used logistic regression analysis and receiver operating characteristic (ROC) curve analysis to assess the association between plasma biomarkers and abnormal amyloid status based on CSF and PET [101]. These analyses provide valuable insights into the relationship between plasma biomarkers and AD and contribute to the understanding of their diagnostic and prognostic value.

B. Genetics

In addition to the methods mentioned above, there are genetic factors associated with AD that need to be considered. The APOE gene has three common alleles: APOE ϵ 2, APOE ϵ 3, and APOE ϵ 4. Among them, the APOE ϵ 4 allele is considered a major genetic risk factor for AD [102].

In AD research, the APOE gene and its polygenic hazard score (PHS) are of great importance. The polygenic hazard score (PHS) is based on 31 single nucleotide polymorphisms (SNPs) and can reliably identify AD risk in individuals of any age [103]. It has been found that the effect of APOE* ϵ 4 on AD risk is mediated through the inhibition of amyloid-beta (A β) clearance and promotion of A β aggregation [104] [105]. Furthermore, carrying the ApoE ϵ 4 allele is associated with accelerated A β deposition, hippocampal atrophy, and disruption of the default mode network, and these effects can manifest from the preclinical stages of AD. APOE4 promotes the pathogenic mechanisms of AD by impairing microglial reactivity, lipid transport [106], synaptic integrity and plasticity [107], glucose metabolism [108], and cerebrovascular integrity and function [109]. Although all isoforms of apolipoprotein E

(APOE) are associated with A β deposition in the brain, the APOE ϵ 4 allele confers a significantly higher risk for AD compared to the presence of APOE ϵ 2 or APOE ϵ 3. Therefore, incorporating APOE genotyping into clinical decision-making and treatment planning can enable more personalized and precise management and treatment of AD.

C. Oxidative Stress and Inflammation

There is growing evidence that oxidative damage plays a significant role in the pathological processes of Alzheimer's disease (AD). Inflammatory processes associated with AD pathology, including microglia and astrocytes surrounding plaques, are also implicated in the disease [110]. In addition to their potential direct involvement, the secretions of these inflammatory cells, such as acute-phase proteins, alpha-1 antichymotrypsin (ACT, also known as SERPINA3 protein), alpha-2 macroglobulin (α 2M), as well as activators of the classical pathway of complement and cytokines like interleukin-1 β (IL-1 β) and tumor necrosis factor-alpha (TNF- α , also known as TNF), persist within plaques [97].

In the pathogenesis of Alzheimer's disease, free radicals also play a significant role in neuronal damage. One important consequence of free radical damage is lipid peroxidation, leading to the generation of F2-isoprostanes, which serve as biomarkers of this pathological mechanism. Multiple studies have shown that levels of F2-isoprostanes in the cerebrospinal fluid of AD patients are higher compared to healthy elderly individuals or non-AD dementia patients [111].

IV. POSTURE

As research on Alzheimer's disease (AD) advances, it has become apparent that relying solely on traditional neuroimaging and biomarker-based methods may not fully capture the early changes and subtle pathological features of AD. Therefore, finding a complementary approach to improve the accuracy and sensitivity of AD diagnosis has become increasingly important. In recent years, there has been growing interest in the use of posture assessment (including facial and gait analysis) in the diagnosis and study of AD. Posture, as a comprehensive expression of movement and emotion, can reflect changes in multiple aspects such as motor control, balance, and cognitive function. It can be utilized to identify early pathological changes in AD and evaluate disease progression.

A. Facial Analysis

In interpersonal communication, the recognition and production of facial expressions play a vital role in conveying emotional experiences and are closely linked to emotions, cognition, and behavioral adaptation. However, as AD progresses, patients may experience difficulties in recognizing emotional facial expressions. Through experiments, researchers have found that AD patients have impaired facial emotion expression abilities [112]. They lose some of their capacity to understand and express emotional rhythms [113]. Significant differences have been observed between AD patients and healthy older adults in emotional and cognitive tasks. When unable to discern the emotions of others, patients may become more confused and agitated, leading to abnormal behaviors such as sadness, reluctance to engage in conversations, or throwing objects. Therefore, studying the emotional recognition abilities of AD patients is of great importance for their treatment and care [114]. Additionally, communication difficulties in the early stages of AD may arise from an inability to correctly identify others' emotions, leading to impaired early emotional attention. Therefore, studying facial emotion recognition in AD patients can contribute to early AD diagnosis.

In addition to the decline in cognitive abilities associated with AD, problems can also arise in the production of facial expressions. The amygdala, which is known to be interconnected with the frontal cortex, plays a crucial role in rapid responses to emotional stimuli and regulates emotion-related autonomic reactions [115]. However, some studies have shown that AD patients perform poorly on emotional tasks and exhibit deficits in emotional processing [116]. This can result in their facial expressions appearing somewhat rigid. Therefore, analyzing specific facial expressions made by AD patients can contribute to AD diagnosis to some extent.

B. Gait Behavior

In normal aging, there is a decline in executive attention. AD patients, who have neurodegenerative disorders, exhibit more pronounced deficits in the nervous system, leading to impairments in gait speed and stride length, resulting in gait abnormalities such as slower walking speed and increased gait variability (i.e., fluctuations in stride length or timing)

that occur in the pre-dementia stage. Gait has been shown to be strongly associated with cognition [117]. Vergheze et al. studied three gait parameters - pace, rhythm, and variability - and proposed that quantitative gait impairments serve as markers of preclinical dementia [118]. Ríona analyzed gait features such as step, swing time, step velocity, step length variability, and stance time asymmetry, suggesting that the association between gait variability and cognitive impairment reflects neurophysiological changes [119]. Therefore, gait detection can be used to predict dementia diagnosis [120]. As gait is a complex cognitive task that requires coordination between widespread brain regions to select and organize motor programs suitable for the desired action [121], even in the early stages of the disease, gait abnormalities may reflect neurodegeneration due to dementia. Therefore, gait analysis may serve as a useful screening tool to differentiate mild dementia from normal aging.

C. Assessment Methods

1) Facial Emotion Understanding: Gressie et al. [112] used the Facial Affect Selection Task (FAST) [122] to assess facial emotion recognition in AD patients and compared their performance with a healthy control group. Participants were shown arrays of seven faces from the same person expressing six basic emotions (happiness, anger, sadness, surprise, fear, and disgust) and a neutral expression, and were asked to indicate the target indicated by verbal prompts.

Similar to the FAST task, Jiskoot et al. [123] developed the Emotion Recognition Task (ERT) to investigate emotion recognition deficits in AD patients. The ERT is a computerized neuropsychological test that can be obtained through the DiagnoseIS neuropsychological assessment system (www.diagnoseis.com). Participants are required to select one of six options (i.e., anger, disgust, fear, happiness, sadness, and surprise) to label facial emotional expressions.

Shirley et al. [114] combined facial emotion recognition with a rhythm task through seven subtasks to study the abilities of AD patients in emotional prosody and emotion recognition. In their study, individuals with AD showed comparable emotional rhythm processing abilities to the healthy control group, but their facial emotion processing abilities were impaired, especially in recognizing negative emotions, with AD patients demonstrating poorer results.

2) Facial Emotion Production: Facial emotion expression can be elicited voluntarily, and many researchers induce spontaneous facial expressions by having participants view movies, video clips, or read books. Keith et al. [116] constructed two sets of stimuli using the International Affective Picture System (IAPS) and recorded participants' facial expression data using facial electromyography (EMG) before and during the presentation of the images. Mograbi et al. [124] developed two novel experimental success-failure manipulation (SFM) paradigms, in which participants were instructed to perform tasks arranged in the experiment, and researchers observed the responses of the normal and AD groups during successful and failed tasks to analyze differences in facial emotion production between AD patients and healthy individuals. The results

revealed that while AD patients did not show significant differences in emotion compared to normal individuals when the tasks were successfully completed, they were more likely to exhibit negative emotions when the tasks failed.

3) Gait Data: Different types of sensors are used to collect real-time statistical data on human gait, which can be broadly categorized into two main classes: wearable sensors and environmental sensors. Wearable sensors are typically placed on various parts of the patient's body, and the captured data is either transmitted via wireless connection or collected on onboard storage devices [125] [126]. Environmental sensors, on the other hand, are installed in the environment and do not require older adults to wear them. For example, Magnus et al. [127] measured objective spatiotemporal gait parameters using an electronic walkway, where participants were instructed to walk on the walkway at a comfortable, self-selected pace. Rosaria et al. [128] collected data using a stereo-photogrammetric system with eight infrared cameras. If the combination of the first two categories is considered, a third category can also be obtained. In this case, wearable and environmental sensors are used together to form a hybrid system.

4) Facial Feature Extraction: Static facial feature extraction methods can be categorized into geometric-based or appearance-based approaches. In geometric-based methods, the shape and location of facial organs are extracted to form feature vectors representing the facial geometry. Appearance-based methods, on the other hand, utilize image filters (such as Gabor wavelets) applied to the entire face or parts of the face to detect variations in facial appearance. Keith et al. [116] used EMG electrodes connected to the face to extract features for facial emotion recognition. By analyzing the electromyographic activity of the zygomaticus major and corrugator muscles, the facial expressions of the participants were analyzed. In addition to using facial electrode sensors, many studies analyze facial expressions by directly capturing images of the person's face. Ekman and Friesen proposed the Facial Action Coding System (FACS) [129], in which facial expressions are composed of 46 action units (AUs) that correspond to specific sets of facial muscles, and the combination of AUs can form different emotions. Mograbi et al. encoded 29 facial behaviors involved in emotional expression [124], and facial feature extraction can be used to classify captured expressions.

In addition to capturing static photos, some researchers diagnose AD by analyzing facial feature videos. Static methods typically extract facial features from a single image or frame within an image sequence, while dynamic methods extract facial features from a series of images with time information, utilizing multiple image sequences as input to classifiers that leverage temporal information [130]. Dynamic methods are sensitive to subtle changes in the face and can effectively utilize temporal information.

5) Gait Feature Extraction: Gait detection features can be categorized into three main types: spatiotemporal features, kinematic features, and kinetic features.

Spatiotemporal features are the most commonly used features in gait analysis and have been widely researched and

tested over the years [127]. Examples of spatiotemporal features include stride length, step width, stance time, swing time, single support time, double support time, step count, step duration, heel strike time, toe strike time, heel-off time, and toe-off time. For instance, Bg et al. [131] used the ProtoKinetics Movement Analysis Software (PKMAS) to extract gait features from the Zenomat and GAITRite systems. They extracted features such as stride time and single support time. For each trial, the mean, standard deviation (SD), and asymmetry of eight gait features, including stride time, step time, single support time, swing time, double support time, stance time, stride length, and step length, were extracted.

Kinematic analysis of gait involves studying joint angle deviations, typically calculated using wearable IMU sensors placed on the upper and lower body or motion capture systems (MCS) that allow for three-dimensional motion analysis. From a gait analysis perspective, the most significant kinematic features are the angle values at the ankle, knee, hip, and trunk, which represent the actual changes in joint function when they undergo maximum flexion/extension. For example, Silvia et al. [132] utilized accelerometer-based monitors to measure kinematic gait features. Another study [128] calculated lower limb joint deviations kinematically by analyzing joint deviations in the sagittal plane at the thigh, knee, and ankle joints to obtain the range of motion.

Kinetic features involve the extraction of joint moments and powers, and data are often evaluated using force and pressure sensors embedded in platforms, instrumented walkways, shoes, or insoles. One study [133] extracted features from the vertical ground reaction force (VGRF) signal to describe the amplitude distribution of foot forces throughout the complete gait cycle, characterizing abnormal tremor movements in neurodegenerative diseases.

V. SOUND

In addition to traditional modalities such as neuroimaging, biomarkers, and posture, sound has emerged as a novel non-invasive biological signal that is widely studied for early diagnosis and monitoring of Alzheimer's disease (AD). Sound, as a medium for information transmission, offers unique advantages and relevance in AD research. It is easy to collect, cost-effective, and can be captured using common audio devices such as smartphones and microphones, without the need for specialized equipment or expensive imaging instruments. This makes sound a feasible tool for large-scale screening, with potential applications in various settings including communities, clinical settings, and homes.

A. Characteristics of Language Impairment in AD

In Alzheimer's disease (AD), the language networks and their underlying connectivity fibers in the cerebral cortex and subcortical regions are extensively damaged, resulting in language impairments. The severity of language impairments is positively correlated with the severity of dementia [134]. In the early stages of AD, individuals often experience noticeable difficulties in word-finding and a reduction in vocabulary. Their speech becomes meaningless due to a lack of substantive

words, and they may compensate by using excessive explanations to compensate for their inability to express certain words. As the disease progresses, the difficulties in word-finding become more pronounced [135]. In the middle to late stages, individuals with AD exhibit reduced spontaneous speech, a decrease in vocabulary, and an inability to name common objects or recall the names of family members. In the advanced stages of Alzheimer's disease, language impairments manifest as output disorders, with individuals often giving unrelated responses, ultimately leading to an inability to communicate effectively.

B. Criteria for Assessing AD through Speech

Various language tests have been utilized to identify and quantify cognitive impairments in individuals with Alzheimer's disease (AD) [136]. Samrah et al. [137] analyzed the speech of individuals with AD across four dimensions: speech production, syntactic complexity, lexical content, and fluency errors. M.F. et al. [138] developed an assessment tool consisting of 37 subscales to measure spontaneous speech, auditory comprehension, repetition, naming, reading comprehension, and other abilities in individuals with AD. Each subscale is assigned a score ranging from 0 (normal) to 6 (most abnormal). Ulla et al. [139] categorized tasks into three main types: spontaneous speech (SS), verbal fluency tasks (VF), and other tasks (OT).

The purpose of the SS task is to elicit participants' spontaneous speech. This is typically achieved by asking participants to describe pictures or engage in conversations. The spontaneous speech task can be used to analyze various linguistic attributes such as word retrieval processes, syntax, semantic and acoustic impairments, and communication errors.

Verbal fluency tasks (VF) consist of two types: phonemic verbal fluency (PVF) and semantic verbal fluency (SVF). In PVF tasks, participants are asked to generate as many words as possible within two minutes that begin with a specific letter, such as words starting with the letter "F." In SVF tasks, participants are instructed to produce as many words as possible within one minute belonging to a specific semantic category, such as animal names. Recently, natural language processing (NLP) techniques have been employed for automated analysis of semantic clusters, while speech processing techniques are utilized to analyze timing and acoustic measures.

Other tasks (OT) include tests unrelated to spontaneous speech and semantic clusters, such as sentence repetition, paragraph reading, story writing, counting backward, pronunciation, or naming tests. These tasks allow examination of different aspects of memory, semantic processing, as well as acoustic and speech measurements.

C. Extraction of Speaker-specific Data and Dataset Creation

VBSD Dataset [40]: This dataset collects voice data from 36 participants, including 254 voice samples from individuals with AD and 250 voice samples from healthy controls (HC). A total of 504 spectrogram features were extracted from these voice data. The purpose of this dataset

is to identify common features among multiple AD patients (<https://github.com/LinLLiu/AD>).

Dem@Care Dataset [140]: This dataset consists of voice data from 24 AD patients and 8 healthy controls. The dataset covers four tasks, including picture description, description of pictures from memory, sentence repetition, and rapid pronunciation repetition. The aim of this dataset is to study the performance of AD patients' speech across different tasks.

Pitt Dataset [141]: The Pitt Corpus collects voice data from participants in Alzheimer's and related dementia research at the University of Pittsburgh School of Medicine. The dataset includes description tasks, word fluency tasks, story recall tasks, and sentence construction tasks. The dataset aims to explore the characteristics of voice data across different tasks.

ADReSS Dataset [142]: This dataset was created for the ADReSS challenge, where participants were matched for age and gender to minimize bias risk in the prediction task. The dataset includes speech recordings and textual oral picture descriptions from participants based on the Cookie Theft picture from the Boston Diagnostic Aphasia Examination.

D. Extraction and Introduction of Voice Features

As the disease progresses, AD patients exhibit semantic, syntactic, and phonetic abnormalities compared to normal individuals. Therefore, traditional machine learning methods often utilize lexical and syntactic structures as classification features. For instance, individuals with AD may experience word-finding difficulties, leading to temporal variations in their speech, such as noticeable pauses and slow tempo. During oral reading, speakers with AD dementia demonstrate decreased speech rate, reduced articulation rate, decreased phonation time, as well as increased number and proportion of pauses. In addition to simple syntactic analysis, researchers have incorporated other analyses, such as additional prosodic analysis on the Pitt dataset, extracting 42 Mel Frequency Cepstral Coefficients (MFCC) features [143].

Chen et al. introduced a novel feature called Multi-Resolution Cochleagram (MRCG), which combines cochleograms of four different temporal resolutions to capture local and contextual information. It is capable of extracting acoustic features from mixtures even under very low signal-to-noise ratio (SNR) conditions [144]. Furthermore, researchers have diagnosed Alzheimer's disease by extracting spectrogram-based voice features, including speech rate, number of pauses, and response time. Objective acoustic measurements reveal a significant reduction in pitch modulation in the Alzheimer's group, with individuals with AD often exhibiting a higher proportion of speech interruptions, a decreased ratio of noise to harmonic content, and a general decrease in speech intensity and energy. Their speech sounds highly dysarthric.

For example, Liu et al. [40] recorded participants' voices using wearable Internet devices and obtained a series of speech waveforms. After collecting the data, spectrogram features were extracted from the audio segments (Figure 5), and the resulting spectrogram data were used to train a model. Finally, this trained model was used to identify whether a new set of data belonged to AD. Florian et al. proposed

TABLE III
VOICE DATASETS

Dataset Name	Description
VBSD Dataset [40]	Contains voice data from AD patients and healthy controls, extracting spectral features.
Dem@Care Dataset [140]	Includes voice data from AD patients and healthy controls across multiple tasks.
Pitt Dataset [141]	Collects voice data from participants in Alzheimer's and related dementia studies.
ADReSS Dataset [142]	Includes voice recordings and descriptions based on pictures.

the openSMILE feature extraction toolkit [38], which can extract acoustic features from speech segments. This toolkit combines feature extraction algorithms from speech processing and can extract various features from speech. In addition to mainstream rhythmic features, Jesús B et al. [145] extracted paralinguistic features, which are related to pitch and spectral energy balance. These features place appropriate linguistic information within the context.

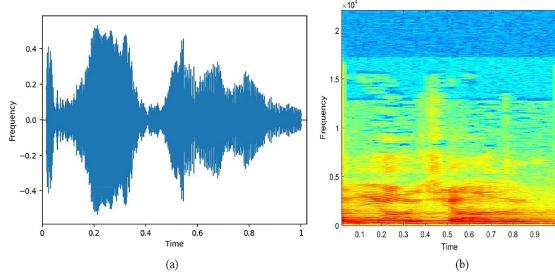


Fig. 5. The spectrum of AD patients' speech, the darker the color, the stronger the speech energy

VI. RESEARCH METHODS

The research methods for early diagnosis of AD using computer-assisted techniques mainly include traditional machine learning methods and deep learning methods. Traditional machine learning methods usually use a series of manually engineered features to classify patient data through classification algorithms. Deep learning methods construct deep neural network models to automatically learn features from raw data and classify them.

Although traditional machine learning methods have achieved some success in the early diagnosis of AD, deep learning methods perform better in processing high-dimensional data and have achieved many important results in other medical fields. Therefore, more and more researchers are applying deep learning methods to the early diagnosis of AD and have achieved some very promising results in this field.

This article will describe research on early classification of AD based on computer-assisted diagnosis from both traditional machine learning and deep learning perspectives.

A. Traditional machine learning methods

1) Feature Extraction: Feature extraction is a common application in pattern recognition and image processing in machine learning. Its goal is to remove the redundant parts caused

by too many features in the original data, in order to save memory space and improve computational power, improve classification performance, and belongs to the dimensionality reduction process.

Feature extraction focuses on the size, shape, and volume of the extracted features for classification research. Early AD diagnosis research mainly involves the hippocampus and the structure of the entorhinal cortex. The common feature extraction objects mainly have the following forms:

- **Volume-based features:** Previous studies have shown that the hippocampus has atrophied in the early stages of AD, and the research shows that about 50% of AD patients have hippocampal atrophy. Therefore, the hippocampal volume can be used as a biomarker feature for AD diagnosis.
- **Thickness-based features:** Lerch et al. found that AD patients have significant differences in the thickness of the entire brain cortex compared to the normal control group NC, reflecting a decrease in the thickness of the cortex in multiple parts of the brain. Clinical studies have found that changes in brain volume caused by brain atrophy are an important biomarker in the early stages of AD.
- **Morphological features:** Changes in brain volume caused by hippocampal and medial temporal lobe atrophy have been used as biomarker features for early detection of AD. Gerardin et al. found in their research that morphological changes were already present in the brains of early-stage AD patients before hippocampal atrophy occurred. Using morphological features as biomarker features for early diagnosis of AD has more sensitive characteristics.
- **Texture feature analysis:** Texture is a visual feature that reflects the homogeneous phenomenon of an image or pattern, reflecting the characterizing attributes of the tissue structure that presents slow or periodic changes on the surface of an object. Freeborough et al. found significant differences between AD patients and normal control group (NC) through texture analysis.
- **Voxel Morphology Analysis:** Voxel (Volume Element) is the minimum unit of digital data in three-dimensional space segmentation, relative to the minimum unit of two-dimensional space, pixel. Voxel-Based Morphometry (VBM) is a comprehensive, objective and automatic analysis method based on voxel units for MRI medical images, which can be used for morphological research on living human brain tissue. By quantitatively calculating the relative changes in gray and white matter density and

volume corresponding to different voxels in MRI images through VBM, it can be used for early diagnosis research of AD. Karas et al. found that there are morphological differences in GM (Gray Matter) voxels between AD patients and normal control group through VBM technology.

2) Classification Algorithm Based: The traditional machine learning methods used for early AD diagnosis and classification research have become increasingly systematic and mature. The algorithms for AD early classification using machine learning assistance mainly include the following categories:

- **Support Vector Machine (SVM):** In the AD early diagnosis and classification research, SVM is the most frequently used algorithm for AD classification. Kim et al. [146] used Boundary Equilibrium Generative Adversarial Networks (BEGAN) to extract features of Alzheimer's disease (AD) and normal cognition (NC) states for stable convergence. Then, the authors trained a support vector machine (SVM) classifier on these features to distinguish AD from NC.
- **Logistic Regression:** Logistic regression can achieve good classification results when learning a large number of training samples. Ieracitano et al. [147] used the logistic regression algorithm based on the brainwave (EEG) CWT and BiS features to classify patients with mild cognitive impairment (MCI) or Alzheimer's disease (AD) in the neurological system.
- **Linear Discriminant Analysis (LDA):** The LDA algorithm is commonly used to evaluate the severity of patients' condition and predict the disease, and can achieve better classification results than logistic regression when the sample data is small. Ahmad et al. [148] used biomarker patterns and resting-state functional magnetic resonance imaging (fMRI) to label the hippocampus (HP), middle temporal gyrus (MTG), olfactory cortex, and posterior cingulate cortex (PCC), and then used LDA for AD classification research.
- **Bayesian Classifier:** The Bayesian classifier has a smaller classification error rate and is commonly used in pattern recognition and other fields. Yubraj et al. [149] proposed using cortical thickness and subcortical volume as biological markers for AD classification, using principal component analysis (PCA) for dimensionality reduction, and using Bayesian classifier for classification research.
- **Random Forest:** Random Forest has higher accuracy than decision tree algorithms and can handle more input data. Dimitriadis et al. [150] proposed using the random forest algorithm to select multiple biomarkers from a subset of feature sets based on preprocessed MRI images, and fusing these biological markers to classify using the random forest algorithm.
- **K-Nearest Neighbor (KNN):** The classification of the KNN algorithm is based on whether the K nearest samples in the feature space of the sample data point belong to the same category. Farouk et al. [151] extracted biological marker information from structural magnetic

resonance images (SMRI), fused texture features based on gray-level co-occurrence matrix and voxel-based morphological features, and used the KNN algorithm for early AD diagnosis and classification.

