Review

Application of Deep Learning for Alzheimer's Prediction

Madhur Singh [1,*], Dr Geetha S [2]

[1] School of Computer Science and Engineering, Vellore Institute of Technology, Chennai; madhur.singh2020@vitstudent.ac.in
[2] Professor and Associate Dean (Research), School of Computer Science and Engineering, Vellore Institute of Technology, Chennai; geetha.s@vit.ac.in
* Correspondence: madhur.singh2020@vitstudent.ac.in; Tel.: +91 9041111001

Citation: To be added by editorial staff during production.

Academic Editor: Firstname Lastname

Received: date
Revised: date
Accepted: date
Published: date

Abstract: In our modern era, Alzheimer's and related diseases pose significant health challenges. The application of deep learning in this interdisciplinary domain has generated considerable excitement and displayed promising potential. This paper conducts a comprehensive survey of deep learning literature pertaining to Alzheimer's disease, mild cognitive impairment, and related conditions spanning from 2010 to early 2023. We delve into the various unsupervised, supervised, and semi-supervised methods developed for diverse tasks in this field, with a focus on recent advancements like recurrent neural networks, graph-neural networks, and generative models. Additionally, we present an overview of data sources, data processing techniques, training protocols, and evaluation methods, offering guidance for future deep learning research in Alzheimer's disease. Despite the impressive performance of deep learning in numerous studies and tasks, it grapples with challenges in interpretation and generalization. This survey touches upon these challenges and provides a glimpse into potential avenues for future research endeavors.

Keywords: Alzheimer's Disease, Mild Cognitive Impairment, Deep Learning

1. Introduction

Deep learning demonstrates a remarkable potential in medical image analysis and knowledge discovery. Over the past 13 years, there has been a notable intersection of interest between the interdisciplinary exploration of Alzheimer's disease (AD) and the application of deep learning. The purpose of this paper is to perform a comprehensive survey of the latest developments in deep learning studies related to diverse aspects of Alzheimer's disease research. The scope is not only limited to the current detection methods but also includes the intricacies of generalization and interpretation pathways. The following section presents the contemporary definition of AD and outlines clinical diagnostic methods. Subsequently, we delve into the primary areas of interest within this interdisciplinary study and shed light on the prevailing challenges.

1.1 Alzheimer's Disease and Mild Cognitive Impairment

Alzheimer's disease is the most prevalent form of dementia, posing a significant health challenge in our current era. Projections made by the authors in [1] anticipate that over 1% of the global population will suffer from Alzheimer's or related diseases by 2050. Alzheimer's manifests as a chronic neurodegenerative disorder in middle-to-old age people, though there are rare instances of early-onset cases affecting individuals aged 45–64 [2]. Progression of the disease causes cognitive decline symptoms, including memory impairment, language dysfunction, and diminishing cognition and judgment. Depending on the stage of the disease, individuals with symptoms may require varying levels of assistance in their daily lives, significantly impacting their quality of life (QOL) and that of their families. Studies on the economic burden of dementia and Alzheimer's highlight the substantial societal demand for elderly care, leading to increased overall socio-economic pressure.

The biological processes that lead to Alzheimer's can initiate more than two decades before symptoms manifest. Current insights into Alzheimer's pathogenesis revolve around the deposition of amyloid peptides and the accumulation and phosphorylation of tau proteins surrounding neurons, ultimately resulting in neurodegeneration and eventual brain atrophy. Factors

associated with Alzheimer's include age, genetic predisposition, Down's syndrome, brain injuries, and cardiorespiratory fitness. The cognitive impairment linked to Alzheimer's spans three stages: (1) preclinical AD, where detectable changes in the brain, cerebral spinal fluid (CSF), and blood plasma are observable; (2) mild cognitive impairment (MCI) due to AD, marked by biomarker evidence of Alzheimer's-related brain changes; and (3) dementia due to AD, where noticeable alterations in the brain coincide with memory, thinking, and behavioral changes impacting daily function.