B. Deep Learning Methods

Deep learning is a machine learning technique based on neural networks, which has achieved remarkable results in many fields. In the diagnosis of Alzheimer's disease, deep learning technology has begun to play an increasingly important role. By analyzing a large amount of medical data of patients, such as neuroimaging, voiceprints, posture, and biomarkers, deep learning technology can improve the accuracy and precision of Alzheimer's disease diagnosis.

1) Neural Network Classification Models:

- **LetNet-5:** Sarraf et al. [152] conducted classification studies of AD and NC based on MRI images using the LeNet model. LetNet-5 can obtain effective representations of original images, and can identify the rules of original images with very few preprocessing, but due to the characteristics of its model, the results are not ideal for processing large-scale data and complex problems.
- **AlexNet:** Afzal et al. [153] used the pre-trained AlexNet network model to select 199 subjects from the Open Access Series of Imaging Studies (OASIS) dataset for classification studies of AD and NC. The classification accuracy reached 98.41%.
- **VGG:** Jain et al. [154] used the VGG-16 network model and MRI single-mode imaging to conduct a three-classification study of NC, MCI, and AD. Experimental data came from the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset, and 50 samples were selected for the experiment. The classification accuracy can reach 99.14%.
- **GoogLeNet:** Sarraf et al. [152] respectively obtained 144 fMRI and 302 MRI sample data from the ADNI dataset based on MRI and fMRI modal imaging data, and used the GoogLeNet network model to extract low to high features from a large number of training images. The experimental training set and test set ratio was 3:1. The average accuracy of GoogLeNet model in AD and NC classification under fMRI images was 94.24%. The average accuracy of GoogLeNet and LeNet models based on MRI images were 98.74% and 97.88%, respectively. GoogLeNet model performed better.
- **ResNet:** Ji et al. [155] conducted AD classification studies based on ResNet network model in single-mode MRI, multi-mode MRI, and FDG-PET images. The classification accuracy reached 97.65% in AD/mild cognitive impairment and 88.37% in mild cognitive impairment/normal control.
- **3D-CNN:** Kompanek et al. [156] used 3D-CNN based on MRI images to classify NC and AD, and used Volumetric Data Augmentation (VDA) to enhance the classification accuracy and generalization ability of the network model.

2) Based on unimodal deep learning methods: Modality refers to the form in which things happen or exist, which can

be information such as sound, images, and text. For the early diagnosis of AD research, it can be data such as brain imaging, voiceprints, and postures.

The application of deep learning methods to biomedical research is one of the main tasks to use different modal biomarkers for early AD classification research. Barbaroux et al. [157] explored AD classification tasks by analyzing the human brain cortex using the SCNN (Spherical CNN) model. SCNN extends the regular CNN from planar data to spherical data by defining convolutional operators that are equidistant to the same 3D rotation on spherical input. The experiment is based on the morphological measurement method of T1-weighted MRI images. The classification accuracy between AD and NC can reach 92.16%, indicating the feasibility and superiority of using the SCNN model directly for discriminating analysis of human brain cortex.

Kim et al. [146] used an improved deep learning method with a GAP layer based on FDG-PET imaging data, and conducted experiments based on the ADNI and Severance dataset imaging data from the nuclear medicine department of the Westfalen Hospital. They achieved accuracy rates of 91.02% and 86.09% in the classification results of AD and NC, respectively. The experiment used slice-selective learning to reduce computation, and at the same time, used transfer learning and a GAP layer instead of a fully connected layer to avoid overfitting and improve the model's generalization ability. The experimental results on different datasets showed that the performance of the GAP layer was statistically significant compared with the fully connected layer in terms of model accuracy, sensitivity, and specificity ($p < 0.01$), and the model's performance on accuracy, sensitivity, and specificity did not show statistically significant differences on different datasets ($p > 0.05$).

Bin et al. [158] proposed using multiple deep 2D-CNNs constructed by separable convolution layers to learn various features of brain regions. Based on the sMRI imaging in the OASIS cross-sectional image dataset for AD classification research, the Xception and Inception version 3 structures were used as transfer learning models. The experimental results showed that the model had superior performance in AD classification.

Suganthe et al. [159] proposed to establish a deep convolutional neural network (DCNN) and a VGG-16-based convolutional neural network (VCNN) model, using MRI images from the ADNI database for classification research on 4 stages: AD, NC, EMCI, and LMCI. By changing the number of convolutional layers, convolutional kernel sizes and numbers, dropout layers, pooling layers, and other parameters to compare the model classification results, the results showed that the DCNN model with 7 convolutional layers had much better classification performance than the DCNN model with 4 convolutional layers. The classification accuracy of AD and LMCI reached 93.76%. The comparative experiments of the 2 models showed that the accuracy of the VCN model was higher than that of the DCNN model, but its number of trainable parameters was higher than that of the DCNN model.

Böhle et al. [160] used layer-wise relevance propagation (LRP) to visualize the decisions of an MRI-based Alzheimer's

disease (AD) convolutional neural network. LRP produces a heatmap in the input space indicating the relevance/importance of each voxel to the final classification result, which can help clinicians interpret neural network decisions for diagnosis of AD (and other diseases) based on structural MRI data.

Zhu et al. [79] proposed a dual-attention multiple instance deep learning network (DA-MIDL) for early diagnosis of AD and mild cognitive impairment (MCI), its prodromal stage. DA-MIDL consists of three main components: 1) Patch-Nets based on spatial attention blocks to extract discriminative features within each sMRI patch, while enhancing abnormal microstructural features of the brain, 2) attention-based multiple instance (MIL) pooling to balance relative contributions of each patch and generate global differentially-weighted representations for the entire brain structure, and 3) attention-based global classifier to further learn holistic features and make AD-related classification decisions. The model was evaluated on 1689 subjects from the ADNI and AIBL datasets, achieving state-of-the-art results.

Liu et al. [10] proposed a deep learning embedded AD diagnosis framework with less prior clinical knowledge to differentiate four stages of AD progression. Due to unsupervised feature representation embedded in the workflow, it has the potential to be extended to more unlabeled data for feature engineering in practice. In the unsupervised pretraining stage, they used SAEs to obtain high-level features. When using multiple neural imaging modalities, they employed a zero-mask strategy to extract synergistic effects between different modalities, processed in a denoising manner. After unsupervised feature engineering, the authors used a softmax regression model. They also adopted a novel method to visualize high-level brain biomarkers for analyzing extracted high-level features. The proposed framework was evaluated in AD stages 2-4 classification. Based on MR and PET ADNI datasets, the authors' framework outperformed state-of-the-art SVM-based methods and other deep learning frameworks.

Lian et al. [73] proposed a hierarchical fully convolutional network (H-FCN) for automated identification of discriminative local regions in the entire brain sMRI. Then, joint learning and fusion of multiscale feature representations were conducted to construct a hierarchical classification model for AD diagnosis. Their proposed H-FCN method was evaluated in a large number of subjects from two independent datasets, i.e., ADNI-1 and ADNI-2, demonstrating good performance in joint discriminative atrophy localization and brain disease diagnosis.

Ebrahimi [68] used a ResNet-18 pretrained on the ImageNet dataset and employed serialized models, time convolutional networks (TCN), and various types of recurrent neural networks. Their proposed TCN model obtained the best classification performance using MRI, with an accuracy of 91.78%, sensitivity of 91.56%, and specificity of 92%. Their results indicate that adopting sequence-based models can improve classification accuracy by 10% compared to 2D and 3D CNNs.

Westman et al. [161] used orthogonal partial least squares (OPLS) to differentiate Alzheimer's disease (AD), mild cognitive impairment (MCI), and elderly control groups, using local and global magnetic resonance imaging (MRI) volume

measurements. The study included 117 AD patients, 122 MCI patients and 112 control subjects (from the AddNeuroMed study). High-resolution sagittal 3D MP-RAGE data sets were obtained from each subject, with automatic region segmentation and manual contouring of the hippocampus for each image. A total of 24 different anatomical structures were used for sample MVP analysis. Seventeen randomly selected AD patients, 12 randomly selected control subjects and 22 MCI patients who converted to AD in the 1-year follow-up were excluded from the initial OPLS analysis to provide a small external test set for model validation. Using only hippocampal measurements compared to controls, the authors found a sensitivity of 87% and a specificity of 90% for AD. Combining global and regional measurements resulted in a sensitivity of 90% and a specificity of 94%. This increase in sensitivity and specificity led to a positive likelihood ratio from 9 to 15. From the external test set, the model correctly predicted 82% of AD patients and 83% of control subjects. Finally, 73% of MCI subjects who converted to AD in 1 year of follow-up were found to be closer to AD patients than control subjects.

Morra et al. [162] introduced a new brain MRI segmentation method called automatic context model (ACM) for automatic segmentation of the hippocampus in images of elderly dementia tissues. During the training stage, the algorithm used 21 manually labeled segmentations to learn the classification of hippocampal and non-hippocampal regions using an improved AdaBoost method based on approximately 18,000 features, including image intensity, location, image curvature, image gradient, tissue classification map of gray/white matter and CSF, mean, standard deviation, and Haar filters of sizes ranging from 1x1x1 to 7x7x7. All brain images were linearly registered to a standard template, and a basic shape prior was designed to capture the global shape of the hippocampus, defined as the point-wise sum of all training masks. They also incorporated shape priors and features of curvature, gradient, mean, standard deviation, and Haar filters of the tissue classification image as features. In each iteration of the ACM, the Bayesian posterior distribution and the surrounding features were fed to AdaBoost as new features. In the validation study, the results were compared with the manual annotations of two experts. Using the leave-one-out method and standard overlap and distance error metrics, the automated segmentation results were highly consistent with those of human evaluators, with any differences comparable to those between trained human evaluators.

In another study [85], a groundbreaking deep belief network (DBN)-based multi-task learning algorithm is proposed for solving classification issues. Specifically, the overfitting problem is addressed by utilizing the dropout technology and zero-masking strategy. These techniques also improve the model's generalization ability and robustness. Subsequently, a novel framework based on DBN-based multi-task learning is constructed for precise diagnosis of AD. After MRI preprocessing, principal component analysis reduces the feature dimension, and multi-task feature selection approach selects the feature set related to all tasks based on the internal relevancy among multiple related tasks. Using data from the ADNI dataset, their

method achieved satisfactory results in six tasks: health control (HC) vs. AD, HC vs. pMCI, HC vs. sMCI, pMCI vs. AD, sMCI vs. AD, and sMCI vs. pMCI, with accuracies of 98.62%, 96.67%, 92.31%, 91.89%, 99.62%, and 87.78%, respectively. Experimental findings demonstrate that the DBN-based MTL algorithm developed in this study is a remarkably efficient, superior, and practical method for AD diagnosis.

Wang et al. [67] firstly propose a VIN, a VGG-inspired network as the backbone, and investigate the utilization of attention mechanisms. They introduce the Alzheimer's Disease VGG-Inspired Attention Network (ADVIAN), where convolutional block attention modules are integrated into the VIN backbone. Additionally, they propose 18-way data augmentation to avoid overfitting. They perform ten runs of 10-fold cross-validation to report unbiased performance. Their results show that sensitivity and specificity reach 97.65 +/- 1.36 and 97.86 +/- 1.55, respectively. ADVIAN achieves a precision and accuracy of 97.87 +/- 1.53 and 97.76 +/- 1.13, respectively. Furthermore, the F1 score, MCC, and FMI obtain 97.75 +/- 1.13, 95.53 +/- 2.27, and 97.76 +/- 1.13, respectively. The AUC is 0.9852. In conclusion, they proposed ADVIAN outperforms 11 state-of-the-art methods. Additionally, their experimental results demonstrate the effectiveness of 18-way data augmentation.

In another study, Liu et al. [163] utilize a deep-learning framework that employs Siamese neural networks, which are trained on paired lateral inter-hemispheric regions. The approach aims to harness the discriminative power of whole-brain volumetric asymmetry. To achieve this objective, they use the MRICloud pipeline to obtain volumetric features of pre-defined atlas brain structures, with a novel non-linear kernel trick applied to normalize these features. This normalization process works to reduce batch effects across various datasets and populations. They work with the low-dimensional features of the brain and show that Siamese networks achieve comparable performance to studies that use whole-brain MR images. This is an advantage, as they reduce complexity and computational time while still preserving the biological information density. Their experimental results demonstrate that Siamese networks outperform traditional prediction methods that do not incorporate the asymmetry in brain volumes on the ADNI and BIOCARD datasets in certain metrics.

In another paper, Lee et al. [164] propose a new method for diagnosing Alzheimer's disease (AD) or mild cognitive impairment (MCI) using magnetic resonance imaging (MRI). Their approach integrates voxel-based, region-based, and patch-based methods into a single framework. They parcellate the brain into predefined anatomical regions using templates and derive complex nonlinear relationships among the voxel intensities, which provide volumetric measurements, within each region. Unlike other approaches that use cubical or rectangular shapes, they focus on the unique, atypical shapes of the regions. Using deep neural networks, they learn the complex nonlinear relationships among the voxels in each region and extract a "regional abnormality representation." Finally, they make a clinical decision by integrating the regional abnormality representations across the whole brain. Their method has an advantage in allowing us to interpret and

visualize symptomatic observations of subjects with AD or MCI in the brain space. On the baseline MRI dataset from the Alzheimer's Disease Neuroimaging Initiative (ADNI) cohort, their method achieves state-of-the-art performance for four binary classification tasks and one three-class classification task.

Cui et al. [165] present a classification framework that combines convolutional and recurrent neural networks for longitudinal analysis of structural MR images in AD diagnosis. First, they construct a Convolutional Neural Network (CNN) to learn the spatial features of MR images for the classification task. Then, they build a recurrent Neural Network (RNN) with three cascaded bidirectional gated recurrent unit (BGRU) layers on the outputs of CNN at multiple time points to extract longitudinal features for AD classification. Their proposed method jointly learns the spatial and longitudinal features along with the disease classifier, achieving optimal performance. Furthermore, their approach models longitudinal analysis using RNN from the imaging data at various time points. They evaluate the proposed method with longitudinal T1-weighted MR images of 830 participants, including 198 AD, 403 mild cognitive impairment (MCI), and 229 normal control (NC) subjects from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. Experimental results show the proposed method achieves a classification accuracy of 91.33% for AD vs. NC and 71.71% for pMCI vs. sMCI. These results demonstrate the promising performance of their method for longitudinal MR image analysis.

In another paper, Ge et al. [166] propose a deep learning method for multi-scale feature learning that is based on segmented tissue areas. They propose a novel deep 3D multi-scale convolutional network scheme that generates multi-resolution features for AD detection. The proposed scheme employs several parallel 3D multi-scale convolutional networks, each of which applies to individual tissue regions (GM, WM and CSF), followed by feature fusions. The proposed fusion is applied in two separate levels: the first level fusion is applied on different scales within the same tissue region, and the second level is applied on different tissue regions. To further reduce the dimensions of features and mitigate overfitting, a feature boosting and dimension reduction method, XGBoost, is utilized before classification. The proposed deep learning scheme has been tested on a moderate open dataset of ADNI (consisting of 1198 scans from 337 subjects) and has shown excellent performance on randomly partitioned datasets (with a top score of 99.67% and an average score of 98.29%) and good performance on subject-separated partitioned datasets (with a top score of 94.74% and an average score of 89.51%).

Cui et al. [167] propose a new method for analyzing the hippocampus that combines global and local features using three-dimensional densely connected convolutional networks and shape analysis for diagnosing Alzheimer's Disease. Their method effectively incorporates local visual features and global shape characteristics to improve classification without requiring tissue segmentation or nonlinear registration. They evaluated their method using T1-weighted structural MRIs from 811 subjects, which included 192 AD patients, 396 MCI patients (of which 231 were stable MCI and 165 were

progressive MCI), and 223 normal control subjects from the Alzheimer's disease neuroimaging initiative database. Their experimental results demonstrate that their proposed method achieved a classification accuracy of 92.29% and area under the ROC curve of 96.95% for AD diagnosis.

The methodology proposed by Ieracitano et al. [168] involves evaluating the power spectral density (PSD) of EEG traces from 19 channels and using this information to generate 2D grayscale images (PSD-images) representing the spectral profiles. A customized Convolutional Neural Network, which includes a processing module comprising of convolution, Rectified Linear Units (ReLU), and pooling layer (CNN1), is designed to extract suitable features from the PSD-images, thereby facilitating two and three-way classification tasks. Compared to more conventional learning machines, the resulting CNN provides better classification performance, achieving an average accuracy of 89.8% for binary classification and 83.3% for three-way classification.

In another paper, Li et al. [169] present a 4D deep learning model, called C3d-LSTM, for AD discrimination. The model is capable of utilizing spatial and time-varying information simultaneously by processing 4D fMRI data directly. The proposed C3d-LSTM combines 3D convolutional neural networks (CNNs) to extract spatial features from each volume of the 3D static image in an fMRI image sequence. Subsequently, the feature maps obtained are fed to the long short-term memory (LSTM) network to capture the time-varying information contained within the data. Because of its design, the C3d-LSTM is an end-to-end data-driven model, making it more convenient for processing 4D fMRI data. The proposed model was evaluated on the AD Neuroimaging Initiative dataset against controlled experiments. This evaluation produced better results in AD detection than the functional connectivity, 2D, or 3D fMRI data methods when utilizing 4D fMRI data directly. The results demonstrate that the proposed method using the C3d-LSTM model provided a far better result in AD detection.

In another study, Chitradevi et al. [170] considered various brain sub-regions to diagnose Alzheimer's disease using four optimization algorithms: Genetic Algorithm (GA), Particle Swarm Optimization algorithm (PSO), Grey Wolf Optimization (GWO), and Cuckoo Search (CS). Among these optimization techniques, GWO produced promising results due to its proper selection of global optimal solution. The segmented regions were classified using a deep learning classifier and validated with Ground Truth (GT) images. The results demonstrate that GWO can segment brain sub-regions with a high accuracy of 98% similarity between the ground truth and segmented regions. The segmented regions were then classified using a deep learning classifier, providing a high accuracy of 95%. The HC sub-region showed better classification accuracy (95%), sensitivity (95%), and specificity (94%) compared to all other regions. Based on the segmentation and classification measures, it is clear that the proposed method performed better than other methods. Finally, clinical validation was performed on normal and AD subjects using mini-mental state examination (MMSE) scores. The proposed pipeline correlated strongly with MMSE scores. Therefore, the results suggest that the HC region is a major factor in diagnosing AD.

The aim of Puente et al. [171]’s study is to develop an automated system for detecting disease presence in sagittal MRI scans, which are underutilized. They utilized sagittal MRIs from ADNI and OASIS datasets and employed Transfer Learning techniques for better accuracy. Their findings indicate that damages related to AD can be identified in sagittal MRI scans and that the results are similar to those from state-of-the-art DL models that use horizontal-plane MRI scans.

In another paper, Chen et al. [172] propose an iterative model for sparse and deep learning (ISDL) that jointly extracts deep features and identifies critical cortical regions for diagnosing AD and MCI. Their first step is to design a Deep Feature Extraction (DFE) module that captures structural information ranging from local to global that is derived from 62 cortical regions. They then incorporate a sparse regression module to identify the critical cortical regions and integrate it into the DFE module. This step helps to exclude irrelevant cortical regions from the diagnostic process. They update the parameters of the two modules alternatively and iteratively in an end-to-end manner. Their experimental results suggest that the ISDL model provides a state-of-the-art solution for both AD-CN classification and MCI-to-AD prediction.

Hazarika et al. [173] believe that, amongst all the models presented, the DenseNet-121 model achieves a convincing result with an average performance rate of 88.78%. However, one of the downsides of the DenseNet model is that it utilizes a lot of convolutional operations, which make the model computationally slower than several other models discussed. Depth-wise convolution is a popular method to make convolutional operation faster and better. Therefore, to enhance the execution time, they propose replacing the convolution layers in the original DenseNet-121 architecture with depth-wise convolution layers. The new architecture not only improved the execution time, but it also improved the performance of the model, with an average rate of 90.22%.

Bae et al. [174] have developed a deep learning model for predicting conversion from MCI (Mild Cognitive Impairment) to DAT (Dementia of Alzheimer’s Type) using structural magnetic resonance imaging scans as input to a 3-dimensional convolutional neural network. This neural network was trained using transfer learning, with normal control and DAT scans used as input in the source task to pretrain the model. Later, they retrained the model on a target task of classifying which MCI patients eventually progressed to DAT. Their model had an outstanding classification accuracy of 82.4% at the target task, surpassing current models in the field. They also employed an occlusion map approach to visualize the brain regions that significantly contribute to the prediction. Surprisingly, they found that among the contributory regions were the pons, amygdala, and hippocampus.

In another study, Zhang et al. [175] propose a novel computer-aided approach for early diagnosis of AD. They introduce an explainable 3D Residual Attention Deep Neural Network (3D ResAttNet) for end-to-end learning from sMRI scans. Their approach is distinct from existing methods in three ways: Firstly, they propose a Residual Self-Attention Deep Neural Network to capture local, global, and spatial information of MR images and improve diagnostic perfor-

mance. Secondly, they introduce an explainable method using Gradient-based Localization Class Activation mapping (Grad-CAM) to improve the interpretability of the proposed method. Thirdly, their work provides a complete end-to-end learning solution for automated disease diagnosis. They evaluate the proposed 3D ResAttNet method on a large cohort of subjects from real datasets for two challenging classification tasks, namely Alzheimer’s disease (AD) vs. Normal cohort (NC) and progressive MCI (pMCI) vs. stable MCI (sMCI). The experimental results demonstrate that their approach has a competitive advantage over state-of-the-art models in terms of accuracy performance and generalizability.

In another paper, Huang et al. [176] propose the use of a convolutional neural network (CNN) to integrate all the multi-modality information contained in both T1-MR and FDG-PET images of the hippocampal area for the diagnosis of Alzheimer’s disease (AD). Unlike traditional machine learning algorithms, their method does not require manually extracted features - instead, it utilizes 3D image-processing CNNs to learn features for the diagnosis or prognosis of AD. To evaluate the performance of the proposed network, they trained the classifier with paired T1-MR and FDG-PET images in the ADNI dataset. The dataset includes 731 cognitively unimpaired subjects labeled as CN, 647 subjects with AD, 441 subjects with stable mild cognitive impairment (sMCI), and 326 subjects with progressive mild cognitive impairment (pMCI). They obtained higher accuracies of 90.10% for the CN vs. AD task, 87.46% for the CN vs. pMCI task, and 76.90% for the sMCI vs. pMCI task. The proposed framework yielded a state-of-the-art performance.

Ghoraani et al. [131] collected data on both “single-tasking” gait (walking) and “dual-tasking” gait (walking with cognitive tasks) from 32 healthy participants, 26 participants with mild cognitive impairment (MCI), and 20 participants with Alzheimer’s disease (AD) using a computerized walkway. Each participant was assessed with the Montreal Cognitive Assessment (MoCA), and a set of gait features (such as mean, variance, and asymmetry) was extracted. Significant gait features for three classifications (MCI/healthy, AD/healthy, and AD/MCI) were identified, and a support vector machine model was trained in a one-vs.-one manner for each classification. The majority vote of the three models was assigned as either healthy, MCI, or AD. Regarding the results, the average accuracy of the classification was 78% (F1-score of 77%), which was reasonable when compared to the MoCA score with an accuracy of 83% (F1-score of 84%). Additionally, the performance of healthy vs. MCI or AD was 86% (F1-score of 88%), which was also comparable to MoCA’s accuracy of 88% (F1-score of 90%).