Mild cognitive impairment (MCI), the stage preceding dementia, is most commonly associated with Alzheimer's. However, not all cases of MCI progress to Alzheimer's. Various studies have examined the demographics and progression of MCI, revealing that 15–20% of individuals aged 65 or older experience MCI from various potential causes [23]. At a two-year follow-up, 15% of those with MCI developed dementia, while 32% developed Alzheimer's and 38% developed dementia at a five-year follow-up [24–26]. Early diagnosis of MCI and its subtypes allows for timely intervention, significantly impacting patient longevity and QOL [27]. Thus, gaining a deeper understanding of the condition and developing effective and precise diagnostic methods holds paramount public interest.

1.2 Existing Diagnostic Methods

The present standard for diagnosing AD and MCI relies on a combination of diverse methods. Cognitive assessments, such as the Mini-Mental State Examination (MMSE), Clinical Dementia Rating, and Cambridge Cognitive Examination, involving a series of questions, are commonly employed alongside physical and neurological examinations. The diagnostic process also considers medical and family history, encompassing psychiatric history and the evolution of cognitive and behavioral changes. Genetic sequencing, specifically for biomarkers like the APOE-e4 allele, aids in determining genetic predisposition.

Neuroimaging plays a crucial role in examining various indicators of brain changes and ruling out alternative causes. Structural magnetic resonance and diffusion tensor imaging are widely used to detect signs of brain atrophy. Various computed tomography (CT) forms are also part of AD and MCI diagnosis. Positron emission tomography (PET) includes FDG-PET, which assesses brain glucose metabolism, and amyloid-PET, used to measure beta-amyloid levels. Single-photon emission computed tomography (SPECT), though prone to false-positive results, can be potentially employed in diagnosis. The combination of multiple imaging modalities is a common practice to leverage the strengths of each.

Recent diagnostic advancements involve CSF and blood plasma biomarkers, such as Amyloid-β42, t-tau, and p-tau, along with neurofilament light protein (NFL), neuron-specific enolase (NSE), and HFABP [41,42]. CSF biomarkers are increasingly integral in some AD diagnostic criteria, although the definitive diagnosis, the "ground truth," can only be made post-mortem.

1.3 Application of Deep Learning

In current medical research, the accepted practice involves detailed preprocessing coupled with refined extraction of biomarkers, followed by rigorous statistical analysis. An illustrative study by Gupta et al. [42] applied statistical analysis to biomarkers extracted using voxel-based morphometry and parcelation methods from T1-weighted MRI scans of Alzheimer's disease (AD), mild cognitive impairment (MCI), and healthy controls (HC). The results highlighted statistical significance in various measures, including hippocampal volume and entorhinal cortex thickness. Confirming these findings, Cui Y et al. [57] utilized large deformation diffeomorphic metric mapping (LDDMM) to analyze regional volumetric changes. Various biomarker information extracted was subjected to statistical analysis methods with varying numbers of variables to detect changes in biomarkers during disease development [59]. Similar studies also incorporated other neuroimaging data, genetic data, and cerebrospinal fluid (CSF) biomarkers, advocating the use of MRI imaging biomarkers in AD and MCI diagnosis and forming the foundation for the development of automated diagnostic algorithms.

Machine learning has gained immense popularity in contemporary automated diagnostic algorithms due to its adaptivity to data and the ability to generalize knowledge with lower requirements for expert experience. Klöppel et al. [63] validated the application of machine learning algorithms in diagnosing dementia, comparing the Support Vector Machine (SVM) classification of local grey matter volumes with human diagnosis by professional radiologists. Janousova et al. [64] introduced penalized regression with resampling to identify discriminative regions aiding Gaussian kernel SVM classification, aligning with previous morphological studies. These breakthroughs paved the way for the development of numerous machine-learning algorithms for AD and MCI detection. Zhang et al. [65] proposed a kernel combination method for the fusion of heterogeneous biomarkers for classification with linear SVM. Liu et al. [66] introduced the Multifold Bayesian Kernelization (MBK) algorithm, wherein a Bayesian framework derives kernel weights, and synthesis analysis offers diagnostic probabilities for each biomarker. Zhang et al. [67] proposed eigenbrain extraction using Welch’s t-test (WTT) combined with a polynomial kernel SVM and particle swarm optimization.