Rucco et al. [128]’s study explored the gait patterns in Alzheimer’s disease (AD) and behavioral variant of Frontotemporal Dementia (bvFTD) patients. The bvFTD group showed more instability and slowness in a single task, and worsened further in a dual motor task. AD patients had lower speed and stride length than healthy subjects. During a cognitive dual task, both groups showed impaired velocity and stability, with AD patients showing a significant deterioration in gait performance.

Emotion processing and social interaction difficulties are present in dementia, particularly in frontotemporal dementia. Corticobasal syndrome is related to frontotemporal dementia, with emotion processing changes found in frontoparietal cortical regions and the basal ganglia, but has not been systematically explored. A study was conducted by Kumfor et al. [122] to examine emotion processing in corticobasal syndrome and Alzheimer's disease. Results showed that emotional processing deficits are present in patients with corticobasal syndrome and are related to changes in brain regions crucial for emotion processing.

These studies demonstrate the potential of unimodal deep learning methods in leveraging specific modalities of data, such as brain imaging, for accurate AD classification. However, these approaches are limited to a single modality, and the fusion of information from multiple modalities can provide more comprehensive insights.

3) Based on Multimodal Deep Learning Methods: Multimodal deep learning refers to the use of multiple single-modal information in deep learning methods to achieve the fusion of information between different modalities.

Many researchers use multimodal data for early classification research of Alzheimer's disease (AD). Huang et al. [176] used various multimodal feature information under the hippocampal TI-weighted magnetic resonance imaging and FDG-PET imaging, using convolutional neural networks for AD classification research. The classification accuracies obtained for NC and PMCI, SMCI and PMCI, AD and NC were 87.46%, 76.90%, and 90.10% respectively. The results showed that the classification results combined with multimodal imaging data were better than those in the single-modal form.

Forouzannzhad et al. [177] used the multimodal imaging technology of positron emission tomography (PET) and magnetic resonance imaging (MRI) and the results of standard neuropsychological testing scores, to classify early diagnosis of AD using deep neural networks (DNN). The classification accuracies for normal control group NC and early mild cognitive impairment EMCI were as high as 84.0%. The classification accuracies for NC and late mild cognitive impairment LMCI, CN and AD, EMCI and LMCI, EMCI and AD, LMCI and AD were 84.1%, 96.8%, 69.5%, 90.3% and 80.2%, respectively. The classification accuracy for NC and EMCI was only 68.0% on MRI alone. The study showed that the multimodal method was better than single-modal imaging analysis.

Kang et al. [178] constructed a transfer learning method of the VGG16 model based on SMRI and diffusion tensor imaging DTI (Diffusion Tensor Imaging) bimodal data for the classification research of EMCI and NC. The data came from the ADNI dataset. The experiment used a multimodal fusion strategy to merge slices with the same index into RGB slices, and input them into the model for training. LASSO (Least Absolute Shrinkage and Selection Operator) algorithm was used to extract the features related to EMCI disease, and the experimental obtained a classification accuracy of 94.2%, with a high sensitivity of 97.3%. The experimental results showed that multimodal data can provide more useful information for

distinguishing EMCI and NC, and from a clinical perspective, DTI is an important biomarker for EMCI.

Khvostikov et al. [179] fused the SMRI and DTI imaging modes in the hippocampal region of interest, and compared the use of single-modal data for the experiment with the AD classification algorithm based on the 3D-CNN model. In the AD and MCI classifications, a high accuracy of 93.3% was obtained under the multimodal situation, which had a significant advantage over the accuracy of 65.8% under the SMRI single-mode. In the experiment, data augmentation was used to balance the sizes of different classes, in order to eliminate the impact of different amounts of data on the network training process.

Lin et al. [180] tackled the problem of expensive missing PET data by first using a reversible generative adversarial network (RevGAN) model to reconstruct missing data. Then, the team proposed a 3D convolutional neural network (CNN) classification model with multimodal inputs for AD diagnosis. Finally, their method was evaluated on the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. The experimental results showed that the proposed framework can effectively map the structure and functional information of brain tissue to images. The synthesized images were very close to real images, and the framework significantly improved the performance of AD diagnosis and MCI conversion prediction.

Martinez-Murcia et al. [70] attempted to explore AD through deep convolutional autoencoders. By fusing the information of neuropsychological test results, diagnosis, and other clinical data with imaging features extracted only through MRI, the study aimed to find the relationship between cognitive symptoms and potential neurodegenerative processes. Then, the extracted features in different combinations were analyzed and visualized, and regression and classification analyses were used to estimate the impact of each coordinate of the autoencoder's manifold on the brain. In terms of AD diagnosis, the study achieved over 80% classification accuracy.

In the study by Shi [181] et al., a multimodal stacked dense pyramid network (MM-SDPN) algorithm was proposed, which consists of two levels of SDPN for feature representation and fusion of multimodal neuroimaging data for AD diagnosis. Specifically, the high-level features of MRI and PET were learned by two SDPNs, and then the features were fed into another SDPN to fuse multimodal neuroimaging information. The proposed MM-SDPN algorithm was applied to the ADNI dataset for binary and multi-class classification tasks. The experimental results showed that MM-SDPN outperformed the state-of-the-art algorithms based on multimodal feature learning for AD diagnosis.

El-Sappagh et al. [182] developed an accurate and interpretable diagnostic and progression detection model for Alzheimer's disease (AD), providing doctors with accurate decisions and a set of explanations for each decision. Specifically, the model integrates 11 modalities from 1048 subjects from the ADNI dataset: 294 cognitively normal, 254 stable mild cognitive impairment (MCI), 232 progressive MCI, and 268 AD. This is actually a two-layer model that uses random forest (RF) as the classifier algorithm. In the first layer, the model performs multi-class classification aimed at early

diagnosis of AD patients. In the second layer, the model applies binary classification to detect MCI to AD progression that may occur within three years after baseline diagnosis. The performance of the model was optimized using key markers selected from a large number of biological and clinical indicators. For interpretability, the authors provide RF classifier global and instance-based explanations for each layer using the SHapley Additive exPlanations (SHAP) feature attribution framework. Additionally, they implemented 22 interpreters based on decision trees and fuzzy rules to provide supplementary explanations for each RF decision in each layer. Moreover, these explanations are presented in natural language to help doctors understand the prediction results. The designed model achieved a cross-validation accuracy of 93.95% and F1 score of 93.94% in the first layer, and a cross-validation accuracy of 87.08% and F1 score of 87.09% in the second layer. Due to the wide consistency of the explanations provided with AD medical literature, the resulting system is not only accurate but also trustworthy, traceable, and medically applicable. The system can help enhance the clinical understanding of AD diagnosis and progression by providing detailed insights into the influence of different modalities on disease risk.

AD is a progressive neurodegenerative disease, and biomarkers based on pathological physiology may provide objective indicators for disease diagnosis and staging. Neuroimaging scans obtained from MRI and metabolic imaging obtained through FDG-PET provide in vivo measurements of brain structure and function (glucose metabolism) in brain life activities. There is currently a hypothesis that combining multiple different image modalities providing complementary information can help improve the ability to diagnose AD early. Lu et al. [9] proposed a novel framework based on deep learning that uses multi-modality and multi-scale deep neural networks to identify individuals with AD. Their method provided 82.4% accuracy in identifying patients with mild cognitive impairment (MCI) who will transition to AD within 3 years (86.4% comprehensive accuracy in transitioning within 1-3 years), 94.23% sensitivity in classifying individuals who may be clinically diagnosed with AD, and 86.3% specificity in classifying non-dementia control groups, which is better than results in previous literature.

Qiu's team [72] reported an interpretable deep learning strategy that uses multi-modal input of MRI, age, gender, and minimum mental state examination scores to describe unique Alzheimer's disease markers. Their framework links a fully convolutional network with a multi-layer perceptron, which can construct high-resolution maps of disease probability from local brain structures to generate accurate, intuitive individual AD risk visualization for accurate diagnosis. The model was trained using patients diagnosed with Alzheimer's disease and cognitively normal subjects from the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset ($n=417$), and validated on three independent cohorts: the Australian Imaging, Biomarkers and Lifestyle representational study (AIBL) ($n=382$), the Framingham Heart Study ($n=102$), and the National Alzheimer's Coordinating Center (NACC) ($n=582$). The performance of the model with multi-modal input remained consistent across datasets, with average area under the curve

values of 0.996, 0.974, 0.876, and 0.954 for the ADNI study, AIBL, Framingham Heart Study, and NACC data sets, respectively. In addition, their method outperformed the diagnostic performance of a multi-institutional team of eleven practicing neurologists, and the high-risk brain areas predicted by the model closely tracked post-mortem pathological findings.

Shangran et al. [183] present a deep learning framework that conducts several diagnostic steps consecutively to identify individuals with normal cognition (NC), mild cognitive impairment (MCI), Alzheimer's disease (AD), and non-AD dementias (nADD). Their framework uses multiple models that can accept flexible combinations of routinely collected clinical information, including demographics, medical history, neuropsychological testing, neuroimaging, and functional assessments. They then demonstrate that their models perform comparably to the diagnostic accuracy of practicing neurologists and neuroradiologists. Additionally, they utilize interpretability methods in computer vision to reveal that disease-specific patterns detected by their models correspond closely with the presence of neuropathological lesions on autopsy and track distinct patterns of degenerative changes throughout the brain. Their work demonstrates methodologies for validating computational predictions against established standards of medical diagnosis.

In another study, Martinez et al. [70] have implemented a novel approach for exploratory data analysis of Alzheimer's Disease (AD) which utilizes deep convolutional autoencoders. Their main objective is to establish connections between cognitive symptoms and the underlying neurodegeneration process by integrating information from neuropsychological test outcomes, clinical diagnoses, and other related data with imaging features extracted using a data-driven decomposition of MRI scans. To accomplish this, they analyze and visualize the distribution of extracted features in various combinations through regression and classification analysis. Furthermore, they estimate the influence of each coordinate of the autoencoder manifold on the brain. Their results show that the imaging-derived markers are able to predict clinical variables with correlations greater than 0.6, particularly in the case of neuropsychological evaluation variables like the MMSE or the ADAS11 scores. Moreover, their approach achieves a classification accuracy above 80 for the diagnosis of AD.

Liu et al. [184] introduce a multi-model deep learning framework based on a convolutional neural network (CNN) for automatic hippocampal segmentation and Alzheimer's disease (AD) classification through structural MRI data. Firstly, they construct a multi-task deep CNN model to learn jointly about hippocampal segmentation and disease classification. Using the segmentation results, they then construct a 3D Densely Connected Convolutional Network (3D DenseNet) to learn the features of 3D patches. Their method utilizes a combination of the features learned from the multi-task CNN and DenseNet models to classify disease status. Their proposed method is evaluated using baseline T1-weighted structural MRI data that they have collected from 97 AD patients, 233 patients with mild cognitive impairment (MCI), and 119 normal control (NC) subjects from the ADNI database. They have achieved excellent results, with a dice similarity coefficient of 87.0%

for hippocampal segmentation. Furthermore, their proposed method has attained an accuracy of 88.9% and an AUC (area under the ROC curve) of 92.5% for classifying AD versus NC subjects, along with an accuracy of 76.2% and an AUC of 77.5% to classify MCI versus NC subjects.

These findings emphasize the advantages of multimodal deep learning methods in leveraging diverse data sources to improve AD classification accuracy. By integrating information from multiple modalities, these approaches provide more comprehensive and robust insights for early AD diagnosis.

VII. EXPERIMENT

In this chapter, we will discuss the relevant researches on the early identification of Alzheimer's disease (AD) patients using various modalities in recent years, focusing on three aspects: 1) Dataset and feature selection, 2) Modality and method, 3) Evaluation index and performance.

A. Dataset and feature selection

The early classification of AD is a complex task that requires specific characteristics in the data samples, including: 1) The dataset needs to have a certain scale; 2) The samples should be labeled by professional doctors; 3) Acquiring a large amount of experimental data and annotations often requires significant human and material resources.

Currently, the most widely used general datasets for AD research include ADNI, OASIS, AIBL, FHS, NACC and BIOCARD. These datasets cover a wide range of data types and can meet the needs of most AD-related studies.

1) ADNI: The Alzheimer's Disease Neuroimaging Initiative (ADNI) is renowned for its longitudinal and multi-center research and is the most common dataset used in AD research. ADNI researchers collect various types of data from patients, including clinical, genetic, MRI, PET images, and biomarkers, with the aim of investigating whether they can measure the progression of Mild Cognitive Impairment (MCI) and early AD. The ADNI series, including ADNI-1, ADNIGO, ADNI-2, and ADNI-3, are supplements and improvements to previous series, as shown in Fig. 6.(Source: <https://adni.loni.usc.edu/>)

2) OASIS: The Open Access Series of Imaging Studies (OASIS) aims to share brain neuroimaging datasets with researchers in related fields. OASIS is available in the following versions: OASIS-1 includes 434 cross-sectional MRI scans from 416 subjects aged 18 to 96 years. OASIS-2 consists of 373 longitudinal MRI scans from 150 subjects. OASIS-3 includes 2,842 longitudinal MRI scans, 2,157 PET scans, and 1,472 CT scans from 1,379 subjects. Each subject typically has two or more scans with an interval of at least one year. Currently, OASIS is the second largest core data resource after ADNI. (Source: <https://www.oasis-brains.org/#about>)

3) AIBL: The Australian Imaging, Biomarker & Lifestyle (AIBL) study was initiated in 2006 to conduct long-term assessments (over 4.5 years) on more than 1,100 patients in order to determine which biomarkers, cognitive features, health, and lifestyle factors contribute to the subsequent development of symptomatic Alzheimer's disease (AD). AIBL collects imaging and medical data from 211 AD patients, 133 patients

with mild cognitive impairment (MCI), and 768 cognitively normal healthy individuals, making it the largest study of its kind in Australia. (Source: <https://aibl.csiro.au/about/>)

4) FHS: The Framingham Heart Study (FHS) [185] is a longitudinal community-based cohort study that has collected extensive clinical data on three generations. Since 1976, the FHS has expanded to include the evaluation of factors leading to cognitive decline, dementia, and Alzheimer's disease. In March 1999, the offspring cohort began undergoing brain MRI scans and neuropsychological (NP) testing. The MRI and NP studies allow for quantitative measurements of brain structure, such as white matter hyperintensities (WMH) volume, and cognitive function, respectively. Combined with its comprehensive clinical data, the FHS has become one of the largest databases in its field.

5) NACC: The National Alzheimer's Coordinating Center (NACC) is responsible for developing and maintaining a database that collects participant information from 29 Alzheimer's Disease Centers (ADCs) funded by the National Institute on Aging (NIA). The database includes standardized clinical and neuropathological research data. The goal is to provide researchers with a longitudinally collected set of standardized assessment procedures to better characterize participants with mild Alzheimer's disease and mild cognitive impairment compared to non-dementia control groups.

6) BIOCARD: The overall objective of this study [186] was to identify variables that can predict the subsequent development of mild to moderate Alzheimer's disease symptoms in cognitively normal individuals. From 1995 to 2005, during the tenure of the National Institutes of Health (NIH) in the United States, participants underwent comprehensive neuropsychological testing annually. MRI scans, cerebrospinal fluid (CSF) samples, and blood specimens were collected approximately every 2 years. In 2009, the research team at Johns Hopkins University continued with annual clinical and cognitive evaluations, collected blood samples, and assessed previously obtained data. Due to funding constraints, CSF and MRI scans were not collected since the study moved to Johns Hopkins University. Please refer to Figure 7 for the timeline of data collection.

However, the aforementioned datasets do not include audio attribute data such as pause counts, reaction time, speech rate, etc. Therefore, in a series of studies using the sound modality for AD patient diagnosis, research teams have to recruit participants and collect voice data autonomously to meet specific research needs. These self-built datasets include the previously mentioned PITT Corpus, VBSD, Dem@Care, ADReSS, and others. Similarly, these datasets also do not contain posture attribute data from various body parts such as chest, hips, wrists, and ankles, resulting from actions such as sitting, standing, walking, and lying down. These data are primarily obtained from sensors worn by patients.

7) Smartphone-based single-sensor dataset: The dataset [187] consists of six Activities of Daily Living (ADLs), including sitting, standing, walking, lying, going upstairs, and going downstairs. These ADLs were performed by 30 participants ranging in age from 19 to 49 years. The participants wore a smartphone attached to their waist and followed a structured

STUDY CHARACTERISTICS	ADNI-1	ADNI-GO (Grand Opportunities)	ADNI-2	ADNI-3
Primary goal	Develop biomarkers as outcome measures for clinical trials	Examine biomarkers in earlier stages of disease	Develop biomarkers as predictors of cognitive decline, and as outcome measures	Study the use of tau PET and functional imaging techniques in clinical trials
Funding	\$40 million federal (NIA), \$27 million industry and foundation	\$24 million American Recovery Act funds	\$40 million federal (NIA), \$27 million industry and foundation	\$40 million federal (NIA), up to \$20 million industry and foundation
Duration/start date	5 years/October 2004	2 years/September 2009	5 years/September 2011	5 years/September 2016
Cohort	200 elderly controls 400 MCI 200 AD	Existing ADNI-1 + 200 early MCI	Existing ADNI-1 and ADNI-GO + 150 elderly controls 150 late MCI 150 AD	Existing ADNI-1, ADNI-GO, ADNI-2 + 133 elderly controls 151 MCI 87 AD

https://blog.osdn.net/C_ambage

Fig. 6. ADNI dataset. The ADNI series includes ADNI-1, ADNI-GO, ADNI-2, and ADNI-3. The figure summarizes them in the following terms: primary goal, funding, duration/start date, and cohort

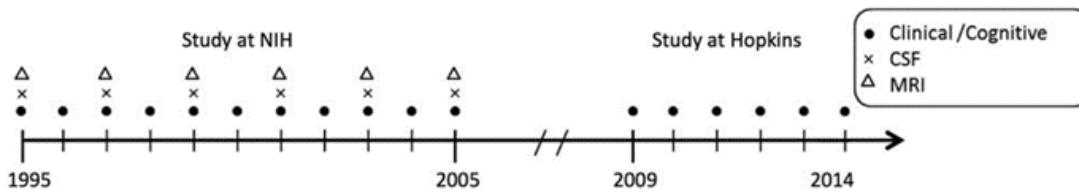


Fig. 7. BIOCARD: the timeline of data collection

protocol to perform the aforementioned ADLs. The dataset includes three-dimensional acceleration signal data (from the accelerometer) and three-dimensional angular velocity data (from the gyroscope). The sampling frequency of the smartphone sensors was 50Hz. For more details, please refer to Fig.IV.

8) Multi-sensor dataset: The dataset [187] was developed using four inertial sensors placed on the chest, right hip, right wrist, and ankle. A total of 19 participants (26 ± 8 years old) took part in the experiment. Data collection was performed using wireless sensors based on blink sensing technology. The sensors captured three-dimensional acceleration and three-dimensional angular velocity signals, with a sampling frequency of 204.8Hz. For more detailed information, please refer to Fig. IV.

At the gene-level granularity, previous studies have utilized gene expression profiles and DNA methylation profiles as datasets for research purposes [188]. The specific utilization of the dataset is depicted in Fig.V.

Apart from the aforementioned datasets, there are also some niche datasets collected independently by universities, research centers, or hospitals to meet specific research needs. These datasets can be seen in the "Dataset" column in the table below.

While general-purpose datasets can fulfill the needs of the majority of research studies, different studies often select subsets from these datasets based on factors such as the number of participants, their distribution, and the types of features available. It is ideal to have a balanced distribution of participants across gender, educational background, and age,

as it reduces sampling bias and enhances the persuasiveness of experimental results, allowing for generalization. Therefore, this paper has compiled the "Datasets and Feature Selection" information from various studies in this field in recent years, as shown in Table VI-IX.

MRI is the most commonly used imaging modality. Based on the size of the selected region of interest (ROI), the most commonly used feature extraction methods can be roughly classified into four categories.

1) Voxel-based methods [70] [156] [189] treat voxel intensity as features. However, analyzing features only on isolated voxels ignores the high correlation between voxels and has limitations such as high data dimensionality, computational complexity and overfitting.

2) Region-based methods [181] [190] [85] [191] [43] [163] [86] [192] [166] [193] [170] [194] [195] [196] [197] [198] extract features from predefined brain regions. Commonly selected regions include the hippocampus (HC), gray matter (GM), white matter (WM), corpus callosum (CC), cerebrospinal fluid (CSF) and cortical thickness (typically represented as vertex-level scales on the cortical surface [172] [199] [195]). The dimensionality of the data is significantly reduced compared to voxel-level features, making it the most commonly used scale for features.

3) Patch-based methods [72] [73] [79] [167] [184] [200] [201] learn new feature representations from local patches, providing a scale between voxel-level and region-level features that can better capture local features in the images.

4) Slice-based methods. Depending on the different directions of the two-dimensional slices, they can be divided into

TABLE IV
TWO POSE DATASETS BASED ON SENSORS

Author	Sampling Frequency	Sensor Location	Experiment Setting (Population)	Recorded Signals	Activities Analyzed
Anguita et al.	50 Hz	Waist-mounted Smartphone	Semi-naturalistic conditions(30 subjects) (19 to 48 years)	3D Acceleration Signal, 3D Angular Velocity	Standing, sitting, lying, walking, walking upstairs, Walking downstairs.
Leutheuser et al.	204.8 Hz	Wrist, hip, chest, ankle	Laboratory settings (23 young adults) (27 ± 7 years)	3D Acceleration Signal, 3D Angular Velocity	Walking, sitting, lying, Standing, stairs up, stairs down

TABLE V
DATASETS OF GENE EXPRESSION PROFILES AND DNA METHYLATION PROFILES

Dataset	Gene expression Rosetta/Merck Human 44k 1.1 microarray		DNA methylation Illumina Human Methylation 450 Bead Chip
GEO ID	GSE33000	GSE44770	GSE80970
Number of normal samples	157	100	68
Number of AD samples	310	129	74
Number of features	19,488 genes		485,577 probes for CpG site

axial [67] [202] [46], sagittal [192] [171], and coronal [203] [187] [198] slices. Some studies, such as [174] [204], use slices from all three directions, but most studies only use axial slices.

Preprocessing steps for MRI often include skull-stripping [173] [175] to optimize the results by removing the skull.

The scores obtained from cognitive assessment tests are commonly used features in cognitive modalities. Such as MMSE, MoCA, CDR-SOB, CDR-GLOB, ADAS-cog, RAVLT, FAQ, CERAD-NB, and DSRS [16] [63] [72] [85] [191] [193] [194] [200] [205] [196] [206]. These measures are interrelated, and under certain conditions, they can be converted or transformed into one another. For example, the conversion from MMSE to MoCA is discussed in [16]. However, each measure focuses on different cognitive aspects, highlighting the need for their combination to achieve complement, thereby providing a more comprehensive description of cognitive abilities.

Gait analysis is a less commonly used modality in cognitive research. It involves having participants wear sensors that detect the motion information of joints and limbs over time using gyroscopes, accelerometers, and other sensors. This information is then used as features in the analysis [131] [207] [187]. Voice modality, on the other hand, often directly utilizes existing feature sets such as eGeMAPS, emobase, ComParE, and MRCG as features [142] [208]. Some studies also use the spectral information of voice segments as features [40]. These modalities and their features have been discussed in detail in the previous text.

Gender, age, education level, and ethnic background are commonly used features in demographic modality [72] [46] [171]. These features provide a more detailed and comprehensive characterization of the participants' profiles. APOE4, SNPs, amyloid-beta protein, and cerebrospinal fluid (CSF)

are commonly used features in biomarker modality [63] [46] [209] [191] [193] [205]. Incorporating genetic and molecular biomarkers related to disease progression significantly improves the prediction results and facilitates pathological research on Alzheimer's disease (AD). This helps address the limited interpretability of deep learning models. Some biomarkers are often represented in the form of images, hence intersecting with the image modality.

B. Modality and method

1) Modality: In the compiled literature, the modalities involved include images, voiceprints, posture/gait, cognition, biomarkers (genes, cerebrospinal fluid, etc.), and demographic characteristics. Sometimes these modalities are further divided into different types of images, such as MRI, PET, DTI, EEG, or different biomarkers, such as cerebrospinal fluid, SNP, APOE4, Amyloid- β protein. Therefore, some papers refer to the combination of multiple types of images or biomarkers as multimodal fusion too. MRI, PET images, and cognitive assessment test scores are the most commonly used. Cognitive assessment test scores are often used with a certain cutoff value to determine if a person is an AD patient [16]. They can also be used to predict the change in test scores at future time points through regression, to determine if MCI subjects will convert to AD patients [193]. Several attempts have been made to combine images and cognition, demonstrating the superiority of both modalities [72] [85] [204] [200] [196] [206]. Since general datasets often include demographic data, this modality is commonly used as an auxiliary modality, combined with other modalities to exchange feature information and achieve comprehensive feature capture [194]. The fields of voiceprints and posture/gait are currently less explored, with most research focusing on a single modality. However, in [187], the combination of posture/gait and image modalities achieved excellent

TABLE VI
DATASET AND FEATURE SELECTION

Ref.	Year	Dataset	Sample	Feature
[16]	2020	Penn Memory Center + Clinical Core of the University of Pennsylvania's Alzheimer's Disease Center	587 subjects(321 AD,126 MCI,140 CN)	MMSE + MoCA + CERAD-NB + DSRS
[63]	2022	ADNI dataset	490 subjects (56 AD, 287 MCI, 147 CN)	MRI + Amyloid- β PET + CSF + Cognitive sub-scores + APOE + SNP
[131]	2021	A retrospective cohort of community-dwelling older adults	78 subjects (20 AD, 26 MCI, 32 CN)	Stride time + Step time + Single Support time + Swing time + Double support time + Stance time + Stride length + Step length
[40]	2020	VBSD dataset Dem@Care dataset	VBSD:36 subjects (23 AD, 13 CN) Dem@Care:32 subjects (24 AD, 8 CN)	Spectrogram features extracted from speech segments
[208]	2021	Pitt corpus	164 subjects (82 AD, 82 non-AD)	Feature sets: eGeMAPS + emobase + ComParE + MRCG
[142]	2020	ADReSS Challenge Dataset	156 subjects (78 AD, 78 non-AD) 4077 speech segments (2122 AD, 1955 non-AD)	Feature sets: eGeMAPS + emobase + ComParE + MRCG + Minimal + Linguistic
[181]	2018	ADNI dataset	202 subjects (51 AD, 99 MCI, 52 CN) 99 MCI (43 MCI-C, 56 MCI-NC)	MRI (Region-level: GM volume) + PET
[9]	2018	ADNI dataset	1242 subjects (360 sCN, 409 sMCI, 18 pCN, 217 pMCI, 238 AD)	FDG-PET + MRI
[210]	2020	ADNI dataset AIBL dataset FHS dataset NACC dataset	ADNI: 417 subjects (188 AD, 229 CN) AIBL: 382 subjects (62 AD, 320 CN) FHS: 102 subjects (29 AD, 73 CN) NACC: 565 subjects (209 AD, 356 CN)	MRI (Patch-level) + Age + Gender + MMSE
[73]	2020	ADNI dataset	1457 subjects (429 CN, 465 sMCI, 205 pMCI, 358 AD)	MRI (Patch-level)
[203]	2021	ImageNet dataset ADNI dataset	ImageNet: 1000 object categories ADNI: 450 subjects (225 AD, 225 CN)	MRI (Coronal slice)
[190]	2008	Database of the Centre Hospitalo-Universitaire of Caen, France	74 subjects (25 AD, 24 MCI, 25 CN)	MRI (Region-level: HC)
[46]	2022	ADNI dataset	Clinical data: 2004 subjects (598 CN, 699 MCI, 707 AD) Imaging: 502 subjects (132 CN, 104 MCI, 266 AD) Genetic: 809 subjects (245 CN, 338 MCI, 226 AD)	Clinical data: Demographics + Neurological exams + Cognitive sub-scores + Biomarkers + Medication + Imaging summary scores Imaging: Cross-sectional MRI Genetics: Whole genome sequencing
[211]	2021	ADNI dataset	361 subjects (92 CN, 82 sMCI, 95 pMCI, 92 AD)	MRI (Region-level: GM volume) + MMSE + ADAS-Cog
[67]	2021	OASIS-1 dataset Local hospitals	OASIS-1: 126 subjects (28 AD, 98 CN) Local hospitals: 70 subjects (70 AD)	MRI (Axial key slices)

TABLE VII
(CONTINUED)

Ref.	Year	Dataset	Sample	Feature
[79]	2021	ADNI dataset AIBL dataset	ADNI: 1193 subjects (389 AD, 172 pMCI, 232 sMCI, 400 CN); AIBL: 496 subjects (79 AD, 17 pMCI, 93 sMCI, 307 CN)	MRI (Patch-level)
[212]	2020	ADNI dataset	479 subjects (168 CN, 212 MCI, 99 AD)	MRI (Voxel-level: voxel intensity of GM and WM)
[213]	2021	ADNI dataset	1086 subjects (362 AD, 308 CN, 183 pMCI, 233 sMCI)	MRI (Region-level: HC) + PET
[209]	2020	ADNI dataset	72 subjects (37 AD, 35 CN)	fMRI + SNP
[207]	2019	Fujidera Assisted Living Facility Tsudou Elderly Care Center Daisen Elderly Care Center	10,833 behavior data from 90 subjects (Male: 32, Female: 68)	Average speed of stepping + The standard deviation of stepping speed + The ratio of correct answers + Average time of answering the calculus questions + The average height of knee joint + The standard deviation of height for the knee joint.
[191]	2015	ADNI dataset	202 subjects (51 AD, 43 pMCI, 56 sMCI, 52 CN)	(MRI + PET) (Region-level: ROI volume) + CSF + MMSE + ADAS-Cog
[188]	2019	Rosetta/Merck Human 44k 1.1 microarray + Illumina Human Methylation 450 Bead Chip	GSE33000: 157 CN, 310 AD GSE44770: 100 CN, 129 AD GSE80970: 68 CN, 74 AD	Gene expression (GEO ID: GSE33000, GSE44770) + DNA methylation (GEO ID: GSE80970)
[43]	2019	ADNI dataset	805 subjects (226 CN, 389 MCI, 190 AD)	MRI (Region-level: GM volume) + PET (Region-level: ROI intensity) + SNP
[163]	2019	ADNI dataset BIOCARD database	1143 subjects (324 from BIOCARD, 819 from ADNI)	MRI (Region-level: volumes of brain structures)
[86]	2019	ADNI dataset	801 subjects (198 AD, 160 pMCI, 214 sMCI, 229 CN)	MRI (Region-level: GM volume)
[192]	2019	ADNI dataset	830 subjects (198 AD, 167 pMCI, 236 sMCI, 229 CN)	Sagittal MRI(Region-level: GM volume)
[166]	2019	ADNI dataset	337 subjects (198 AD, 139 CN)	MRI(Region-level: GM + WM + CSF)
[193]	2019	ADNI dataset AIBL dataset	2146 subjects (1711 from ADNI, 435 from AIBL)	MRI(Region-level: HC) + MMSE + ADAS-Cog13 + APOE4 + RAVLT + FAQ + Amyloid-positive status + Age + Gender
[167]	2019	ADNI dataset	811 subjects (192 AD, 165 pMCI, 231 sMCI, 223 CN)	MRI (Patch-level: HC)
[184]	2020	ADNI dataset	449 subjects (97 AD, 233 MCI, 119 CN)	MRI (Patch-level: HC)
[168]	2018	IRCCS Centro Neurolesi Bonino-Pulejo of Messina (Italy)	189 EEG recordings (63 AD, 63 MCI, 63 CN)	EEG
[169]	2020	ADNI dataset	389 subjects (116 AD, 99 MCI, 174 CN)	fMRI
[214]	2020	ADNI dataset	1302 subjects (243 AD, 666 MCI, 393 CN)	FDG PET

TABLE VIII
(CONTINUED)

Ref.	Year	Dataset	Sample	Feature
[170]	2020	Chettinad Health City, Chennai	200 subjects (100 AD, 100 CN)	T2 flair weighted MRI (Region-level: WM, GM, CC, HC)
[171]	2020	ADNI dataset OASIS dataset ImageNet dataset	2179 subjects (436 from OASIS, 1743 from ADNI)	Sagittal MRI + Age + Gender
[194]	2019	ADNI dataset	785 subjects (192 AD, 181 pMCI, 228 sMCI, 184 CN)	CDRSB + ADAS11 + ADAS13 + RAVLT + APOe4 + Demographics + MRI (Region-level: parietal, temporal and frontal lobes)
[215]	2021	ADNI dataset	54 subjects (27 AD, 27 CN)	fMRI + PET
[172]	2021	ADNI dataset	1248 subjects (347 AD, 179 pMCI, 305 sMCI, 417 CN)	MRI (Vertex-level: cortical surface)
[173]	2023	ADNI dataset	210 subjects (70 AD, 70 MCI, 70 CN)	MRI (Skull stripped)
[174]	2021	ADNI dataset	1080 subjects (1406 AD scans, 2084 CN scans)	MRI (Sagittal + coronal + transverse view)
[175]	2022	ADNI dataset	1407 subjects (353 AD, 172 pMCI, 232 sMCI, 650 CN)	MRI (Skull stripped)
[187]	2019	ADNI dataset OASIS dataset Single-sensor dataset Multi-sensor dataset	ADNI: 432 subjects (232 AD, 200 CN) OASIS: 416 subjects (100 AD, 316 CN) Single-sensor: 30 subjects Multi-sensor: 19 subjects	MRI (Coronal slices) + 3D acceleration signal data + 3D angular velocity data
[195]	2019	ADNI dataset	2145 subjects (647 AD, 326 pMCI, 441 sMCI, 731 CN)	MRI (Region-level: HC) + PET (Region-level: HC/HC + Cortices)
[189]	2019	ADNI dataset	624 subjects	MRI (Voxel-level: whole volume)
[204]	2019	ADNI dataset	392 subjects (91 AD, 200 MCI, 101 CN)	MRI (2D slices of different views) + PET + MMSE + CDR
[200]	2020	ADNI dataset	1469 subjects	MRI (Patch-level) + CDR-SB + ADAS-Cog11 + ADAS-Cog13 + MMSE
[205]	2021	ADNI dataset	1371 subjects (339 AD, 140 pMCI, 473 sMCI, 419 CN)	MRI + PET + Cognitive sub-scores + APOE4 + Neuropathology data+ Neuropsychological battery + CSF
[196]	2021	ADNI dataset	805 subjects	MRI (Region-level: ROI volume) + Cognitive sub-scores
[197]	2019	ADNI dataset	1737 subjects, 12741 visits (2745 CN, 4058 MCI, 2108 AD, 3830 unlabeled)	MRI (Region-level: Ventricle + HC + Whole brain + Entorhinal cortex + Fusiform + Middle temporal gyrus)
[216]	2019	IRCCS Centro Neurolesi Bonino-Pulejo of Messina (Italy)	189 subjects (63 AD, 63 MCI, 63 CN)	EEG
[217]	2019	Beijing Easy monitor Technology	12 subjects (4 AD, 4 MCI, 4 CN)	EEG
[218]	2019	ADNI dataset	1303 subjects (243 AD, 667 MCI, 393 CN)	PET + Age
[198]	2021	ADNI dataset	798 subjects (187 AD, 382 MCI, 229 CN)	MRI (Region-level: Coronal slices of GM, WM and CSF)

TABLE IX
(CONTINUED)

Ref.	Year	Dataset	Sample	Feature
[219]	2020	ADNI dataset Severance Hospital	ADNI: 486 subjects (139 AD, 347 CN) Severance: 141 subjects (73 AD, 68 CN)	PET
[201]	2020	OASIS dataset	174 subjects (87 AD, 87 CN)	MRI (Patch-level)
[206]	2020	ADNI dataset	805 subjects	(MRI + Cognitive sub-scores) (Longitudinal data)
[153]	2019	OASIS dataset ImageNet dataset	OASIS: 150 subjects (75 CN, 75 AD) ImageNet: >1.2 billion labeled images	MRI
[220]	2019	ADNI dataset ImageNet dataset	ADNI: 150 subjects (50 AD, 50 MCI, 50 CN)	MRI (The most informative slices)
[156]	2019	ADNI dataset	706 subjects (279 AD, 427 CN)	MRI (Voxel-level)
[199]	2020	ADNI dataset	589 subjects (188 AD, 136 pMCI, 114 sMCI, 151 CN)	MRI (Vertex-level: cortical surface)
[202]	2020	OASIS dataset ImageNet dataset	Dataset-1: 180 subjects (90 AD, 90 CN) Dataset-2: 114 subjects (30 AD, 84 CN) Dataset-3: 200 subjects (100 AD, 100 CN)	Cross-sectional MRI
[159]	2020	ADNI dataset	1000 subjects (250 AD, 250 pMCI, 250 sMCI, 250 CN)	Sagittal MRI
[178]	2020	ADNI dataset	120 subjects (70 sMCI, 50 CN)	MRI + DTI

classification results with accuracy above 95% on several datasets, providing new ideas for future research to explore unexplored combinations of multiple modalities. There is no strict differentiation between genes and biomarkers, and they are often used to describe the same type of features. Studies that only use gene-level features are relatively scarce, such as [188], and they tend to focus more on clinical perspectives rather than computer-aided approaches, which do not meet the requirements of contemporary automated recognition. Therefore, genes and biomarkers are also often used as one of the modalities in multimodal fusion, adding richness to the data structure. The most common combination is still with image, cognition, and demographic modalities [193] [194] [205].

2) Method: In the compiled literature, several commonly used deep learning architectures are frequently applied.

Deep Feedforward Neural Network (DFFNN): DFFNN consists of multiple fully connected layers. Due to its high computational cost, it can be challenging to handle high-dimensional input data. Therefore, preprocessing of the input data is often required for dimensionality reduction. For example, in [188], the Limma package was used for feature extraction, [43] extracting features from predefined regions of interest (ROIs), while [163] focused on volume features of various ROIs in MRI. [86] employed a divide-and-conquer strategy by segmenting the gray matter (GM) into sub-blocks and using the voxels as input features for weak classifiers in DFFNN.

Convolutional Neural Network (CNN): CNN is the most widely used network architecture. In practical research, transfer learning methods are often employed to improve efficiency

and reduce costs by using pre-trained 2D CNN models directly for classification tasks. This requires converting the neuroimages into appropriate 2D slices, as mentioned in the "Feature Selection" section. [220] provided specific guidelines for selecting the most informative slices. Common input information for images includes voxel-level, patch-level, region-level, 2D slice and 1D numerical volume features. When dealing with multimodal information, preprocessing is often applied to concatenate the selected features as input, as done in [193], [194], [204] and [200].

Recurrent Neural Network (RNN): RNN is well-suited for utilizing sequential information. For example, [169] analyzed fMRI data as a time series of neuroimages, aligning with the concept framework of RNN for processing sequential data. [196] predicted the cognitive test scores at a target moment by inputting the scores of previous time periods, allowing for the prediction of whether MCI patients will convert to AD. [197] used Long Short-Term Memory (LSTM) to handle missing values in incomplete data. [205] utilized bidirectional LSTM by inputting images and cognitive modal data from four different time points to predict AD progression. Compared to the convolutional nature of CNN, RNN is less compatible with image data and is more suited for sequences such as speech, text and gait. However, the literature reviewed here does not include applications of RNN in these modalities. When neuroimages are decomposed into slices, patches, regions or other sub-elements, RNN can capture the spatial dependencies among these sub-elements.

Generative Networks (GN): [217] used a deep Boltzmann

machine (DCssCDBM), with integrated clear label information to guide feature extraction. [218] employed a Variational Autoencoder (VAE) to define anomaly scores based on the distance between given brain images and normal data. The model was trained using only normal brain PET images to identify abnormal patterns. [198] utilized Generative Adversarial Networks (GANs) as one of the weak classifiers in the voting decision. [219] employed Boundary Equilibrium GAN (BEGAN) with the encoder part of the discriminator network serving as input for an SVM classifier. [201] applied Supervised Switching Autoencoders (multi-channel reconstruction AEs) to individual slices of MRI patches and made predictions based on majority voting.

Deep Polynomial Network (DPN): DPN requires prior feature selection/extraction to find the most discriminative input features for training the network. For example, [206] used LASSO for feature extraction and then employed DPN for feature encoding.

This paper compiles the "Modeling methods" content from various studies in this field in recent years, as shown in Table X-XIV.

C. Evaluation index and performance

1) Evaluation index: In studies on AD classification based on computer-aided techniques, the effectiveness, reliability, and generalizability of the methods can be evaluated using several metrics to assess their applicability for clinical assistance and diagnosis. The main evaluation criteria include the following:

1. Accuracy (ACC): It represents the ratio of correctly classified samples to the total number of samples in a given test dataset. ACC reflects the probability of the classifier accurately classifying AD, MCI, and CN. Generally, higher values indicate better classifier performance.

2. Sensitivity (True Positive Rate, TPR): Also known as recall or hit rate, it represents the probability of correctly detecting all true positive cases. TPR measures the rate of correctly identified AD cases and indicates how well the classifier identifies AD. Larger values imply fewer cases of AD being missed.

3. Specificity (True Negative Rate, TNR): TNR, also known as the true negative rate, represents the probability of correctly detecting all true negative cases. It measures the classifier's ability to accurately identify true negatives, which reflects the classifier's capability to correctly classify CN. Generally, higher values indicate a lower probability of misdiagnosing normal individuals as AD.

4. Precision (Positive Predictive Value, PPV): It is the ratio of true positive samples to all samples classified as positive. Precision measures the proportion of correctly identified true positive cases among all samples classified as positive.

5. Area Under the ROC Curve (AUC): AUC represents the area under the curve plotted with TPR (true positive rate) on the y-axis and FPR (false positive rate) on the x-axis. A larger AUC value indicates better classification performance of the classifier.

6. F1-score: F1-score is the harmonic mean of precision and recall (sensitivity). Since precision and recall are conflicting

metrics, using a single precision or recall measure may be biased. Therefore, F1-score provides a balanced evaluation metric by considering both precision and recall.

When dealing with multi-class classification tasks, the calculation of accuracy, sensitivity, and specificity needs to consider the corresponding values for each class. Micro and Macro are two different weighting approaches for calculating these metrics, taking sensitivity as an example.

The calculation formulas are as follows:

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$Sens = Recall = TPR = \frac{TP}{TP + FN} \quad (2)$$

$$Spec = TNR = \frac{TN}{TN + FP} \quad (3)$$

$$PPV = \frac{TP}{TP + FP} \quad (4)$$

$$F1 - score = \frac{2 \times PPV \times TPR}{PPV + TPR} \quad (5)$$

$$FPR = \frac{FP}{TN + FP} \quad (6)$$

$$Sens_{macro} = \frac{\sum_{i=1}^n Sens_i}{n} \quad (7)$$

$$Sens_{micro} = \frac{\sum_{i=1}^n TP_i}{\sum_{i=1}^n TP_i + \sum_{i=1}^n FN_i} \quad (8)$$

Among them, TP (True Positive) represents the number of samples predicted as positive and are actually positive; FP (False Positive) represents the number of samples predicted as positive but are actually negative; FN (False Negative) represents the number of samples predicted as negative but are actually positive; TN (True Negative) represents the number of samples predicted as negative and are actually negative.

In addition to these commonly used performance metrics, some studies have adopted other evaluation metrics to better align with their model performance.

1. Matthews Correlation Coefficient (MCC): MCC is a measure used to evaluate the classification performance of binary classification tasks. It takes into account true positives, true negatives, false positives, and false negatives. MCC is considered a balanced metric that can be applied even when there is a large difference in sample sizes between the two classes. It essentially describes the correlation coefficient between the actual classification and the predicted classification, with values ranging from -1 to 1. A value of 1 indicates perfect prediction, 0 indicates predictions that are no better than random, and -1 indicates complete disagreement between predicted and actual classifications.

2. Fowlkes-Mallows Index (FMI): FMI is an external evaluation method used to determine the similarity between two clusters. A higher FMI indicates greater similarity between the clustering and the reference classification. The minimum possible value of FMI is 0, which corresponds to the worst

TABLE X
MODALITY AND METHOD

Ref.	Year	Modality	Method	Comment
[16]	2020	Cognition	Diagnostic accuracy and optimal cut-off scores were calculated for each measure, and a method for converting MoCA to MMSE scores was presented.	The MoCA is more sensitive and accurate. Complementing the MMSE or the MoCA with the DSRs significantly improved diagnostic accuracy.
[63]	2022	Biomarker + Image + Cognition + Genetics	Constructed subject similarity networks for all modalities and applied multilayer community detection to find groups with shared similarities across modalities.	It modeled each modality as a layer of a multidimensional network, so it could be equipped to handle interactions across modalities and find relation which might had been overlooked.
[131]	2021	Gait	A SVM model in a one-vs-one manner was trained for each classification, and the majority vote of the three models was assigned as CN, MCI, or AD.	Gait-based cognitive screening can be easily adapted into clinical settings.
[40]	2020	Acoustic/Linguistic data	Using the spectrogram features extracted from speech data and machine learning model LogisticRegressionCV.	Patient's speech data can effectively reduce the medical cost, and the data can be collected in real-time and accurately.
[208]	2021	Acoustic/Linguistic data	Using a new active data representation (ADR) method for feature extraction, together with "hard fusion" of feature sets to identify AD through decision tree.	A more comprehensive acoustic feature set is able to capture a wider range of speech subtleties. Besides, a larger number of features allows for the use of more sophisticated classifiers.
[142]	2020	Acoustic/Linguistic data	Operating linear discriminant analysis (LDA) on automatically extracted voice features to identify AD.	One strength of the method is its speaker independent nature.
[181]	2018	Image	Two SDPNs were first used to learn high-level features of MRI and PET, respectively, which were then fed to another SDPN to fuse multimodal neuroimaging information (MM-SDPN).	The proposed MM-SDPN can be a powerful representation algorithm for not only multimodal neuroimaging data but also other medical data.
[9]	2018	Image	Using a novel DNN based method that utilized multi-scale and multi-modal information (MMDNN) combining metabolism (FDG-PET) and regional volume (T1-MRI)	The approach preserved the structural and metabolism information at multiple scales.
[210]	2020	Image + Cognition	Developing a deep learning framework which linked a fully convolutional network (FCN) to a multilayer perceptron (MLP) and generates high resolution disease probability map.	The intuitive local probabilities outputted by the model were readily interpretable.
[73]	2020	Image	Proposing a hierarchical fully convolutional network (H-FCN) to identify discriminative patches and regions in sMRI.	The intuitive local probabilities outputted by the model were readily interpretable.
[203]	2021	Image	The CNN was ResNet-18 pre-trained on an ImageNet dataset. The employed sequence-based models were the temporal convolutional network (TCN) and different types of RNN.	The proposed models can be applied to any other image-based tasks
[190]	2008	Image	Automatically segmenting the Hippocampus with high spatial resolution, then adapting bootstrap method.	This automated method can serve as an alternative to manual tracing.
[46]	2022	Image + Genetics	Using stacked denoising auto-encoders for EHR and SNP, and 3D CNNs for MRI. Identifying top-performing features learned by the deep-models with clustering and perturbation analysis	Novel perturbation and clustering-based feature extraction assisting DL model interpretations were capable of AD stage prediction.
[211]	2021	Image + Cognition	Proposing a new framework based on the developed DBN-based MTL algorithm. The PCA and MTFS were introduced for the feature selection, and the dropout technique together with the zero-masking technology was utilized.	MTL strategy could further improve the performance by taking into account the correlation and difference among multiple tasks.