There is also substantial interest in applying deep learning to the realm of AD and related diseases. Deep learning integrates the two-step feature extraction and classification process into neural networks, serving as universal approximators based on backpropagation parameter training. Deep learning has witnessed significant strides in the medical data domain, such as breast cancer, tuberculosis, and others. Instead of manual feature crafting, models, and optimizers, deep learning utilize the layered structure of neural networks for the automated abstraction of various levels of features. For instance, Feng et al. [68] utilized a proposed deep learning model to extract biomarkers for MRI in neuroimaging. Their study demonstrated that the deep learning approach surpassed other neuroimaging biomarkers in discerning amyloid and tau pathology and neurodegeneration in prodromal AD.

1.4 Regions of Interest

The goal of this review is to detect and predict neurodegeneration, with a focus on providing early detection and accurate prognoses to facilitate treatment and intervention. The primary areas of interest within this interdisciplinary field can be broadly categorized into three domains:

1. **Classification of Various Stages of AD:** This area centers around the diagnosis and efficient monitoring of disease progression. Current studies primarily concentrate on classifying different stages of Alzheimer's disease (AD), including Mild Cognitive Impairment (MCI) subtypes and normal cognitive controls (NC). Some studies extend their focus to include the subjective cognitive decline (SCD) stage that precedes MCI.
2. **Predicting MCI Conversion:** The emphasis here is on predicting the conversion of MCI, often framed as a classification problem that defines MCI converters and non-converters based on a time threshold from the initial diagnosis. Certain studies delve into predicting the time-to-conversion from MCI to AD.

Additionally, there are other areas of interest, including:

1. **Knowledge Discovery:** Some studies aim to comprehend Alzheimer's disease through data, contributing to knowledge discovery in the field [69].
2. **Phenotyping and Sample Enrichment for Clinical Trials:** Deep learning models are employed to select patients likely to respond to treatment, preventing ineffective or unnecessary interventions in clinical trials [70,71].
3. **Segmentation and Preprocessing:** Deep learning models are applied in segmentation and preprocessing tasks to achieve higher performance or efficiency compared to conventional pipelines [72].

1.5 Challenges in Research

Diagnosing or prognosing of AD and related diseases is marked by uncertainty due to evolving diagnostic criteria and scientific understanding. While deep learning approaches show promise in various areas of interest, several challenges persist, leaving room for improvement:

1. Unavailability of a Comprehensive Dataset: While data for AD and related diseases are abundant, the number of subjects is moderate compared to large datasets like Image-Net, hindering optimal generalization.
2. Numerical Representation of AD Stages: Quantifying differences between AD stages poses challenges. Discrepancies in cognitive decline severity calculations exist, and distinguishing between late mild cognitive impairment and early stages can be subtle. Studies have revealed false cases in clinical and post-mortem diagnoses, complicating the identification of definitive AD signs [73–78].
3. Difficulty in Preprocessing: Preprocessing medical data, particularly neuroimaging data, involves intricate pipelines without standardized procedures. The lack of a universal preprocessing standard leads to variability, and the quality of preprocessing is subject to clinicians' subjective judgment.
4. Privacy Concerns: The utilization of medical data, including neuroimaging and genetic information, brings forth significant privacy concerns. Safeguarding the confidentiality and anonymity of individuals in deep learning studies is crucial to adhere to ethical standards and legal regulations.
5. Ethical Implications: Deep learning studies may involve the use of identifiable patient data, posing ethical challenges related to informed consent, data ownership, and ensuring that individuals' rights are protected throughout the research process.
6. Legal Compliance: Adhering to privacy laws and regulations, such as the Health Insurance Portability and Accountability Act (HIPAA) in the United States, adds complexity to deep learning research. Ensuring compliance with legal frameworks is essential to avoid legal repercussions.
7. Data Security: Protecting medical data from unauthorized access, breaches, or unintended disclosures is a paramount concern. The implementation of robust security measures is necessary to maintain the integrity and confidentiality of sensitive information.
8. Differences in Diagnostic Criteria: Variations in diagnostic criteria, especially in studies predating accessible methods like CSF biomarkers and genetic sequencing, present challenges in establishing ground truth labels.
9. Lack of Reproducibility: Many frameworks and models lack open-source code, impeding transparency in implementation details, dataset selection, preprocessing procedures, and evaluation metrics. Comprehensive frameworks for benchmarking different models based on standardized processing and testing are limited.
10. Lack of Expert Knowledge: Researchers proficient in DL often lack medical backgrounds, creating challenges in understanding complex medical data, particularly in preprocessing and identifying relevant brain regions of interest.

11. Generalizability and Interpretability: DL models face issues of information leakage, limited measures of generalizability to real-world populations, and the "black box" nature hindering interpretation and feedback of knowledge to clinicians.
12. Other Practical Challenges: Subjectivity in cognitive assessments, invasiveness of diagnostic techniques like lumbar puncture, and the high cost of neuroimaging, such as MRI, add to the practical challenges in AD research.

The above sub-section is summarized below in Table 1.

Table 1. Summary of Challenges Faced During AD Prediction

Challenge	Description
Unavailability of a comprehensive dataset	Despite abundant data the number of subjects is moderate, limiting optimal generalization
Numerical Representation of AD Stages	Quantifying differences between AD stages is challenging due to cognitive decline discrepancies and subtle distinctions between late and early stages
Difficulty in Preprocessing	Lack of standardized procedures in preprocessing, especially for neuroimaging, introduces variability, relying on subjective clinician judgment
Privacy Concerns	Utilizing medical data, including neuroimaging and genetics, raises significant privacy concerns, necessitating safeguarding for ethical and legal compliance
Ethical Implications	Involving identifiable patient data introduces ethical concerns related to informed consent, data ownership, and protecting individuals' rights
Legal Compliance	Adhering to privacy laws, like HIPAA, complicates deep learning research, requiring compliance to prevent legal repercussions
Data Security	Protecting medical data from unauthorized access is crucial, necessitating robust security measures for confidentiality
Differences in Diagnostic Criteria	Variations in diagnostic criteria, especially in older studies, challenge consistent ground truth label establishment
Lack of Reproducibility	Limited open-source code transparency hinders framework understanding, and comprehensive benchmarking frameworks are scarce
Lack of Expert Knowledge	Deep learning researchers without medical backgrounds face challenges in understanding medical data, particularly in preprocessing and identifying relevant brain regions
Generalizability and Interpretability	Deep learning models face challenges in generalizing to real-world populations, with a "black box" nature hindering interpretation and knowledge transfer
Practical Challenges	Subjectivity in cognitive assessments, invasiveness of diagnostic techniques, and the high cost of neuroimaging contribute to practical challenges

1.6 Taxonomy

The structure of this paper unfolds as follows: In Section 2, we delve into a comprehensive literature survey on Alzheimer's Disease (AD) prediction. Moving forward in Section 3, we present the data types incorporated in deep learning research, along with potential data sources and the different data modalities used in medical image analysis, succeeded by an exploration of four categories of data processing for neural network data input in Section 4.

Sections 5 and 6 constitute the crux of this paper, delving into deep learning architectures and methods. We categorize these methods into unsupervised, semi-supervised, and supervised learning categories. Each category encompasses typical models and recent advancements, with a specific focus on recent developments in generative models, recurrent neural networks, and graph neural networks.

Section 7 introduces a spectrum of techniques, including transfer learning, ensemble learning, and multi-modal fusion. In Section 8, we introduce an innovative methodology for AD prediction, laying the groundwork for potential future research directions and emphasizing interpretability and generalization. A comprehensive taxonomy of the survey is visually depicted in Fig 1.

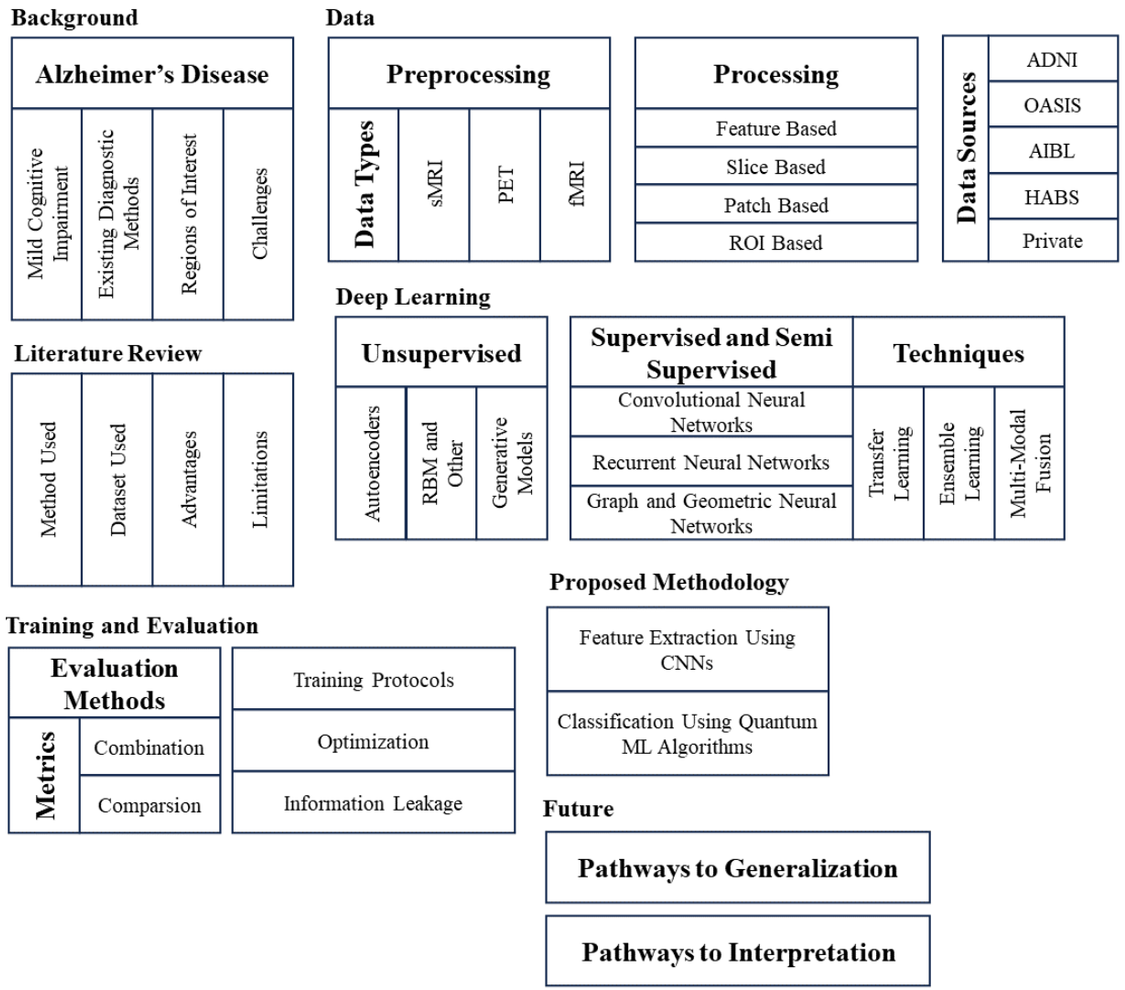

Figure 1. Concept Map of Review Paper

2. Literature Review

Table 2. Literature Survey

S.No	Title	Journal	Year	Method Used	Dataset Link	No. of Images Used	Advantages	Limitations
1	Medial Temporal Lobe Atrophy Predicts Alzheimer's Disease in Patients with Minor Cognitive Impairment	J Neurol Neurosurg Psychiatry	2002	Volumetry of the Hippocampus, Volumetry of the Parahippocampal Gyrus, and Qualitative Rating of Medial Temporal Lobe Atrophy	31 Patients from the Maastricht Memory Clinic		Prospective Follow-up: The study features a prospective follow-up assessment conducted 1 to 3 years after the initial assessment, providing valuable longitudinal data on cognitive decline. MRI Methodology: The use of a three-dimensional volumetric scan and inversion recovery scan for MRI allows for detailed imaging of the hippocampus and parahippocampal gyrus, contributing to precise MTA score determination. Neuropsychological Assessment: The inclusion of a neuropsychological assessment involves standardized clinical tests, such as the AVLT and SCWT, enhancing the depth and reliability of cognitive evaluation. Clinical Applicability: The study suggests that volumetry of the hippocampus is a preferred predictor but acknowledges the time and	Different MRI Scan Axes: The study used different scan axes for measuring hippocampal volume and parahippocampal gyrus volumetry and MTA scoring. While the slice thickness was thin, potential bias due to the difference in scan axis is acknowledged. Long-Term Predictive Models: The study acknowledges the need for further investigation to determine the simplest model for predicting long-term outcomes beyond the follow-up period. Selective Age Group: The mean age of the study sample is lower than in comparable studies, potentially impacting the positive predictive value of medial temporal lobe atrophy as the conversion rate to dementia is lower in younger patients. Depression Consideration: Patients with mild to moderate depression