TABLE XI
(CONTINUED)

Ref.	Year	Modality	Method	Comment
[67]	2021	Image	Proposing a novel VGG-inspired network as the mainstay and combining the attention mechanism with VIN to produce a new ADVIAN deep-learning model. The 18-way DA was harnessed to prevent overfitting in the training set.	The model did not go through strict clinical environment tests.
[79]	2021	Image	Proposing a dual attention multi-instance deep learning network (DA-MIDL) including: 1) PatchNets with spatial attention blocks,2) an attention MIL pooling operation, and 3) an attention-aware global classifier.	The size of input patches was fixed.However, the structural changed in the cerebrum caused by brain atrophy may occur across multiple regions with different scales.Using the fixed size could not represent various local features.
[212]	2020	Image	Proposing a deep convolutional autoencoder (CAE) architecture, a tool that could perform an automatic non-linear decomposition of a very large dataset (over 2000 images).	It was a complete data-driven approach, self-supervised so that the final result was guided only by the particularities of the dataset, and very little pre-processing was required.
[213]	2021	Image	Adopting the Reversible Generative Adversarial Network (RevGAN) to reconstruct the missing data, the 3D CNN with multi-modality input to classify AD.	When the data was missing, the performance of AD diagnosis and MCI conversion prediction could be significantly improved using this method.
[209]	2020	Image + Genetics	Applying the correlation analysis to detect the associations between brain regions and genes, then proposing CERF to extract the discriminative fusion features between CN and AD.	The proposed model found out the abnormal brain regions and pathogenic genes of AD such as thalamic, lingual gyrus, angular gyrus, precuneus, insula DAB1 and LRP1B gene.
[207]	2019	Gait	Utilizing Kinect device for collecting gait data to extract 12 features of single and dual-task performance, then developing a method for automatic dementia score estimation to investigate which characteristics are the most important.	The combination of single and dual-task performance was useful for MMSE score classification,
[191]	2015	Image + Cognition + Biomarker	Utilizing the dropout technique to to prevent weight co-adaptation. In addition, incorporating stability selection, an adaptive learning factor, and a multi-task learning strategy into restricted Boltzmann machine (RBM).	Classical deep learning did not perform well on this small dataset, but with the dropout technique, the average accuracy was improved by 5.9% on average.
[188]	2019	Genetics	Using Limma package to identify DEGs and DMFs, then integrating them by intersecting, followed by a DFFNN whose hyper-parameters were optimized by Bayesian Optimization	It is supposed that such study will receive much attention owingconnectionng to the possibility of biomedical explanation and connection to the mechanism study.
[43]	2019	Image + Genetics	Learning latent representations for MRI, PET and Genetic data by 3 DFFNNs, then leaning the resultant feature representations for each pair of modality combination by another 3 DFFNNs and finally, feeding all these feature maps into the last DFFNN.	The method was focusing on using ROI features as input to the deep learning model. However, such handcrafted features may limit the richness of structural and functional brain information from MRI and PET images
[163]	2019	Image	Siamese structure adopted on DFFNN to quantify the asymmetry of right and left brain hemispheres as a sign of MCI or AD concatenating all the outputs into a vector.	Siamese networks had the advantage of reduced complexity and computational time.
[86]	2019	Image	Segmenting GM into regions using atlas, then frequently and randomly selecting some of the voxels in each region and feeding them into DFFNN as some weak classifiers. Then concatenating all the outputs into a vector.	This region-based abnormality method is visualizable and interpretable.
[192]	2019	Image	Applying independent 3D CNNs for each time point and feeding the output feature maps to 3 stacked bidirectional GRUs,concatenating all the outputs in an FC layer followed by Softmax.	The proposed method is a data-driven method to jointly learn the spatial and longitudinal features and disease classifier model.

TABLE XII
(CONTINUED)

Ref.	Year	Modality	Method	Comment
[166]	2019	Image	Independent multi-scale 3D CNN for GM, WM and CSF inputs to extract feature maps of different levels and applying XGBoost on the concatenated features for determining the most important features followed FC and Softmax layers	The dataset used in experiments was moderate in size, which led to some overfitting.
[193]	2019	Image + Cognition + Biomarker + Genetics	CNN to extract image features and combining them with non-image biomarkers for time-to-invert prediction by LASSO Cox regression for MCI subjects	It found that amyloid-positive MCI with molecular evidence of prodromal AD had higher predicted AD dementia progression risks than amyloid-negative MCI.
[167]	2019	Image	Applying Parallelized 3D DenseNet on the central patch and MLPb on the shape description extracted by SPHARM-PDM tool for bilateral Hippocampi and combining 4 resultant feature maps in FC layers followed by Softmax.	The proposed method can make full use of the visual and shape features of hippocampus.
[184]	2020	Image	Multi-task CNN with ResNet block to segment hippocampus and using different levels of feature maps, concatenating them with that of hippocampus obtained by another CNN.	Multi-task deep model and DenseNet learned complementary features for disease classification.
[168]	2018	Image	Evaluating the power spectral density (PSD) of the 19-channels EEG traces and representing the related spectral profiles into 2D gray scale images (PSD-images).	The originality of the proposed method lies in mapping the power spectrum of each subject into a 2D gray scale image.
[169]	2020	Image	CNN for feature representation of different time nodes of 4D fMRI and feed them to an LSTM layer followed by Softmax.	The proposed method can extract the spatio-temporal property of fMRI data fully for diagnosis of AD.
[214]	2020	Image	3D-CNN to determine the likelihood of progression to AD for MCI subjects	The deep learning-based cognitive function evaluation model could be transferred to multiple disease domains.
[170]	2020	Image	Gray wolf optimization (GWO) for brain segmentation and AlexNet for classification.	The pipeline of the process adopted in this study can be extended to other neuro developmental disorders.
[171]	2020	Image	Using TL technique, sagittal MRI images and ANN ResNet feature extractor with the SVM classifier.	TL allows experiments with little data as well as data augmentation.
[194]	2019	Image + Demographic +Genetics + Cognition	Consider the classification performance of the proposed network on four different input biomarker combinations.	The approach is flexible and can in principle integrate other imaging modalities and diverse other sets of clinical data.
[215]	2021	Image	VGG-16 for extracting features, SVM, Linear Discriminate, Kmeans clustering and Decision tree for classification.	Main objective of this paper is to include a preprocessing method before CNN model to increase accuracy.
[172]	2021	Image	Parallel ResNet-10 networks for each cortical region and integrating them to a sparse regression while these two modules are updated alternatively and iteratively	It incorporated a sparsity mask generated by sparse regression module into a deep learning module for simultaneous AD diagnosis and identification of critical regions.
[173]	2023	Image	An experimental analysis of twenty popular CNNs was performed using MRI. It was observed that the overall classification accuracy of DenseNet-121 is the best.	To improve the execution time, it had proposed replacing the convolution layers in the original DenseNet-121 architecture with depth-wise convolution layers.
[174]	2021	Image	ResNet-29 was trained by transfer learning. CN and AD scans were used to pretrain the model, then model was retrained on the target task of pMCI vs sMCI.	It avoided problems previously neglected such as data shortage, high variance and data leakage.
[175]	2022	Image	Proposing an explainable 3D Residual Attention Deep Neural Network (3D ResAttNet) for end-to-end learning from sMRI.	The explainable mechanism is able to identify and highlight the contribution of the important brain parts.

TABLE XIII
(CONTINUED)

Ref.	Year	Modality	Method	Comment
[187]	2019	Image + Gait	Monitoring the subjects' activities of daily living using body worn inertial sensors to get physical activity data, then classify target together with MRI.	This paper correlates the importance of high accuracy AD diagnosis with physical activities and how these activities can be logged without human intervention with high accuracy.
[195]	2019	Image	Two paralleled VGG11 networks for each modality (MRI + PET) and concatenating the feature maps to being fed into some FC layers followed by Softmax layer.	It indicated that the hippocampal area with no segmentation can be chosen as the input.
[189]	2019	Image	Dense connections were introduced to maximize the information flow. And probability-based fusion method was employed to combine 3D-DenseNets with different architectures.	Dense connections improve the information and gradients propagation throughout the network, then make the network easier to train due to less parameters.
[204]	2019	Image + Cognition	Two independent CNNs were used to extract features of MRI and PET, calculating correlation coefficient and combining the outputs with the results of clinical psychological diagnosis.	The method combined the neuroimaging diagnosis with the clinical neuropsychological diagnosis, so the process is closer to the process of clinician's diagnosis and easy to implement.
[200]	2020	Image + Cognition	Using two scales of landmark patches as two channels into a densely connected CNNs for each landmark and concatenating these feature maps to being fed into some FC layers to predict cognitive scores of four follow-up time points.	The current network primarily worked for estimating multiple types of clinical scores, without considering the underlying association between clinical scores and class labels of subjects.
[205]	2021	Image+ Cognition+ Genetics+Biomarker	Applying Bidirectional LSTM on the multi modal time series data for feature representation and using them in two models.	Although the DFBL architecture is more accurate than the MRBL architecture, the latter is more interpretable.
[196]	2021	Image + Cognition	Combining group LASSO and correntropy to reduce dimension and obtain features, then using features in a stacked independent RNN to predict cognitive scores.	By stacking multiple layers IndRNN, it was possible to study the cross-channel information in time and obtain deep characteristic information.
[197]	2019	Image	Applying a modified LSTM to work with incomplete data while LDA providing class labels.	This study highlighted the potential of RNNs for modeling the progression of AD using longitudinal measurements.
[216]	2020	Image	Exploiting the potential of both CWT and Bis features extracted from time-frequency and bispectrum of EEG.	This study focused on the AD/MCI/ICN classification task and did not address prediction of early onset of AD.
[217]	2019	Image	Using multi-task learning strategy based on discriminative convolutional high-order Boltzmann Machine with hybrid feature maps.	The designed model was well-performed in comparison to other generative model since it bridged the connection between feature extraction and classification.
[218]	2019	Image	Using variational autoencoder, abnormality score was defined as how far a given brain image is from the normal data.	The approach has advantages in the flexible application for heterogeneous patients, even for uncommon disorders.
[198]	2021	Image	A majority voting ensemble was used to combine multi-slice predictions with complementary results VGG16, ResNet50, and DCGAN were used to construct an ensemble classifier.	This technique may enhance AD diagnostics when the sample size is limited.
[219]	2020	Image	Applying slice selective learning to reduce computational cost and to extract unbiased features. Extracting features by a Boundary Equilibrium Generative Adversarial Network	When there were insufficient datasets to train deep neural networks with single-source training datasets, this approach was a feasible alternative to use datasets from various hospitals.
[201]	2020	Image	Applying Supervised Switching AEs on extracted patches of a single MRI slice while prediction is based on a majority rule from the ensemble of patch predictions	It used only one scan slice per subject, instead of the whole 3D volume.

TABLE XIV
(CONTINUED)

Ref.	Year	Modality	Method	Comment
[206]	2020	Image + Cognition	LASSO for feature extraction, deep polynomial network for feature encoding followed by an SVR to predict cognitive scores.	The proposed framework can effectively reveal the relationship between clinical scores and MRI data.
[153]	2019	Image	Proposing a pretrained AlexNet model based on TL technique Using the main view of the brain and the MRI with extensive image augmentation techniques to avoid overfitting issue.	Extensive augmentation approaches can prevent overfitting issues in class balance dataset.
[220]	2019	Image	Using a pretrained VGG16 network for transfer learning.	Transfer learning effectively reduces computational cost.
[156]	2019	Image	Presenting an effective strategy of augmenting volumetric data in a 3D-CNN neural network model.	It had shown that using a set of augmentation methods postponed the overfitting without any other regularization technique.
[199]	2020	Image	Using spherical CNN formulation as an effective deep learning framework for modeling human cortex.	This work demonstrated the feasibility and superiority of the spherical CNN directly applied on the spherical representation in the discriminative analysis of the human cortex.
[202]	2020	Image	The whole brain image was passed through two transfer learning models: Inception V3 and Xception, as well as a custom CNN with separable convolutional layers.	Experimental results show that the transfer learning approaches exceed the performance of non-transfer learning based approaches.
[159]	2020	Image	Using DCNN and VGG-16 inspired CNN (VCNN) models.	VCNN performed better.
[178]	2020	Image	A CNN based on transfer learning is developed to extract features of MRI and DTI, where an L1-norm is introduced to reduce the feature dimensionality by LASSO algorithm.	DTI data is an effective supplement to sMRI, and it acts as an significant biomarker of MCI.

binary classification where all elements are misclassified. The maximum possible value is 1, which corresponds to the best binary classification where all elements are perfectly classified.

3. Mean Absolute Error (MAE): MAE measures the average absolute difference between predicted values and true values. It has a range of $[0, +\infty)$, where a value of 0 indicates a perfect model with predictions matching the true values exactly, and larger values indicate larger errors.

4. Pearson Correlation Coefficient (R): R measures the linear correlation between two sets of data. It has a range of $[-1, 1]$, where 0 represents no correlation, negative values indicate negative correlation, and positive values indicate positive correlation.

5. Balanced Accuracy (BACC): BACC is a performance metric that is better than raw accuracy when the sample sizes of each class (normal and symptomatic) are imbalanced.

6. C-index (Concordance Index): C-index is used to evaluate the predictive ability of a model. It represents the proportion of pairs where the predicted results are consistent with the actual results among all pairs of patients. It estimates the probability of concordance between predicted and observed outcomes.

The calculation formulas/methods are as follows:

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (9)$$

$$FMI = \sqrt{PPV \cdot TPR} \quad (10)$$

$$R(X, Y) = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y} \quad (11)$$

$$BACC = \frac{TPR + TNR}{2} \quad (12)$$

The C-index is calculated by randomly pairing all the research subjects. Taking the prediction of AD as an example, in each pair, at least one individual will progress to AD. If the predicted probability of developing AD is higher for the individual who actually progresses to AD compared to the other individual, it is considered a consistent prediction with the actual outcome.

2) Performance: The literature covered in this article mainly involves several types of prediction: binary classification, multiclass classification and regression prediction. The most common classification involves categorizing subjects into AD, sMCI, pMCI and CN (also known as NC/HC). Among them, the binary classification of sMCI and pMCI predicts whether MCI subjects will progress to AD in the future.

Due to the varying evaluation metrics used in different studies and the differences in selected modalities, datasets and target requirements, the performance of the same model can vary significantly across different classification tasks. Therefore, it is not appropriate to use a single metric to evaluate the superiority or inferiority of all models. The selection of a suitable model should be based on specific requirements. For example, although the accuracy of [131] is lower than that of similar studies in both classification tasks (less than 87%), its gait feature collection cost is much lower than medical

imaging. Similarly, although the classification accuracy of [40] is not top-notch, speech data has the advantages of low cost and real-time availability. [163] achieved significant results in reducing problem complexity and computation time, but at the expense of sacrificing some accuracy. [86] proposed a method with the advantages of result visualization and interpretability, although its performance metrics are excellent, but not state-of-the-art. On the other hand, although [67] achieved an accuracy of 97.76% in the AD vs CN task, with other metrics also being extremely high, the experiment did not undergo rigorous clinical testing.

However, from the experimental results, it can be observed that combining multiple modalities often outperforms using a single modality as features. This approach can balance and improve overall performance, compensating for the limitations of single modality data in capturing comprehensive features. For example, [85] combined imaging with cognitive modalities, achieving accuracy higher than 90% in all binary classification tasks, with CN vs AD and sMCI vs AD reaching accuracy levels of 98.62% and 99.62% respectively, although other metric results were not provided. [187] combined imaging with gait modalities and achieved classification accuracy above 95% for AD vs CN on various datasets, as well as sensitivity, specificity, and precision above 90%.

This article summarizes the "Evaluation index and Performance" of various studies in this field in recent years, as shown in Table XV-XX.

VIII. CONCLUSION

This article discusses the disease mechanism of Alzheimer's disease and an overview of diagnostic methods, analyzes the advantages and disadvantages of each diagnostic method, and summarizes the prospects of multimodal diagnostic methods in the diagnosis of AD.

The study started with ancillary diagnostic techniques for earlier classification studies, such as cognitive assessment, voice detection, and posture method. Each method has its own merits. Cognitive assessment can directly measure a patient's cognitive function and has been widely used in AD diagnosis and monitoring, demonstrating high clinical validation and reliability. Some methods based on blood or cerebrospinal fluid samples have lower invasiveness and risk compared to other imaging techniques. Similarly, posture and sound methods are non-invasive approaches, and their data can be conveniently obtained through devices. Neuroimaging techniques can obtain brain images at high resolution, revealing subtle changes in brain structure and function. Researchers can quantitatively analyze the images using quantitative analysis methods to support AD diagnosis and monitoring.

But these methods also have some obvious limitations. Some methods may encounter interference during the data collection process. For example, cognitive assessment may be influenced by individual factors such as educational level. Sound detection may be disrupted by environmental noise, unclear articulation, and other factors, affecting feature extraction and interpretation. Other methods require expensive equipment and specialized operations, making them difficult to

TABLE XV
EVALUATION INDEX AND PERFORMANCE

Ref.	Year	Target	Performance				
			Acc (%)	Sens (%)	Spec (%)	AUC	others
[16]	2020	CN vs MCI(MMSE+DSRS)	97			0.97	
		CN vs MCI(MoCA+DSRS)	97			0.97	
		MCI vs AD(MMSE+DSRS)	93			0.91	
		MCI vs AD(MoCA+DSRS)	92			0.9	
[63]	2022	AD vs non-AD	≈90				Precision (%): 86.9
		MCI->CN prediction	88.5	>85	>85		
		MCI->AD prediction	90.8	>85	>85		
[131]	2021	CN vs MCI vs AD	78				F1-score: 0.77
		CN vs MCI/AD	86				F1-score: 0.88
[40]	2020	CN vs AD: VBSD dataset	83.3	86.9			Precision (%)
		Dem@Care dataset	84.4	87.5			86.9 91.3
							0.869 0.894
[208]	2021	AD vs non-AD	78.7	80.5	76.8		
[142]	2020	AD vs non-AD: Acoustic	62	75			Precision (%)
		Linguistic	75	62			60 83
							0.67 0.71
[181]	2018	CN vs AD(a)	97.13	95.93	98.53	0.972	
		MCI vs CN(a)	87.24	97.91	67.04	0.901	
		MCI-C vs MCI-CN(a)	78.88	68.04	86.81	0.801	
		AD vs MCI-C vs MCI-NC vs NC(a)	57	53.65	85.05		
[9]	2018	sMCI vs pMCI	82.93	79.69	83.84		
		sCN vs AD	84.6	80.2	91.8		
		sCN vs (pMCI & AD)	86	85.7	86.5		
		sCN vs. (pCN, pMCI and AD)	86.4	86.5	86.3		
[210]	2020	CN vs AD: ADNI dataset	96.8	95.7	97.7	0.996	F1-score (%)
		AIBL dataset	93.2	87.7	94.3	0.974	96.5
		FHS dataset	79.2	74.2	80.8	0.867	0.937
		NACC dataset	85.2	92.4	81	0.954	81.4
							0.517
[73]	2020	CN vs AD	90.3	82.4	96.5	0.951	
		pMCI vs sMCI	80.9	52.6	85.4	0.781	
[203]	2021	CN vs AD	91.78	91.56	92		
[190]	2008	CN vs AD	84	84	84	0.913	
		MCI vs CN	73	75	70	0.808	
		AD vs MCI	69	67	71	0.721	
[46]	2022	CN vs MCI vs AD: EHR + SNP + Random Forest	78	79			Precision (%)
		EHR + Imaging + Random Forest	77	77			78 76
							0.78 0.77

TABLE XVI
(CONTINUED)

Ref.	Year	Target	Performance				
			Acc (%)	Sens (%)	Spec (%)	AUC	others
[211]	2021	sMCI vs pMCI	87.78				
		CN vs AD	98.62				
		CN vs sMCI	92.31				
		CN vs pMCI	96.67				
		sMCI vs AD	99.62				
		pMCI vs AD	91.89				
[67]	2021	CN vs AD	97.76	97.65	97.86	0.9852	Precision: 97.87 F1-score: 0.9775 MCC: 0.9553 FMI: 97.76
[79]	2021	CN vs AD on ADNI dataset	92.4	91	93.8	0.965	
		pMCI vs sMCI on ADNI dataset	80.2	77.1	82.6	0.851	
		pMCI vs CN on ADNI dataset	89.5	82.4	92.5	0.917	
		sMCI vs CN on ADNI dataset	82.5	80.4	83.8	0.86	
		CN vs AD on AIBL dataset	90.2	84.8	91.5	0.939	
		pMCI vs sMCI on AIBL dataset	80.9	70.6	82.8	0.824	
[212]	2020	Model: CAE-SVC					F1-score
		CN vs AD (Tissue: GM)	77.1	75.9	78.3		0.768
		sMCI vs pMCI (Tissue: WM)	70.4	69.3	72.6		0.709
[213]	2021	CN vs AD(a), sMCI vs pMCI(b)					
		MRI + PET(a)	92.28	90.38	94.37	0.9276	
		PET + 100% synthetic MRI(a)	90.77	90.58	90.98	0.916	
		MRI + PET(b)	74.1	75	73.08	0.766	
[209]	2020	MRI + 100% synthetic PET(b)	73.78	64.52	85.92	0.7509	
		CN vs AD (Pearson + CERF)	86.2				
[207]	2019	MMSE \geqslant 24 vs MMSE $<$ 24:					
		Single-task features (Physical)		38.4	87.8		
		Combined features		75.3	79.9		
[191]	2015	AD vs CN	91.4				
		MCI vs CN	77.4				
		AD vs MCI	70.1				
		MCI.C vs MCI.NC	57.4				
[188]	2019	AD vs CN	82.3				
[43]	2019	CN vs MCI vs AD	\approx 65				
		CN vs sMCI vs pMCI vs AD	\approx 55				
		AD vs CN	\approx 92				
		CN vs MCI	\approx 76				
[163]	2019	CN vs MCI/AD		88.39	95.84		BACC: 0.9220 F1-score: 0.9259

TABLE XVII
(CONTINUED)