							resource limitations in clinical settings, making MTA scoring a practical alternative with good predictive accuracy.	were not excluded, possibly introducing a confounding factor. However, analyses with correction for depression severity yielded similar results.
2	Qualitative Estimates of Medial Temporal Atrophy as a Predictor of Progression From Mild Cognitive Impairment to Dementia	Arch Neurol	2007	Qualitative Study	769 participants were recruited from 69 Alzheimer's Disease Cooperative Studycenters in the United States and Canada.	190 images	Blinded Prospective Design: The study design's blinded and prospective nature enhances the reliability and objectivity of the findings, minimizing bias in the evaluation process. Well-Defined Patient Cohort: Inclusion of a well-defined patient cohort with detailed surveillance by experienced physicians provides robust and clinically relevant data for analysis. Physician Expertise: Physicians with dementia experience conducted detailed surveillance, ensuring a high level of expertise in evaluating patients, contributing to the study's credibility. Simple and Translatable Method: The simplicity and good reliability of the method used make it easily translatable into standard clinical practice, potentially facilitating widespread adoption.	Selective Study Cohort: The study cohort's select nature, including well-educated individuals with moderately severe memory impairments, may limit the generalizability of findings to broader populations. Effect of Experience: The study suggests that experience with the scale may modestly affect its predictive value, particularly with a cutoff score greater than 2.0, indicating potential limitations in generalizability. Small Number of Participants with High Scores: The study's small number (n = 15) of participants with mean MTA scores greater than 2.0 may contribute to differences among raters, emphasizing the need for larger sample sizes.
3	Combining MR Imaging, Positron-Emission	American Journal of Neuroradiology	2010	Regression Analyses	Alzheimer's Disease Neuroimaging Initiative (ADNI)	AD: 38 MCI: 73 NC: 42	Multimodal Approach: The study employs a multimodal approach, integrating MR imaging	Limited Plasma Biomarkers: The study focuses primarily on CSF biomarkers, neglecting potential