Ref.	Year	Target	Performance				
			Acc (%)	Sens (%)	Spec (%)	AUC	others
[86]	2019	AD vs CN	92.75	91.89	93.47	0.9804	
		MCI vs CN	89.22	93.33	82.55	0.9573	
		AD vs MCI	81.46	68.66	88.24	0.8954	
		pMCI vs sMCI	88.52	87.5	89.22	0.9568	
		AD vs MCI vs CN(Macro)	71.18	54.97	72		
		AD vs MCI vs CN(Micro)	71.18	55.7	71.18		
[192]	2019	AD vs CN	91.33	86.87	95.2	0.9322	
		pMCI vs sMCI	71.71	65.27	76.27	0.7303	
[166]	2019	AD vs CN	94.74				
[193]	2019	AD vs CN on ADNI-GO&2	90			0.956	
		AD vs CN on AIBL	92.9			0.958	
		From MCI to AD?				0.813	C-index: 0.864
[167]	2019	AD vs CN	92.29	90.63	93.72	0.9695	
		pMCI vs sMCI	75	73.33	76.19	0.797	
		MCI vs CN	74.64	77.27	69.96	0.777	
[184]	2020	AD vs CN	88.9	86.6	90.8	0.925	
		MCI vs CN	76.2	79.5	69.8	0.775	
[168]	2018	AD vs CN	92.95	95.3		0.97	Precision (%)
		AD vs MCI	84.62	85.04		0.93	F1-score
		MCI vs CN	91.88	92.31		0.97	91.02
		AD vs MCI vs CN	83.33	82.91		0.94	0.8468
[169]	2020	AD vs MCI	92.11			0.92	
		MCI vs CN	88.12			0.89	
		AD vs CN	97.37			1	
		AD vs CN vs MCI	89.47				
[214]	2020	AD vs CN				0.94	
		pMCI vs sMCI				0.82	
[170]	2020	AD vs CN	95	95	94		
[171]	2020	OASIS: AD vs pMCI vs sMCI vs CN	86.81	37.81	88.25		Precision (%)
		ADNI: AD vs MCI vs CN	78.64	58.28	80.06		F1-score
[194]	2019	pMCI vs sMCI	86	87.5	85	0.925	
[215]	2021	AD vs CN on average (fMRI)	99.95			1	
		AD vs CN on average (PET)	73.46			0.75	
[172]	2021	AD vs CN	95.32	91.18	93.94		
		pMCI vs sMCI	77.6	71.62	75.85		

TABLE XVIII
(CONTINUED))

Ref.	Year	Target	Performance					
			Acc (%)	Sens (%)	Spec (%)	AUC	others	
[173]	2023	CN vs MCI for 60-69 years	93	92			Precision (%)	F1-score
		CN vs MCI for 70-79 years	92	92			93	0.94
		CN vs MCI for 80+ years	87	87			94	0.93
		MCI vs AD for 60-69 years	92	92			90	0.9
		MCI vs AD for 70-79 years	89	91			90	0.93
		MCI vs AD for 80+ years	87	86			87	0.86
		CN vs AD for 60-69 years	92	94			94	0.93
		CN vs AD for 70-79 years	90	90			91	0.88
		CN vs AD for 80+ years	86	87			87	0.86
		CN vs MCI vs AD for 60-69 years	90	89			89	0.87
		CN vs MCI vs AD for 70-79 years	87	88			90	0.91
		CN vs MCI vs AD for 80+ years	87	85			89	0.86
[174]	2021	pMCI vs sMCI	82.4					
[175]	2022	AD vs CN	91	91	92	0.984		
		pMCI vs smCI	82	81	81	0.92		
[187]	2019	AD vs CN: OASIS	95.93	92.75	92.51		Precision (%)	F1-score
		ADNI	98.74	98.5	98.21		96.94	0.947
		Single sensor on smartphone	99	96.95	96.41		98.81	0.9865
		Multi-sensor	99.6	98.41	99.71		97	0.9697
[195]	2019	AD vs CN	90.1	90.85	89.21	0.9084		
		CN vs pMCI	87.46	90.73	80.61	0.8761		
		sMCI vs pmCI	76.9	68.15	83.93	0.7961		
[189]	2019	AD vs MCI	93.61	92.45			Precision(%):	94.59
		MCI vs CN	98.42	98.34			Precision(%):	98.37
		AD vs CN	98.83	98.7			Precision(%):	98.70
		AD vs CN vs MCI	97.52					
[204]	2019	AD vs CN	98.47	96.58	95.39	0.9861		
		MCI vs CN	85.74	90.11	91.82	0.8815		
		AD vs MCI	88.2	97.43	84.31	0.8801		
[200]	2020	Correlation Coefficients of cognitive scores at M24 (Model was trained on ADNI-1, and tested on ADNI-2)					SDR-SB:	0.527
							ADAS-Cog11:	0.552
							ADAS-Cog13:	0.583
							MMSE:	0.537
[205]	2021	AD vs CN vs MCI	84.95	86.5			Precision (%)	F1-score
		AD vs CN vs pMCI vs sMCI	86.08	87.47			85.65	0.8607
							87.2	0.8733

TABLE XIX
(CONTINUED))

Ref.	Year	Target	Performance				
			Acc (%)	Sens (%)	Spec (%)	AUC	others
[196]	2021	Using CLSIndRNN to predict M36:					MAE
		MMSE					1.56
		CDR-SOB					0.81
		CDR-GLOB					0.13
		ADAS-Cog					3.33
[197]	2019	CN vs MCI				0.5914	R
		CN vs AD				0.9029	0.76
		MCI vs AD				0.7844	0.67
		CN vs MCI vs AD				0.7596	0.85
[216]	2019	AD vs CN	96.95	94.91			Precision (%)
		AD vs MCI	90.24	85.71		0.965	94.72
		MCI vs CN	96.24	95.86			0.9485
		AD vs MCI vs CN	89.22	80.99			87.14
[217]	2019	AD vs MCI vs CN	95.04			0.96	0.8646
[218]	2019	AD vs CN				0.9	95.31
[198]	2021	AD vs CN	90.36	93.94	83.78	0.8972	0.9558
		AD vs MCI	77.19	68.97	54.06	0.7118	
		MCI vs CN	72.36	74.71	84.42	0.6829	
		pMCI vs sMCI	66.7				
[219]	2020	AD vs CN on ADNI dataset	94.82	91.78	97.06	0.98	
		AD vs CN on Sev.	94.33	92.11	97.45	0.98	
[201]	2020	AD vs CN	90	95	85	0.92	BACC: 0.90
[206]	2020	Using CTDE to predict M36:					MAE
		MMSE					1.756
		CDR-SOB					0.816
		CDR-GLOB					0.13
		ADAS-Cog					4.981
[153]	2019	AD vs CN	98.41				
[220]	2019	AD vs MCI vs CN	95.73				Precision (%)
		AD vs CN	99.14	98.75			99.68
		AD vs MCI	99.3	98.75			100
		CN vs MCI	99.22	99.06			99.37
[156]	2019	AD vs CN	88.8	90			F1-score
[199]	2020	AD vs CN	92.16	89.73	95.05	0.9459	
		pMCI vs sMCI	76.47	82.62	57.39	0.7445	
[202]	2020	AD vs CN	64.87	64.88	64.86		MCC
		Dataset-1 (Xception, Fold-9)	82.79				0.2975
		Dataset-2 (CNN, Fold-5)	99.45				0.6488
Dataset-3 (Inception V3, Fold-5)							

TABLE XX
(CONTINUED))

Ref.	Year	Target	Performance				
			Acc (%)	Sens (%)	Spec (%)	AUC	others
[159]	2020	CN vs sMCI	96.19				
		AD vs CN	93.78				
		CN vs pMCI	94.87				
		AD vs sMCI	92.8				
		AD vs pMCI	94.79				
		pMCI vs sMCI	92.91				
[178]	2020	CN vs sMCI	94.2	97.3	92.9		

be widely used in clinical and large-scale applications, such as various neuroimaging techniques. Additionally, some methods may not directly provide information about brain structure and function, limiting the comprehensive evaluation of AD, such as posture detection. Using voice, pose, biomarkers, cognitive assessment scores, and their fusion modalities can often bring more stable and robust diagnostic results to the model. Among them, we discuss in detail such as the fusion of PET and MRI, the fusion of pose and text, the fusion of prosodic features and speech content and so on. These fusion methods offer significant advantages over single-modality approaches. They leverage the complementary information captured by different modalities, leading to more stable and robust diagnostic results. The integration of diverse data sources allows for a more comprehensive assessment of AD, considering both structural and functional changes, as well as cognitive and physical impairments.

In the process of investigating the multi-modal fusion model, we found that there are currently two mainstream fusion solutions, one is the fusion of different data levels of a single modality, and the other is the data fusion of different modalities. The former requires more attention to combine features extracted from different levels of data within a single modality and combine the decisions or predictions made independently at different data levels within a single modality. The latter focus on the integration of data from different modalities to leverage their complementary information. It also need to learn joint representations or features that capture shared information across different modalities.

In general, multimodal fusion methods have shown a lot of potential, but there are still many problems. We believe that the main challenges in this field focus on (1) how to align different modality data so that they can work better together and improve the accuracy of diagnosis. (2) How the modal data interact with each other, discover the correlation between the modal and the modal, and make more contributions to the interpretability of the diagnosis. (3) Existing medical images, pose data, and sound data are all separate, and there is no set of relatively complete multi-modal data sets for researchers to use.

REFERENCES

- [1] “2023 alzheimer’s disease facts and figures,” *Alzheimer’s & Dementia*, vol. 19, no. 4, pp. 1598–1695, 2023.
- [2] H. Zhang, J. Chen, F. Jiang, F. Yu, Z. Chen, J. Li, G. Chen, X. Wu, Z. Zhang, Q. Xiao, X. Wan, B. Wang, and H. Li, “Huatuogpt, towards taming language model to be a doctor,” 2023.
- [3] H. Wang, C. Liu, N. Xi, Z. Qiang, S. Zhao, B. Qin, and T. Liu, “Huatu: Tuning llama model with chinese medical knowledge,” 2023.
- [4] H. Xiong, S. Wang, Y. Zhu, Z. Zhao, Y. Liu, L. Huang, Q. Wang, and D. Shen, “Doctorglm: Fine-tuning your chinese doctor is not a herculean task,” 2023.
- [5] S. Fathi, M. Ahmadi, and A. Dehnad, “Early diagnosis of alzheimer’s disease based on deep learning: A systematic review,” *Computers in Biology and Medicine*, vol. 146, p. 105634, May 2022.
- [6] M. Khojaste-Sarakhs, S. S. Haghghi, S. F. Ghomi, and E. Marchiori, “Deep learning for alzheimer’s disease diagnosis: A survey,” *Artificial Intelligence in Medicine*, p. 102332, Jun 2022.
- [7] A. Alberdi, A. Aztiria, and A. Basarab, “On the early diagnosis of alzheimer’s disease from multimodal signals: A survey,” *Artificial Intelligence in Medicine*, vol. 71, p. 1–29, Jun 2016.
- [8] S. Teipel, A. Drzezga, M. J. Grothe, H. Barthel, G. Chételat, N. Schuff, P. Skudlarski, E. Cavedo, G. B. Frisoni, and W. a. Hoffmann, “Multimodal imaging in alzheimer’s disease: validity and usefulness for early detection,” *Lancet Neurology*, pp. 1037–1053, 2015.
- [9] D. Lu, K. Popuri, and D. et al., “Multimodal and multiscale deep neural networks for the early diagnosis of alzheimer’s disease using structural mr and fdg-pet images,” *Scientific Reports*, Apr 2018.
- [10] S. Liu, S. Liu, W. Cai, H. Che, S. Pujol, R. Kikinis, D. Feng, M. J. Fulham, and ADNI, “Multimodal neuroimaging feature learning for multiclass diagnosis of alzheimer’s disease,” *IEEE Transactions on Biomedical Engineering*, p. 1132–1140, Nov 2014.
- [11] M. F. Folstein, S. E. Folstein, and P. R. McHugh, ““mini-mental state”: a practical method for grading the cognitive state of patients for the clinician,” *Journal of psychiatric research*, vol. 12, no. 3, pp. 189–198, 1975.
- [12] Z. S. Nasreddine, N. A. Phillips, V. Bédirian, S. Charbonneau, V. Whitehead, I. Collin, J. L. Cummings, and H. Chertkow, “The montreal cognitive assessment, moca: a brief screening tool for mild cognitive impairment,” *Journal of the American Geriatrics Society*, vol. 53, no. 4, pp. 695–699, 2005.
- [13] W. Rosen, R. Mohs, and K. Davis, “A new rating scale for alzheimer’s disease,” *American Journal of Psychiatry*, vol. 141, p. 1356–1364, Dec 2014.
- [14] G. M. Peavy, “Neuropsychological assessment of severely demeted elderly: the severe cognitive impairment profile,” *Archives of Neurology*, vol. 53, p. 367, Nov 2011.
- [15] J. Saxton and A. A. Swihart, “Neuropsychological assessment of the severely impaired elderly patient,” *Clinics in Geriatric Medicine*, vol. 5, p. 531–543, Jul 2018.
- [16] D. R. Roalf, P. J. Moberg, S. X. Xie, D. A. Wolk, S. T. Moeller, and S. E. Arnold, “Comparative accuracies of two common screening instruments for classification of alzheimer’s disease, mild cognitive impairment, and healthy aging,” *Alzheimer’s & Dementia*, vol. 9, no. 5, pp. 529–537, 2013.

- [17] P. Robert, S. Ferris, S. Gauthier, R. Ihl, B. Winblad, and F. Tennigkeit, "Review of alzheimer's disease scales: is there a need for a new multi-domain scale for therapy evaluation in medical practice?," *Alzheimer's Research&#amp; Therapy*, vol. 2, p. 24, Aug 2010.
- [18] G. B. Frisoni, M. Boccardi, F. Barkhof, K. Blennow, S. Cappa, K. Chiotis, J.-F. Démontet, V. Garibotti, P. Giannakopoulos, A. Giett, et al., "Strategic roadmap for an early diagnosis of alzheimer's disease based on biomarkers," *The Lancet Neurology*, vol. 16, no. 8, pp. 661–676, 2017.
- [19] R. J. Perrin, A. M. Fagan, and D. M. Holtzman, "Multimodal techniques for diagnosis and prognosis of alzheimer's disease," *NATURE -LONDON-*, 2009.
- [20] T. Tapiola, I. Alafuzoff, S.-K. Herukka, L. Parkkinen, P. Hartikainen, H. Soimininen, and T. Pirttilä, "Cerebrospinal fluid β -amyloid 42 and tau proteins as biomarkers of alzheimer-type pathologic changes in the brain," *Archives of neurology*, vol. 66, no. 3, pp. 382–389, 2009.
- [21] O. Hansson, H. Zetterberg, P. Buchhave, E. Londos, K. Blennow, and L. Minthon, "Association between csf biomarkers and incipient alzheimer's disease in patients with mild cognitive impairment: a follow-up study," *The Lancet Neurology*, vol. 5, no. 3, pp. 228–234, 2006.
- [22] H. M. Snyder, M. C. Carrillo, F. Grodstein, K. Henriksen, A. Jeromin, S. Lovestone, M. M. Mielke, S. O'Bryant, M. Sarasa, M. Sjgren, H. Soares, J. Teeling, E. Trushina, M. Ward, T. West, L. J. Bain, D. W. Shineman, M. Weiner, and H. M. Fillit, "Developing novel blood-based biomarkers for alzheimer's disease," *Alzheimer's&#amp; Dementia*, vol. 10, p. 109–114, Dec 2013.
- [23] B. T. M. D. M. Fádel, R. L. S. D. Carvalho, T. T. B. A. D. Santos, and M. C. N. Dourado, "Facial expression recognition in alzheimer's disease: A systematic review," *Journal of Clinical and Experimental Neuropsychology*, vol. 41, no. 5, pp. 1–12, 2018.
- [24] L. C. Jiskoot, J. M. Poos, M. E. Vollebergh, S. Franzen, J. van Hemmen, J. M. Papma, J. C. van Swieten, R. P. C. Kessels, and E. van den Berg, "Emotion recognition of morphed facial expressions in presymptomatic and symptomatic frontotemporal dementia, and alzheimer's dementia," *Journal of Neurology*, vol. 268, p. 102–113, Jul 2020.
- [25] F. Kumfor, L.-A. Sapey-Triomphe, C. E. Leyton, J. R. Burrell, J. R. Hodges, and O. Piguet, "Degradation of emotion processing ability in corticobasal syndrome and alzheimer's disease," *Brain*, vol. 137, p. 3061–3072, Sep 2014.
- [26] R. Mc Ardle, B. Galna, P. Donaghy, A. Thomas, and L. Rochester, "Do alzheimer's and lewy body disease have discrete pathological signatures of gait?," *Alzheimer's&#amp; Dementia*, vol. 15, p. 1367–1377, Sep 2019.
- [27] L. C. Kourtis, O. B. Regele, J. M. Wright, and G. B. Jones, "Digital biomarkers for alzheimer's disease: the mobile/ wearable devices opportunity," *npj Digital Medicine*, vol. 2, Feb 2019.
- [28] M. Higuma, N. Sanjo, H. Mitoma, M. Yoneyama, and T. Yokota, "Whole-day gait monitoring in patients with alzheimer's disease: A relationship between attention and gait cycle.," *Journal of Alzheimer's Disease Reports*, vol. 1, p. 1–8, Feb 2017.
- [29] G. Cicirelli, D. Impedovo, V. Dentamaro, R. Marani, G. Pirlo, and T. R. D'Orazio, "Human gait analysis in neurodegenerative diseases: a review," *IEEE Journal of Biomedical and Health Informatics*, p. 229–242, Jun 2021.
- [30] L. Gonnerman, J. Aronoff, A. Almor, D. Kempler, and E. Andersen, "From beetle to bug: Progression of error types in naming in alzheimer's disease," Jan 2004.
- [31] S. Weintraub, A. H. Wicklund, and D. P. Salmon, "The neuropsychological profile of alzheimer disease," *Cold Spring Harbor Perspectives in Medicine*, vol. 2, p. a006171–a006171, Apr 2013.
- [32] M. Bondi, D. Salmon, and A. Kaszniak, "The neuropsychology of dementia.," Jan 1996.
- [33] S. Misiewicz, A. M. Brickman, and G. Tosto, "Prosodic impairment in dementia: Review of the literature.," *Current Alzheimer Research*, vol. 15, p. 157–163, Nov 2017.
- [34] M. L. B. Pulido, J. B. Alonso, M. A. Ferrer-Ballester, C. M. Travieso-González, J. Mekyska, and Z. Smekal, "Alzheimer's disease and automatic speech analysis: A review," *Expert Syst. Appl.*, vol. 150, p. 113213, 2020.
- [35] C. Poellabauer, N. Yadav, L. Daudet, S. L. Schneider, C. Busso, and P. J. Flynn, "Challenges in concussion detection using vocal acoustic biomarkers," *IEEE Access*, vol. 3, pp. 1143–1160, 2015.
- [36] F. Haider, S. Garcia, and S. Luz, "An assessment of paralinguistic acoustic features for detection of alzheimer's dementia in spontaneous speech," *IEEE Journal of Selected Topics in Signal Processing*, vol. 14, no. 2, pp. 272–281, 2020.
- [37] F. Eyben and S. et al., "The geneva minimalistic acoustic parameter set (gemaps) for voice research and affective computing," *IEEE Transactions on Affective Computing*, vol. 7, p. 190–202, Jul 2015.
- [38] F. Eyben, M. Wllmer, and B. Schuller, "Opensmile: the munich versatile and fast open-source audio feature extractor.," in *AcM International Conference on Multimedia*, 2010.
- [39] F. Eyben, F. Weninger, F. Groß, and B. Schuller, "Recent developments in opensmile, the munich open-source multimedia feature extractor," *Proceedings of the 21st ACM international conference on Multimedia*, 2013.
- [40] L. Liu, S. Zhao, H. Chen, and A. Wang, "A new machine learning method for identifying alzheimer's disease," *Simul. Model. Pract. Theory*, vol. 99, 2020.
- [41] T. Zhou, M. Liu, K.-H. Thung, and D. Shen, "Latent representation learning for alzheimer's disease diagnosis with incomplete multimodality neuroimaging and genetic data," *IEEE Transactions on Medical Imaging*, vol. 38, p. 2411–2422, Apr 2019.
- [42] S. M. Landau, D. Harvey, and C. M. a. Madison, "Comparing predictors of conversion and decline in mild cognitive impairment," *Neurology*, vol. 75, p. 230–238, Jul 2010.
- [43] T. Zhou, K.-H. Thung, X. Zhu, and D. Shen, "Effective feature learning and fusion of multimodality data using stage-wise deep neural network for dementia diagnosis," *Human brain mapping*, vol. 40, no. 3, pp. 1001–1016, 2019.
- [44] X.-a. Bi, X. Hu, H. Wu, and Y. Wang, "Multimodal data analysis of alzheimer's disease based on clustering evolutionary random forest," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, p. 2973–2983, Feb 2020.
- [45] D. Zhang, Y. Wang, L. Zhou, H. Yuan, and D. Shen, "Multimodal classification of alzheimer's disease and mild cognitive impairment," *NeuroImage*, p. 856–867, Apr 2011.
- [46] J. Venugopalan, L. Tong, H. R. Hassanzadeh, and M. D. Wang, "Multimodal deep learning models for early detection of alzheimer's disease stage.," *Scientific Reports*, Feb 2021.
- [47] P. S. Aisen, R. C. Petersen, M. C. Donohue, A. Gamst, R. Raman, R. G. Thomas, S. Walter, J. Q. Trojanowski, L. M. Shaw, and L. A. a. Beckett, "Clinical core of the alzheimer's disease neuroimaging initiative: progress and plans.," *Alzheimer's & dementia: the journal of the Alzheimer's Association*, no. 3, p. 6, 2010.
- [48] B. Dubois, H. Hampel, H. H. Feldman, P. Scheltens, P. Aisen, S. Andrieu, H. Bakardjian, H. Benali, L. Bertram, K. Blennow, K. Broich, E. Cavedo, S. Crutch, J.-F. Dartigues, C. Duyckaerts, S. Epelbaum, G. B. Frisoni, S. Gauthier, R. Genton, A. A. Gouw, M.-O. Habert, D. M. Holtzman, M. Kivipelto, S. Lista, J.-L. Molinuevo, S. E. O'Bryant, G. D. Rabinovici, C. Rowe, S. Salloway, L. S. Schneider, R. Sperling, M. Teichmann, M. C. Carrillo, J. Cummings, and C. R. Jack, "Preclinical alzheimer's disease: Definition, natural history, and diagnostic criteria," *Alzheimer's & Dementia*, vol. 12, no. 3, 2016.
- [49] K. Jellinger, "Morphology of alzheimer's disease and related disorders," *Journal of Neural Transmission - Parkinson's Disease and Dementia Section*, vol. 1, no. 1, pp. 11–11, 1989.
- [50] R. Cappai and A. R. White, "Amyloid β ," *The International Journal of Biochemistry & Cell Biology*, vol. 31, no. 9, pp. 885–889, 1999.
- [51] C. R. Jack, M. A. Bernstein, N. C. Fox, P. Thompson, and M. W. Weiner, "The alzheimer's disease neuroimaging initiative (adni): Mri methods.," *Journal of Magnetic Resonance Imaging*, vol. 27, no. 4, pp. 685–691, 2010.
- [52] M. W. Weiner, D. P. Veitch, P. S. Aisen, L. A. Beckett, N. J. Cairns, R. C. Green, D. Harvey, C. R. Jack, W. Jagust, E. Liu, J. C. Morris, R. C. Petersen, A. J. Saykin, M. E. Schmidt, L. Shaw, L. Shen, J. A. Siuciak, H. Soares, A. W. Toga, and J. Q. Trojanowski, "The alzheimer's disease neuroimaging initiative: A review of papers published since its inception," *Alzheimer's & Dementia*, 2013.
- [53] C. R. J. A, M. A. B. A, B. J. B. A, J. L. G. A, N. C. F. B, P. M. T. C, N. S. D, G. K. E, R. J. K. F, and D. C. a. G, "Update on the magnetic resonance imaging core of the alzheimer's disease neuroimaging initiative," *Alzheimer's & Dementia*, vol. 6, no. 3, pp. 212–220, 2010.
- [54] A. D. Leow, S. Zhu, K. Memahon, G. I. D. Zubicaray, M. Meredith, M. Wright, and A. P. M. Thompson, "The tensor distribution function," in *IEEE International Symposium on Biomedical Imaging: from Nano to Macro*, 2009.
- [55] T. M. Nir, N. Jahanshad, J. E. Villalon-Reina, A. W. Toga, C. R. Jack, M. W. Weiner, and P. M. Thompson, "Effectiveness of regional dti measures in distinguishing alzheimer's disease, mci, and normal aging," *NeuroImage Clinical*, vol. 3, pp. 180–195, 2013.
- [56] C. Zarow, H. V. Vinters, W. G. Ellis, M. W. Weiner, and H. C. Chui, "Correlates of hippocampal neuron number in alzheimer's disease and