	Tomography, and CSF Biomarkers in the Diagnosis and Prognosis of Alzheimer Disease			morphometry, FDG-PET, and CSF biomarkers, providing a comprehensive view of neurodegenerative changes in Alzheimer's disease (AD). Diagnostic Sensitivity: The study demonstrates sensitivity to diagnostic status across morphometry, metabolism, and CSF biomarkers, highlighting the potential of these measures in distinguishing between normal controls (NC) and AD. Individual Prognostic Potential: MR imaging morphometry measures, in particular, show promise for individual prognostic use, potentially aiding in the identification of patterns of atrophy predictive of conversion to AD. Comparative Analysis: The study compares the contributions of MR imaging morphometry, FDG-PET, and CSF biomarkers, providing insights into their unique and redundant aspects in both diagnostic accuracy and clinical prediction.	contributions from plasma biomarkers. The inclusion of plasma biomarkers could enhance the understanding of peripheral indicators of AD. Heterogeneous MCI Group: The MCI group in the study might be more heterogeneous compared to previous CSF studies, impacting the predictive power of CSF biomarkers. The study acknowledges potential variability in the MCI population. Discrepancies in CSF Predictions: The study notes discrepancies in the predictive value of CSF measures for clinical decline, possibly influenced by the heterogeneity of the MCI group and the use of continuous behavioral measures rather than conversion. Scatter in Individual Prognostic Measures: While MR imaging morphometry measures show potential for individual prognostic use, there is considerable scatter in regression plots, indicating uncertainty in individual predictions. Sample Selection Bias: Participants were selected based on their willingness and ability to undergo

								specific imaging and biomarker assessments, potentially introducing selection bias and limiting the generalizability of findings.
4	Learning brain connectivity of Alzheimer's disease by sparse inverse covariance estimation	NeuroImage	2010	Sparse Inverse Covariance Estimation (SICE)	Alzheimer's Disease Neuroimaging Initiative (ADNI)	AD: 49 MCI: 116 NC: 67	Connectivity Model Identification: The proposed SICE method successfully identifies connectivity model structures from PET data, providing insights into the functional brain connectivity of different groups (AD, MCI, NC). Order of Inter-Region Connections: With the aid of a quasi-measure, SICE can determine the order of inter-region connections in terms of connection strength, contributing to a more detailed understanding of network characteristics. Group Comparison Insights: The application of SICE to ADNI FDG-PET data revealed significant findings, including decreased connections in the temporal lobe in Alzheimer's disease (AD) and increased connections in the frontal lobe, suggesting compensatory effects. Clinical Trial Application: The methodology proposed in the paper holds potential for application in clinical trials. It can be used to assess drug efficacy by comparing connectivity patterns between	Data Heterogeneity in MCI Group: The MCI group used for brain connectivity modeling is heterogeneous, consisting of subjects with varying outcomes (e.g., conversion to AD, remaining as MCI, or reverting to normal). The reliability of connectivity models for MCI may be affected by this heterogeneity. Lack of Longitudinal Diagnostic Results: The diagnostic results of the MCI subjects at later checkups were not available during the development of the paper. Future work plans to address this limitation by splitting the MCI group based on later diagnostic results, leading to more reliable models. Linear Interaction Modeling: SICE provides a model for linear interactions between brain regions based on the covariance matrix. Future research directions could explore nonlinear interactions by incorporating discretized measurements and building graphical models.