- ischemic vascular dementia," *Annals of Neurology*, vol. 57, no. 6, pp. 896–903, 2010.
- [57] R. J. Killiany, T. Gomez-Isla, M. Moss, R. Kikinis, T. Sandor, F. Jolesz, R. Tanzi, K. Jones, B. T. Hyman, and M. S. Albert, "Use of structural magnetic resonance imaging to predict who will get alzheimer's disease," *Annals of Neurology*, vol. 47, no. 4, pp. 430–439, 2010.
- [58] E. G. Kehoe, J. P. McNulty, P. G. Mullins, and A. L. W. Bokde, "Advances in mri biomarkers for the diagnosis of alzheimer's disease," *Biomarkers in medicine*, vol. 8 9, pp. 1151–69, 2014.
- [59] D. S. Karow, L. K. McEvoy, C. Fennema-Notestine, D. J. Hagler, and A. M. Dale, "Relative capability of mr imaging and fdg pet to depict changes associated with prodromal and early alzheimer disease," *Radiology*, vol. 256, no. 3, pp. 932–942, 2010.
- [60] M. D. Greicius, G. Srivastava, A. L. Reiss, and V. Menon, "Default-mode network activity distinguishes alzheimer's disease from healthy aging: Evidence from functional mri," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 13, pp. 4637–4642, 2004.
- [61] V. L. Villemagne and G. Chételat, "Neuroimaging biomarkers in alzheimer's disease and other dementias," *Ageing Research Reviews*, pp. 4–16, 2016.
- [62] M. Meyer-Lindemann, T. L. Spires-Jones, C. Prada, M. Garcia-Alloza, A. D. Calignon, A. Rozkalne, J. Koenigsknecht-Talbot, D. M. Holtzman, B. J. Bacska, and B. T. Hyman, "Rapid appearance and local toxicity of amyloid- β plaques in a mouse model of alzheimer's disease," *Nature*, vol. 451, no. 7179, pp. p.720–724, 2008.
- [63] R. O. S. M. H. H. f. t. A. D. N. I. Bárbara Avelar-Pereira, Michael E. Belley, "Decoding the heterogeneity of alzheimer's disease diagnosis and progression using multilayer networks," *molecular psychiatry*, no. 3, p. 6, 2022.
- [64] B. Olsson, R. Lautner, U. Andreasson, A. ?hrfelt, E. Portelius, M. Bjerke, M. H?Ltt?, C. Rosén, C. Olsson, and G. a. Strobel, "Csf and blood biomarkers for the diagnosis of alzheimer's disease: a systematic review and meta-analysis," *Lancet Neurology*, pp. 673–684, 2016.
- [65] M. A. Ebrahimiaghahnavieh and R. Chiong, "Deep learning to detect alzheimer's disease from neuroimaging: A systematic literature review," *Computer Methods and Programs in Biomedicine*, vol. 187, p. 105242, 2019.
- [66] A. Farooq, S. M. Anwar, M. Awais, and S. Rehman, "A deep cnn based multi-class classification of alzheimer's disease using mri," in *2017 IEEE International Conference on Imaging Systems and Techniques (IST)*, 2017.
- [67] S. Wang, Q. Zhou, M. Yang, and Y. D. Zhang, "Advian: Alzheimer's disease vgg-inspired attention network based on convolutional block attention module and multiple way data augmentation," *Frontiers in Aging Neuroscience*, vol. 13, p. 687456, 2021.
- [68] A. Ebrahimi, S. Luo, R. Chiong, and A. D. N. Initi, "Deep sequence modelling for alzheimer's disease detection using mri," *COMPUTERS IN BIOLOGY AND MEDICINE*, vol. 134, JUL 2021.
- [69] T. Zhang and C. Duan, "Multi-stream deep convolutional neural network for pet preform surface defects detection and classification," *IEEE Access*, 2021.
- [70] F. Martinez-Murcia, A. Ortiz, J. Gorri?z, J. Ramírez, and D. Castillo-Barnes, "Studying the manifold structure of alzheimer's disease: A deep learning approach using convolutional autoencoders," *IEEE journal of biomedical and health informatics*, vol. 24, no. 1, pp. 17–26, 2020.
- [71] S. Rathore, M. Habes, M. Aksam Iftikhar, A. Shacklett, and C. Davatzikos, "A review on neuroimaging-based classification studies and associated feature extraction methods for alzheimer's disease and its prodromal stages," *Neuroimage*, p. 530, 2017.
- [72] Q. Shangran, P. S. Joshi, M. I. Miller, X. Chonghua, Z. Xiao, K. Cody, G. H. Chang, A. S. Joshi, D. Brigid, and Z. a. Shuhan, "Development and validation of an interpretable deep learning framework for alzheimer's disease classification," *Brain*, no. 6, p. 6, 2020.
- [73] C. Lian, M. Liu, J. Zhang, and D. Shen, "Hierarchical fully convolutional network for joint atrophy localization and alzheimer's disease diagnosis using structural mri," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1–1, 2018.
- [74] A. Payan and G. Montana, "Predicting alzheimer's disease: a neuroimaging study with 3d convolutional neural networks," 2015.
- [75] H. Li, M. Habes, and Y. Fan, "Deep ordinal ranking for multi-category diagnosis of alzheimer's disease using hippocampal mri data," *Computer Vision and Pattern Recognition*, 2017.
- [76] F. J. Martinez-Murcia, J. M. Górriz, Javier.Ramírez, and A. Ortiz, "Convolutional neural networks for neuroimaging in parkinson's disease: Is preprocessing needed?," *International journal of neural systems*, vol. 28, no. 10, p. 1850035, 2018.
- [77] Ruoxuan, Cui, Manhua, and Liu, "Hippocampus analysis by combination of 3d densenet and shapes for alzheimer's disease diagnosis," *IEEE Journal of Biomedical & Health Informatics*, 2018.
- [78] Geert, Litjens, Thijs, Kooi, Babak, Ehteshami, Bejnordi, Arnaud, Arindra, and Adiyoso, "A survey on deep learning in medical image analysis," *Medical Image Analysis*, 2017.
- [79] W. Zhu, L. Sun, J. Huang, L. Han, and D. Zhang, "Dual attention multi-instance deep learning for alzheimer's disease diagnosis with structural mri," *IEEE Transactions on Medical Imaging*, vol. PP, no. 99, pp. 1–1, 2021.
- [80] F. J. Martinez-Murcia, J. M. Górriz, J. Ramírez, F. Segovia, D. Salas-Gonzalez, D. Castillo-Barnes, and A. Ortiz, "Assessing mild cognitive impairment progression using a spherical brain mapping of magnetic resonance imaging," *Journal of Alzheimer's Disease*, pp. 1–17, 2018.
- [81] I. A. Illan, J. M. G?3Rriz, R. R. Javier, and M. B. Anke, "Spatial component analysis of mri data for alzheimer's disease diagnosis: a bayesian network approach," *Frontiers in Computational Neuroscience*, vol. 8, no. 156, p. 156, 2014.
- [82] P. J. Markiewicz, J. C. Matthews, J. Declerck, and K. Herholz, "Robustness of multivariate image analysis assessed by resampling techniques and applied to fdg-pet scans of patients with alzheimer's disease," *Neuroimage*, vol. 46, no. 2, pp. 472–485, 2009.
- [83] Rolf, A., Heckemann, , , Shiva, Keihaninejad, , , Paul, Aljabar, and and, "Improving intersubject image registration using tissue-class information benefits robustness and accuracy of multi-atlas based anatomical segmentation - sciencedirect," *NeuroImage*, vol. 51, no. 1, pp. 221–227, 2010.
- [84] S. Liu, W. Cai, L. Wen, and D. Feng, "Neuroimaging biomarker based prediction of alzheimer's disease severity with optimized graph construction," in *2013 IEEE 10th International Symposium on Biomedical Imaging*, 2013.
- [85] N. Zeng, H. Li, and Y. Peng, "A new deep belief network-based multi-task learning for diagnosis of alzheimer's disease," *Neural Computing and Applications*, no. 12, 2021.
- [86] E. Lee, J.-S. Choi, M. Kim, H.-I. Suk, A. D. N. Initiative, et al., "Toward an interpretable alzheimer's disease diagnostic model with regional abnormality representation via deep learning," *Neuroimage*, vol. 202, p. 116113, 2019.
- [87] K. Blennow, H. Zetterberg, L. Minthon, L. Lannfelt, S. Strid, P. Annas, H. Basun, and N. Andreasen, "Longitudinal stability of csf biomarkers in alzheimer's disease," *Neuroscience Letters*, vol. 419, no. 1, pp. 18–22, 2007.
- [88] R. Motter, C. Vigo-Pelfrey, D. Kholodenko, R. Barbour, K. Johnson-Wood, D. Galasko, L. Chang, B. Miller, C. Clark, and R. Green, "Reduction of β -amyloid peptide42in the cerebrospinal fluid of patients with alzheimer's disease," *Annals of neurology*, 1995.
- [89] Mawuenyega, Kwasi, G., Sigurdson, Wendy, Ovod, Vitaliy, Munsell, Ling, and K. and, "Decreased clearance of cns β -amyloid in alzheimer's disease..," *Science*, 2010.
- [90] A. M. Fagan, M. A. Mintun, R. H. Mach, S. Y. Lee, C. S. Dence, A. R. Shah, G. N. Larossa, M. L. Spinner, W. E. Klunk, and C. A. Mathis, "Inverse relation between in vivo amyloid imaging load and cerebrospinal fluid a β 42 in humans," *Annals of Neurology*, 2006.
- [91] S. H. Freeman, R. Susan, B. T. Hyman, M. P. Frosch, and M. C. Irizarry, "Plasma a β levels do not reflect brain a β levels," *Journal of Neuropathology and Experimental Neurology*, no. 4, pp. 264–271, 2007.
- [92] H. Vanderstichele, E. V. Kerschaver, C. Hesse, P. Davidsson, M. A. Buyse, N. Andreasen, L. Minthon, A. Wallin, K. Blennow, and E. Vanmechelen, "Standardization of measurement of beta-amyloid(1-42) in cerebrospinal fluid and plasma," *Amyloid : the international journal of experimental and clinical investigation : the official journal of the International Society of Amyloidosis*, vol. 7, no. 4, pp. 245–58, 2000.
- [93] T. Sobów, M. Flirski, I. Koszewska, and P. P. Liberski, "Plasma levels of abeta peptides are altered in amnestic mild cognitive impairment but not in sporadic alzheimer's disease," *Acta neurobiologiae experimentalis*, vol. 65, no. 65, pp. 117–124, 2005.
- [94] R. Mayeux, L. S. Honig, M. X. Tang, J. Manly, Y. Stern, N. Schupf, and P. D. Mehta, "Plasma a[math]\beta40 and a[math]\beta42 and alzheimer's disease: relation to age, mortality, and risk.,," *Neurology*, vol. 61, no. 9, pp. 1185–1190, 2003.
- [95] M. Pesaresi, C. Lovati, P. Bertora, E. Maillard, and C. Mariani, "Plasma levels of beta-amyloid (1-42) in alzheimer's disease and mild cognitive impairment," *Neurobiology of Aging*, vol. 27, no. 6, pp. 904–905, 2006.
- [96] Richard, Mayeux, Ming-Xin, Tang, Diane, M., Jacobs, Jennifer, Manly,

- and Karen, "Plasma amyloid β -peptide 1–42 and incipient alzheimer's disease," *Annals of Neurology*, 1999.
- [97] A. Nakamura, N. Kaneko, V. L. Villemagne, T. Kato, J. Doecke, V. Doré, C. Fowler, Q.-X. Li, R. Martins, C. Rowe, T. Tomita, K. Matsuzaki, K. Ishii, K. Ishii, Y. Arahat, S. Iwamoto, K. Ito, K. Tanaka, and C. L. M. . K. Yanagisawa, "High performance plasma amyloid-beta biomarkers for alzheimer's disease," *Nature*, 2018.
- [98] P. C. E. T. P. a Person Envelope, I. M. W. V. A., E. H. T. A., L. V. A., P. O. H. D. E. P. H. Z. P. f g h i j, P. W. M. van der Flier PhD b, P. M. M. M. K. and M. D. C. A. C, "Blood-based biomarkers for alzheimer's disease: towards clinical implementation," *Lancet Neurol*, 2022.
- [99] O. Hansson, H. Zetterberg, E. Vanmechelen, H. Vanderstichele, and K. Blennow, "Evaluation of plasma $\alpha\beta$ 40 and $\alpha\beta$ 42 as predictors of conversion to alzheimer's disease in patients with mild cognitive impairment," *Neurobiology of aging*, vol. 31, no. 3, pp. 357–367, 2008.
- [100] I. M. W. V. MSc, R. E. S. MD, S. C. J. V. PhD, H. H. BSc, N. D. P. MD, B. N. M. van Berckel MD, P. S. MD, C. E. Teunissen, and W. M. van der Flier PhD, "Plasma amyloid as prescreener for the earliest alzheimer pathological changes," *Annals of Neurology*, 2018.
- [101] T. West, K. M. Kirmess, M. R. Meyer, M. S. Holubasch, S. S. Knapik, Y. Hu, J. H. Contois, E. N. Jackson, S. E. Harpstrite, and R. J. a. Bateeman, "A blood-based diagnostic test incorporating plasma $\alpha\beta$ 42/40 ratio, apoe proteotype, and age accurately identifies brain amyloid status: findings from a multi cohort validity analysis," *Molecular Neurodegeneration*, no. 1, 2021.
- [102] Y. Yamazaki, N. Zhao, T. R. Caulfield, C. C. Liu, and G. Bu, "Apolipoprotein e and alzheimer disease: pathobiology and targeting strategies," *Nature Reviews Neurology*, 2019.
- [103] P. M. Staff, "Genetic assessment of age-associated alzheimer disease risk: Development and validation of a polygenic hazard score," *Plos Medicine*, vol. 14, no. 3, p. e1002258, 2017.
- [104] D. Z. Christensen, T. Schneider-Axmann, P. J. Lucassen, T. A. Bayer, and O. Wirths, "Accumulation of intraneuronal $\alpha\beta$ correlates with apoe4 genotype," *Acta Neuropathologica*, vol. 119, no. 5, pp. 555–566, 2010.
- [105] J. C. J. Morris, C. M. C. Roe, C. C. Xiong, A. M. A. Fagan, A. M. A. Goate, D. M. D. Holtzman, and M. A. M. Mintun, "Apoe predicts amyloid-beta but not tau alzheimer pathology in cognitively normal aging," *Annals of Neurology*, vol. 67, no. 1, p. 122, 2010.
- [106] J. Zhao, M. D. Davis, Y. Atagi, M. Shinohara, and G. Bu, "Apoe ϵ 4/ ϵ 4 diminishes neurotrophic function of human ipsc-derived astrocytes," *Human Molecular Genetics*, vol. 26, no. 14, p. 2690, 2017.
- [107] T. Arendt, C. Schindler, M. K. B. ckner, K. Eschrich, and V. Bigl, "Plastic neuronal remodeling is impaired in patients with alzheimer's disease carrying apolipoprotein 4 allele," *THEJOURNAL OF NEUROSCIENCE*, 1997.
- [108] N. Zhao, C. C. Liu, A. J. Van Ingelgom, Y. A. Martens, C. Linares, J. A. Knight, M. M. Painter, P. M. Sullivan, and G. Bu, "Apolipoprotein e4 impairs neuronal insulin signaling by trapping insulin receptor in the endosomes," *Neuron*, vol. 96, no. 1, p. 115–129.e5, 2017.
- [109] J. H. Sun, L. Tan, H. F. Wang, M. S. Tan, L. Tan, J. Q. Li, W. Xu, X. C. Zhu, T. Jiang, and J. T. Yu, "Genetics of vascular dementia: Systematic review and meta-analysis," *Journal of Alzheimer's Disease*, 2015.
- [110] S. Venneti, B. J. Lopresti, and C. A. Wiley, "The peripheral benzodiazepine receptor (translocator protein 18kda) in microglia: from pathology to imaging," *Progress in Neurobiology*, vol. 80, no. 6, pp. 308–322, 2006.
- [111] T. J. Montine, J. Quinn, J. Kaye, and J. D. Morrow, "F2-isoprostanes as biomarkers of late-onset alzheimer's disease," *Journal of Molecular Neuroscience Mn*, vol. 33, no. 1, p. 114, 2007.
- [112] K. Gressie, F. Kumfor, H. Teng, D. Foxe, E. Devenney, R. M. Ahmed, and O. Piguet, "Error profiles of facial emotion recognition in frontotemporal dementia and alzheimer's disease," *International Psychogeriatrics*, p. 1–10, 2023.
- [113] S. Heilman, K.M.and Nadeau, "Emotional and neuropsychiatric disorders associated with alzheimer's disease," *Neurotherapeutics*, vol. 19, no. 99, p. 116, 2022.
- [114] Shirley, A., Radford, , Romola, S., and Bucks, "Emotion processing in alzheimer's disease," *Neurobiology of Aging*, 2000.
- [115] V. E. Sturm, J. S. Yokoyama, W. W. Seeley, J. H. Kramer, and K. P. Rankin, "Heightened emotional contagion in mild cognitive impairment and alzheimer's disease is associated with temporal lobe degeneration," *Pnas*, vol. 110, no. 24, pp. 9944–9949, 2013.
- [116] K. W. Burton and A. W. Kaszniak, "Emotional experience and facial expression in alzheimer's disease," *Aging, Neuropsychology, and Cognition*, 2007.
- [117] Rosie, Morris, Sue, Lord, Jennifer, Bunce, David, Burn, Lynn, and Rochester, "Gait and cognition: Mapping the global and discrete relationships in ageing and neurodegenerative disease," *Neuroscience & Biobehavioral Reviews*, 2016.
- [118] J. Verghese, C. Wang, R. B. Lipton, R. Holtzer, and X. Xue, "Quantitative gait dysfunction and risk of cognitive decline and dementia," *Journal of Neurology Neurosurgery & Psychiatry*, vol. 78, no. 9, p. 929, 2007.
- [119] A. Rma, B. Bga, A. Pd, A. At, and C. Lra, "Do alzheimer's and lewy body disease have discrete pathological signatures of gait? - sciencedirect," *Alzheimer's & Dementia*, vol. 15, no. 10, pp. 1367–1377, 2019.
- [120] O. Beauchet, C. Annweiler, M. L. Callisaya, A.-M. De Cock, J. L. Helbostad, R. W. Kressig, V. Srikanth, J.-P. Steinmetz, H. M. Blumen, J. Verghese, and G. Allali, "Poor gait performance and prediction of dementia: Results from a meta-analysis," *Journal of the American Medical Directors Association*, vol. 17, no. 6, pp. 482–490, 2016.
- [121] J. G. Nutt, "Higher-level gait disorders: An open frontier," *Movement Disorders Official Journal of the Movement Disorder Society*, vol. 28, no. 11, pp. 1560–1565, 2013.
- [122] Fiona, Kumfor, Laurie-Anne, Sapey-Triomphe, Cristian, E., Leyton, James, R., and Burrell, "Degradation of emotion processing ability in corticobasal syndrome and alzheimer's disease," *Brain*, 2014.
- [123] L. C. Jiskoot, J. M. Poos, M. E. Vollebergh, S. Franzen, J. V. Hemmen, J. M. Papma, J. C. V. Swieten, R. P. C. Kessels, and E. V. D. Berg, "Emotion recognition of morphed facial expressions in presymptomatic and symptomatic frontotemporal dementia, and alzheimer's dementia," *Journal of neurology*, vol. 268, no. 1, pp. 102–113, 2021.
- [124] D. C. Mograbi, R. G. Brown, C. Salas, and R. G. Morris, "Emotional reactivity and awareness of task performance in alzheimer's disease," *Neuropsychologia*, vol. 50, no. 8, 2012.
- [125] M. A. Case, H. A. Burwick, K. G. Volpp, and M. S. Patel, "Accuracy of smartphone applications and wearable devices for tracking physical activity data," *Journal of the American Medical Association*, vol. 313, no. 6, pp. 625–6, 2015.
- [126] L. Piwek, D. A. Ellis, S. Andrews, and A. Joinson, "The rise of consumer health wearables: Promises and barriers," *PLoS Medicine*, vol. 13, no. 2, p. e1001953, 2016.
- [127] M. Lindh-Rengifo, S. B. Jonasson, S. Ullén, E. Stomrud, S. Palmqvist, N. Mattsson-Carlgren, O. Hansson, and M. H. Nilsson, "Components of gait in people with and without mild cognitive impairment," *Gait & Posture*, vol. 93, pp. 83–89, 2022.
- [128] R. Rucco, V. Agosti, F. Jacini, P. Sorrentino, P. Varriale, M. De Stefan, G. Milan, P. Montella, and G. Sorrentino, "Spatio-temporal and kinematic gait analysis in patients with frontotemporal dementia and alzheimer's disease through 3d motion capture," *Gait & Posture*, vol. 52, no. Complete, pp. 312–317, 2017.
- [129] E. L. E. Ekman, Paul (Ed) Rosenberg, "What the face reveals: Basic and applied studies of spontaneous expression using the facial action coding system (facs) (2nd ed)," *Oxford University Press*, 2005.
- [130] A. T. Lopes, E. D. Aguiar, A. Souza, and T. Oliveira-Santos, "Facial expression recognition with convolutional neural networks: Coping with few data and the training sample order," *Pattern Recognition*, vol. 61, pp. 610–628, 2017.
- [131] B. Ghoraani, L. N. Boettcher, M. D. Hssayeni, A. Rosenfeld, M. I. Tolea, and J. E. Galvin, "Detection of mild cognitive impairment and alzheimer's disease using dual-task gait assessments and machine learning," *Biomedical signal processing and control*, vol. 64, p. 102249, 2021.
- [132] S. D. Din, A. Godfrey, and L. Rochester, "Validation of an accelerometer to quantify a comprehensive battery of gait characteristics in healthy older adults and parkinson's disease: Toward clinical and at home use," *IEEE Journal of Biomedical and Health Informatics*, 2015.
- [133] A. Smgb and B. Pg, "A robust, cost-effective and non-invasive computer-aided method for diagnosis three types of neurodegenerative diseases with gait signal analysis," *Measurement*, vol. 156, 2020.
- [134] J. Meilan, F. Martinez-Sanchez, J. Carro, N. Carcavilla, and O. Ivanova, "Voice markers of lexical access in mild cognitive impairment and alzheimer's disease," *Current Alzheimer Research*, vol. 14, no. 2, pp. 111–119, 2017.
- [135] Mirheidari, Bahman, Harkness, Kirsty, Reuber, Markus, Christensen, Heidi, Blackburn, and Daniel, "Toward the automation of diagnostic conversation analysis in patients with memory complaints," *Journal of Alzheimers Disease Jad*, 2017.
- [136] S. Greta, H. Ildiko, V. Veronika, K. Janos, and P. Magdolna, "Speaking in alzheimer's disease, is that an early sign? importance of changes in

- language abilities in alzheimer's disease," *Frontiers in Aging Neuroscience*, vol. 7, 2015.
- [137] S. Ahmed, A. M. F. Haigh, C. A. D. Jager, and P. Garrard, "Connected speech as a marker of disease progression in autopsy-proven alzheimer's disease," *Brain*, 2013.
- [138] M. F. Mendez, D. G. Clark, J. S. Shapira, and J. L. Cummings, "Speech and language in progressive nonfluent aphasia compared with early alzheimer's disease," *Neurology*, vol. 61, no. 8, p. 1108, 2003.
- [139] U. Petti, S. Baker, and A. Korhonen, "A systematic literature review of automatic alzheimer's disease detection from speech and language," *Journal of the American Medical Informatics Association*, 2020.
- [140] A. Karakostas, A. Briassoulis, K. Avgerinakis, I. Kompatiari, and M. Tsolaki, "The dem@care experiments and datasets: a technical report," *arXiv preprint arXiv:1701.01142*, 2016.
- [141] J. T. Becker, F. Boiler, O. L. Lopez, J. Saxton, and K. L. Mcgonigle, "The natural history of alzheimer's disease: Description of study cohort and accuracy of diagnosis," *Archives of Neurology*, vol. 51, no. 6, pp. 585–94, 1994.
- [142] S. Luz, F. Haider, S. de la Fuente, D. Fromm, and B. MacWhinney, "Alzheimer's dementia recognition through spontaneous speech: The address challenge," *arXiv preprint arXiv:2004.06833*, 2020.
- [143] Fraser, Kathleen, C. Rudzicz, Frank, Meltzer, Jed, and A., "Linguistic features identify alzheimer's disease in narrative speech," *Journal of Alzheimer's disease: JAD*, vol. 49, no. 2, pp. 407–422, 2016.
- [144] J. Chen, Y. Wang, and D. L. Wang, "A feature study for classification-based speech separation at low signal-to-noise ratios," *IEEE Press*, 2014.
- [145] J. B. Alonso, J. Cabrera, M. Medina, and C. M. Travieso, "New approach in quantification of emotional intensity from the speech signal: emotional temperature," *Expert Systems with Applications*, vol. 42, no. 24, pp. 9554–9564, 2015.
- [146] H. W. Kim, H. E. Lee, S. Lee, K. T. Oh, and S. K. Yoo, "Slice-selective learning for alzheimer's disease classification using a generative adversarial network: a feasibility study of external validation," *European journal of nuclear medicine and molecular imaging*, 2020.
- [147] A. Ci, B. Nm, C. Ah, and A. Fcm, "A novel multi-modal machine learning based approach for automatic classification of eeg recordings in dementia - sciencedirect," *Neural Networks*, vol. 123, pp. 176–190, 2020.
- [148] F. Ahmad, H. Zulifqar, and T. Malik, "Classification of alzheimer disease among susceptible brain regions," *International Journal of Imaging Systems and Technology*, vol. 29, no. 3, pp. 222–233, 2019.
- [149] Y. Gupta, K. H. Lee, K. Y. Choi, J. J. Lee, B. C. Kim, and G. R. Kwon, "Alzheimer's disease diagnosis based on cortical and subcortical features," *Journal of Healthcare Engineering*, vol. 2019, pp. 1–13, 2019.
- [150] 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 healthy elderly, mci, cmci and alzheimer's disease patients: from the alzheimer's disease neuroimaging initiative (adni) database," *J Neurosci Methods*, pp. 14–23, 2017.
- [151] Y. Farouk and S. Rady, "Supervised classification techniques for identifying alzheimer's disease," in *International Conference on Advanced Intelligent Systems and Informatics*, 2019.
- [152] S. Sarraf, D. D. Desouza, J. A. E. Anderson, and G. Tofighi, "Deepad: Alzheimer's disease classification via deep convolutional neural networks using mri and fmri," *Cold Spring Harbor Laboratory*, 2016.
- [153] S. Afzal, M. Maqsood, F. Nazir, U. Khan, F. Aadil, K. M. Awan, I. Mahmood, and O.-Y. Song, "A data augmentation-based framework to handle class imbalance problem for alzheimer's stage detection," *IEEE access*, vol. 7, pp. 115528–115539, 2019.
- [154] R. Jain, N. Jain, A. Aggarwal, and D. J. Hemanth, "Convolutional neural network based alzheimer's disease classification from magnetic resonance brain images," *Cognitive Systems Research*, vol. 57, no. OCT., pp. 147–159, 2019.
- [155] H. Ji, Z. Liu, W. Yan, and R. Klette, "Early diagnosis of alzheimer's disease using deep learning," in *the 2nd International Conference*, 2019.
- [156] M. Kompanek, M. Tamajka, and W. Benesova, "Volumetric data augmentation as an effective tool in mri classification using 3d convolutional neural network," in *2019 International Conference on Systems, Signals and Image Processing (IWSSIP)*, 2019.
- [157] H. Barbaroux, X. Feng, J. Yang, A. F. Laine, and E. D. Angelini, "Encoding human cortex using spherical cnns - a study on alzheimer's disease classification," in *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, 2020.
- [158] A. B. Tufail, Q. N. Zhang, and Y. K. Ma, "Binary classification of alzheimer disease using smri imaging modality and deep learning," *Journal of Digital Imaging*, 2018.
- [159] R. C. Suganthe, R. S. Latha, M. Geetha, and G. R. Sreekanth, "Diagnosis of alzheimer's disease from brain magnetic resonance imaging images using deep learning algorithms," *Advances in Electrical and Computer Engineering*, vol. 20, no. 3, pp. 57–64, 2020.
- [160] M. Böhle, F. Eitel, M. Weygandt, and K. Ritter, "Layer-wise relevance propagation for explaining deep neural network decisions in mri-based alzheimer's disease classification," *Frontiers in Aging Neuroscience*, vol. 11, pp. 194–, 2019.
- [161] E. Westman, A. Simmons, Y. Zhang, J.-S. Muehlboeck, C. Tunnard, Y. Liu, L. Collins, A. Evans, P. Mecocci, B. Vellas, M. Tsolaki, I. Kloszewska, H. Soininen, S. Lovestone, C. Spenger, L.-O. Wahlund, and A. Consortium, "Multivariate analysis of mri data for alzheimer's disease, mild cognitive impairment and healthy controls," *NEUROIMAGE*, vol. 54, pp. 1178–1187, JAN 15 2011.
- [162] J. H. Morra, Z. Tu, L. G. Apostolova, A. E. Green, C. Avedissian, S. K. Madsen, N. Parikhshak, X. Hua, A. W. Toga, C. R. Jack, Jr., M. W. Weiner, P. M. Thompson, and A. D. Neuroimaging, "Validation of a fully automated 3d hippocampal segmentation method using subjects with alzheimer's disease mild cognitive impairment, and elderly controls," *NEUROIMAGE*, vol. 43, pp. 59–68, OCT 15 2008.
- [163] C.-F. Liu, S. Padhy, S. Ramachandran, V. X. Wang, A. Efimov, A. Bernal, L. Shi, M. Vaillant, J. T. Ratnanather, A. V. Faria, et al., "Using deep siamese neural networks for detection of brain asymmetries associated with alzheimer's disease and mild cognitive impairment," *Magnetic resonance imaging*, vol. 64, pp. 190–199, 2019.
- [164] E. Lee, J.-S. Choi, M. Kim, H.-I. Suk, and A. D. N. Initiia, "Toward an interpretable alzheimer's disease diagnostic model with regional abnormality representation via deep learning," *NEUROIMAGE*, vol. 202, NOV 15 2019.
- [165] R. Cui, M. Liu, and A. D. N. Initiia, "Rnn-based longitudinal analysis for diagnosis of alzheimer's disease," *COMPUTERIZED MEDICAL IMAGING AND GRAPHICS*, vol. 73, pp. 1–10, APR 2019.
- [166] C. Ge, Q. Qu, I. Y.-H. Gu, and A. S. Jakola, "Multi-stream multi-scale deep convolutional networks for alzheimer's disease detection using mr images," *Neurocomputing*, vol. 350, pp. 60–69, 2019.
- [167] R. Cui and M. Liu, "Hippocampus analysis by combination of 3-d densenet and shapes for alzheimer's disease diagnosis," *IEEE journal of biomedical and health informatics*, vol. 23, no. 5, pp. 2099–2107, 2018.
- [168] C. Ieracitano, N. Mammone, A. Bramanti, A. Hussain, and F. C. Morabito, "A convolutional neural network approach for classification of dementia stages based on 2d-spectral representation of eeg recordings," *Neurocomputing*, vol. 323, pp. 96–107, 2019.
- [169] W. Li, X. Lin, and X. Chen, "Detecting alzheimer's disease based on 4d fmri: An exploration under deep learning framework," *Neurocomputing*, vol. 388, pp. 280–287, 2020.
- [170] D. Chitra Devi and S. Prabha, "Analysis of brain sub regions using optimization techniques and deep learning method in alzheimer disease," *Applied Soft Computing*, vol. 86, p. 105857, 2020.
- [171] A. Puente-Castro, E. Fernandez-Blanco, A. Pazos, and C. R. Munteanu, "Automatic assessment of alzheimer's disease diagnosis based on deep learning techniques," *Computers in biology and medicine*, vol. 120, p. 103764, 2020.
- [172] Y. Chen and Y. Xia, "Iterative sparse and deep learning for accurate diagnosis of alzheimer's disease," *Pattern Recognition*, vol. 116, p. 107944, 2021.
- [173] R. A. Hazarika, D. Kandar, and A. K. Maji, "An experimental analysis of different deep learning based models for alzheimer's disease classification using brain magnetic resonance images," *Journal of King Saud University-Computer and Information Sciences*, vol. 34, no. 10, pp. 8576–8598, 2022.
- [174] J. Bae, J. Stocks, A. Heywood, Y. Jung, L. Jenkins, V. Hill, A. Katsagelos, K. Popuri, H. Rosen, M. F. Beg, et al., "Transfer learning for predicting conversion from mild cognitive impairment to dementia of alzheimer's type based on a three-dimensional convolutional neural network," *Neurobiology of aging*, vol. 99, pp. 53–64, 2021.
- [175] X. Zhang, L. Han, W. Zhu, L. Sun, and D. Zhang, "An explainable 3d residual self-attention deep neural network for joint atrophy localization and alzheimer's disease diagnosis using structural mri," *IEEE journal of biomedical and health informatics*, vol. 26, no. 11, pp. 5289–5297, 2021.
- [176] Y. Huang, J. Xu, Y. Zhou, T. Tong, X. Zhuang, and A. Initiative, "Diagnosis of alzheimer's disease via multi-modality 3d convolutional neural network," *Frontiers in Neuroscience*, 2019.

- [177] P. Forouzannezhad, A. Abbaspour, C. Li, M. Cabrerizo, and M. Adjouadi, "A deep neural network approach for early diagnosis of mild cognitive impairment using multiple features," in *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, 2018.
- [178] L. Kang, J. Jiang, J. Huang, and T. Zhang, "Identifying early mild cognitive impairment by multi-modality mri-based deep learning," *Frontiers in Aging Neuroscience*, vol. 12, p. 206, 2020.
- [179] A. Khvostikov, K. Aderghal, J. Benois-Pineau, A. Krylov, and G. Catheline, "3d cnn-based classification using smri and md-dti images for alzheimer disease studies," 2018.
- [180] W. Lin, W. Lin, G. Chen, H. Zhang, and M. Du, "Bidirectional mapping of brain mri and pet with 3d reversible gan for the diagnosis of alzheimer's disease," *Frontiers in Neuroscience*, vol. 15, p. 646013, 2021.
- [181] J. Shi, X. Zheng, Y. Li, Q. Zhang, and S. Ying, "Multimodal neuroimaging feature learning with multimodal stacked deep polynomial networks for diagnosis of alzheimer's disease," *IEEE journal of biomedical and health informatics*, vol. 22, no. 1, pp. 173–183, 2017.
- [182] S. El-Sappagh, J. M. Alonso, S. M. R. Islam, A. M. Sultan, and K. S. Kwak, "A multilayer multimodal detection and prediction model based on explainable artificial intelligence for alzheimer's disease," *SCIENTIFIC REPORTS*, vol. 11, JAN 29 2021.
- [183] S. Qiu, M. Miller, I. P. S. Joshi, J. C. Lee, C. Xue, Y. Ni, Y. Wang, I. De Anda-Duran, P. H. Hwang, J. A. Cramer, B. C. Dwyer, H. Hao, M. C. Kaku, S. Kedar, P. H. Lee, A. Z. Mian, D. L. Murman, S. O'Shea, A. B. Paul, M.-H. Saint-Hilaire, E. A. Sartor, A. R. Saxena, L. C. Shih, J. E. Small, M. J. Smith, A. Swaminathan, C. E. Takahashi, O. Taraschenko, H. You, J. Yuan, Y. Zhou, S. Zhu, M. L. Alosco, J. Mez, T. D. Stein, K. L. Poston, R. Au, and V. B. Kolachalam, "Multimodal deep learning for alzheimer's disease dementia assessment," *NATURE COMMUNICATIONS*, vol. 13, JUN 20 2022.
- [184] M. Liu, F. Li, H. Yan, K. Wang, Y. Ma, L. Shen, M. Xu, A. D. N. Initiative, et al., "A multi-model deep convolutional neural network for automatic hippocampus segmentation and classification in alzheimer's disease," *Neuroimage*, vol. 208, p. 116459, 2020.
- [185] J. M. Massaro, R. B. D. Srem, L. M. Sullivan, A. Beiser, C. DeCarli, R. Au, M. F. Elias, and P. A. Wolf, "Managing and analysing data from a large-scale study on framingham offspring relating brain structure to cognitive function," *Statistics in medicine*, vol. 23, no. 2, pp. 351–367, 2004.
- [186] A. Soldan, C. Pettigrew, Y. Lu, M.-C. Wang, O. Selnes, M. Albert, T. Brown, J. T. Ratnanather, L. Younes, M. I. Miller, et al., "Relationship of medial temporal lobe atrophy, apoε genotype, and cognitive reserve in preclinical alzheimer's disease," *Human brain mapping*, vol. 36, no. 7, pp. 2826–2841, 2015.
- [187] M. Raza, M. Awais, W. Ellahi, N. Aslam, H. X. Nguyen, and H. Le-Minh, "Diagnosis and monitoring of alzheimer's patients using classical and deep learning techniques," *Expert Systems with Applications*, vol. 136, pp. 353–364, 2019.
- [188] C. Park, J. Ha, and S. Park, "Prediction of alzheimer's disease based on deep neural network by integrating gene expression and dna methylation dataset," *Expert Systems with Applications*, vol. 140, p. 112873, 2020.
- [189] H. Wang, Y. Shen, S. Wang, T. Xiao, L. Deng, X. Wang, and X. Zhao, "Ensemble of 3d densely connected convolutional network for diagnosis of mild cognitive impairment and alzheimer's disease," *Neurocomputing*, vol. 333, pp. 145–156, 2019.
- [190] O. Colliot, G. Chételat, M. Chaput, B. Desgranges, B. Magnin, H. Benali, B. Dubois, L. Garnero, F. Eustache, and S. Lehéricy, "Discrimination between alzheimer disease, mild cognitive impairment, and normal aging by using automated segmentation of the hippocampus," *Radiology*, vol. 248, no. 1, pp. 194–201, 2008.
- [191] F. Li, L. Tran, K.-H. Thung, S. Ji, D. Shen, and J. Li, "A robust deep model for improved classification of ad/mci patients," *IEEE journal of biomedical and health informatics*, vol. 19, no. 5, pp. 1610–1616, 2015.
- [192] R. Cui, M. Liu, A. D. N. Initiative, et al., "Rnn-based longitudinal analysis for diagnosis of alzheimer's disease," *Computerized Medical Imaging and Graphics*, vol. 73, pp. 1–10, 2019.
- [193] H. Li, M. Habes, D. A. Wolk, Y. Fan, A. D. N. Initiative, et al., "A deep learning model for early prediction of alzheimer's disease dementia based on hippocampal magnetic resonance imaging data," *Alzheimer's & Dementia*, vol. 15, no. 8, pp. 1059–1070, 2019.
- [194] S. Spasov, L. Passamonti, A. Duggento, P. Lio, N. Toschi, A. D. N. Initiative, et al., "A parameter-efficient deep learning approach to predict conversion from mild cognitive impairment to alzheimer's disease," *Neuroimage*, vol. 189, pp. 276–287, 2019.
- [195] Y. Huang, J. Xu, Y. Zhou, T. Tong, X. Zhuang, and A. D. N. I. (ADNI), "Diagnosis of alzheimer's disease via multi-modality 3d convolutional neural network," *Frontiers in neuroscience*, vol. 13, p. 509, 2019.
- [196] B. Lei, E. Liang, M. Yang, P. Yang, F. Zhou, E.-L. Tan, Y. Lei, C.-M. Liu, T. Wang, X. Xiao, et al., "Predicting clinical scores for alzheimer's disease based on joint and deep learning," *Expert Systems with Applications*, vol. 187, p. 115966, 2022.
- [197] M. M. Ghazi, M. Nielsen, A. Pai, M. J. Cardoso, M. Modat, S. Ourselin, L. Sørensen, A. D. N. Initiative, et al., "Training recurrent neural networks robust to incomplete data: Application to alzheimer's disease progression modeling," *Medical image analysis*, vol. 53, pp. 39–46, 2019.
- [198] W. Kang, L. Lin, B. Zhang, X. Shen, S. Wu, A. D. N. Initiative, et al., "Multi-model and multi-slice ensemble learning architecture based on 2d convolutional neural networks for alzheimer's disease diagnosis," *Computers in Biology and Medicine*, vol. 136, p. 104678, 2021.
- [199] H. Barbaroux, X. Feng, J. Yang, A. F. Laine, and E. D. Angelini, "Encoding human cortex using spherical cnns-a study on alzheimer's disease classification," in *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, pp. 1322–1325, IEEE, 2020.
- [200] M. Liu, J. Zhang, C. Lian, and D. Shen, "Weakly supervised deep learning for brain disease prognosis using mri and incomplete clinical scores," *IEEE transactions on cybernetics*, vol. 50, no. 7, pp. 3381–3392, 2019.
- [201] R. Mendoza-Léon, J. Puentes, L. F. Uriza, and M. H. Hoyos, "Single-slice alzheimer's disease classification and disease regional analysis with supervised switching autoencoders," *Computers in biology and medicine*, vol. 116, p. 103527, 2020.
- [202] A. B. Tufail, Y.-K. Ma, and Q.-N. Zhang, "Binary classification of alzheimer's disease using smri imaging modality and deep learning," *Journal of digital imaging*, vol. 33, pp. 1073–1090, 2020.
- [203] A. Ebrahimi, S. Luo, R. Chiong, A. D. N. Initiative, et al., "Deep sequence modelling for alzheimer's disease detection using mri," *Computers in Biology and Medicine*, vol. 134, p. 104537, 2021.
- [204] F. Zhang, Z. Li, B. Zhang, H. Du, B. Wang, and X. Zhang, "Multi-modal deep learning model for auxiliary diagnosis of alzheimer's disease," *Neurocomputing*, vol. 361, pp. 185–195, 2019.
- [205] T. Abuhmed, S. El-Sappagh, and J. M. Alonso, "Robust hybrid deep learning models for alzheimer's progression detection," *Knowledge-Based Systems*, vol. 213, p. 106688, 2021.
- [206] B. Lei, M. Yang, P. Yang, F. Zhou, W. Hou, W. Zou, X. Li, T. Wang, X. Xiao, and S. Wang, "Deep and joint learning of longitudinal data for alzheimer's disease prediction," *Pattern Recognition*, vol. 102, p. 107247, 2020.
- [207] T. Matsuura, K. Sakashita, A. Grushnikov, F. Okura, I. Mitsugami, and Y. Yagi, "Statistical analysis of dual-task gait characteristics for cognitive score estimation," *Scientific Reports*, vol. 9, no. 1, p. 19927, 2019.
- [208] F. Haider, S. De La Fuente, and S. Luz, "An assessment of paralinguistic acoustic features for detection of alzheimer's dementia in spontaneous speech," *IEEE Journal of Selected Topics in Signal Processing*, vol. 14, no. 2, pp. 272–281, 2019.
- [209] X.-a. Bi, X. Hu, H. Wu, and Y. Wang, "Multimodal data analysis of alzheimer's disease based on clustering evolutionary random forest," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 10, pp. 2973–2983, 2020.
- [210] S. Qiu, P. S. Joshi, M. I. Miller, C. Xue, X. Zhou, C. Karjadi, G. H. Chang, A. S. Joshi, B. Dwyer, S. Zhu, et al., "Development and validation of an interpretable deep learning framework for alzheimer's disease classification," *Brain*, vol. 143, no. 6, pp. 1920–1933, 2020.
- [211] N. Zeng, H. Li, and Y. Peng, "A new deep belief network-based multi-task learning for diagnosis of alzheimer's disease," *Neural Computing and Applications*, pp. 1–12, 2021.
- [212] F. J. Martinez-Murcia, A. Ortiz, J.-M. Gorriz, J. Ramirez, and D. Castillo-Barnes, "Studying the manifold structure of alzheimer's disease: a deep learning approach using convolutional autoencoders," *IEEE journal of biomedical and health informatics*, vol. 24, no. 1, pp. 17–26, 2019.
- [213] W. Lin, W. Lin, G. Chen, H. Zhang, Q. Gao, Y. Huang, T. Tong, M. Du, and A. D. N. Initiative, "Bidirectional mapping of brain mri and pet with 3d reversible gan for the diagnosis of alzheimer's disease," *Frontiers in Neuroscience*, vol. 15, p. 646013, 2021.
- [214] H. Choi, Y. K. Kim, E. J. Yoon, J.-Y. Lee, D. S. Lee, and A. D. N. Initiative, "Cognitive signature of brain fdg pet based on deep learning:

- domain transfer from alzheimer's disease to parkinson's disease," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 47, pp. 403–412, 2020.
- [215] R. Janghel and Y. Rathore, "Deep convolution neural network based system for early diagnosis of alzheimer's disease," *Irbm*, vol. 42, no. 4, pp. 258–267, 2021.
- [216] C. Ieracitano, N. Mammone, A. Hussain, and F. C. Morabito, "A novel multi-modal machine learning based approach for automatic classification of eeg recordings in dementia," *Neural Networks*, vol. 123, pp. 176–190, 2020.
- [217] X. Bi and H. Wang, "Early alzheimer's disease diagnosis based on eeg spectral images using deep learning," *Neural Networks*, vol. 114, pp. 119–135, 2019.
- [218] H. Choi, S. Ha, H. Kang, H. Lee, and D. S. Lee, "Deep learning only by normal brain pet identify unheralded brain anomalies," *EBioMedicine*, vol. 43, pp. 447–453, 2019.
- [219] H. W. Kim, H. E. Lee, S. Lee, K. T. Oh, M. Yun, and S. K. Yoo, "Slice-selective learning for alzheimer's disease classification using a generative adversarial network: a feasibility study of external validation," *European journal of nuclear medicine and molecular imaging*, vol. 47, pp. 2197–2206, 2020.
- [220] R. Jain, N. Jain, A. Aggarwal, and D. J. Hemanth, "Convolutional neural network based alzheimer's disease classification from magnetic resonance brain images," *Cognitive Systems Research*, vol. 57, pp. 147–159, 2019.