						groups receiving and not receiving a certain drug, with the advantage of producing reliable models with small sample sizes. Potential for fMRI Modeling: The approach may be extended for functional brain connectivity modeling based on fMRI data. Subject-level connectivity models could be identified, and connectivity-based biomarkers for AD and MCI may be explored, complementing existing region-based biomarkers.	Preprocessing Procedure: The paper acknowledges the use of the default SPM5 registration in the current preprocessing procedure. Future work aims to explore improved image registration algorithms in SPM5/8 (DARTEL).
5	Multimodal Classification of Alzheimer's Disease and Mild Cognitive Impairment	NeuroImage	2011	Alzheimer's Disease Neuroimaging Initiative (ADNI)	AD: 51 NC: 52 MCI: 99	Use of Multiple Modalities: The authors used a multi-kernel SVM to integrate multiple modalities, namely MRI, PET and CSF. 93 features each were extracted from MRI and PET each and 3 features from CSF biomarkers. Such integration is advantageous as different biomarkers provide complementary information which is helpful in diagnosis of AD or MCI. Consideration of Diversity of Individual Modalities: Jaccard Similarity Coefficient and Kappa Index were used as quantitative measurements of diversity on any two modalities (MRI vs PET, PET vs CSF and CSF vs MRI). These results indicate that CSF and PET have the highest complementary information, while MRI	Lack of Data: Besides MRI, PET, and CSF, there are also other modalities of data, i.e., APOE. However, since not every subject has data on all modalities and the number of subjects with all modalities available is too small for reasonable classification, the current study does not consider APOE for multimodal classification. Unable to discriminate among multiple stages of dementia: In the current study, investigation has been done only on the classification between one stage of dementia (either MCI or AD) and healthy controls. It does not test the ability of the classifier to simultaneously discriminate multiple stages of dementia, i.e., multi-class

							and PET have the highest similar information for classification. Ensemble Method for Data Fusion: Multiple SVM models are trained on multiple kernel matrices from different modalities. For any particular test sample, each model will make a prediction and the final output is decided using majority voting. Prominent Evaluation Metrics: The method discussed in the paper, achieved a high accuracy (93.2%) for AD classification, a relatively high sensitivity (81.8%) for MCI classification, and especially a high sensitivity (91.5%) for classification of MCI converters.	classification of AD, MCI, and healthy controls.
6	Resting-state Multi-spectrum Functional Connectivity Networks for Identification of MCI Patients	PloS One	2012	Resting-state Multi-spectrum Functional Connectivity Networks	Data was generated by studying subjects recruited by the Duke-UNC Brain Imaging and Analysis Center (BIAC), Durham, North Carolina, USA.	MCI: 12 NC: 25	Utilization of multi-spectrum functional connectivity networks for identification of MCI patients: The study leverages the power of multi-spectrum functional connectivity networks, allowing for a comprehensive analysis of various functional connectivity patterns within the brain. By incorporating multiple spectra, the research can potentially capture more nuanced and comprehensive information about the functional connectivity changes associated with mild cognitive	Possible challenges in generalizing the results to diverse datasets: As the dataset is developed solely for the particular paper, the generalizability of the findings to different populations or settings may be limited. Addressing these potential variations and ensuring the robustness of the findings across diverse datasets is essential for establishing the broader applicability of the proposed methodology in identifying individuals with MCI.

							impairment (MCI), aiding in the early identification and diagnosis of at-risk individuals. Provides insights into the use of resting-state functional connectivity for early detection of cognitive impairment: By focusing on resting-state functional connectivity networks, the research offers valuable insights into the potential use of intrinsic brain activity patterns as biomarkers for the early detection of cognitive impairment. This approach enables the identification of specific functional connectivity alterations that may serve as early indicators of cognitive decline, facilitating timely intervention and management strategies for individuals at risk of developing MCI.	
7	Multi-modal Multi-task Learning for Joint Prediction of Multiple Regression and Classification Variables in Alzheimer's Disease	NeuroImage	2012	Multi-modal Multi-task Learning	Alzheimer's Disease Neuroimaging Initiative (ADNI)	186 ADNI subjects with all MRI, PET and CSF data AD: 45 MCI: 91 NC: 50	Innovative Multi-Modal Multi-Task Learning: The proposed M3T learning method introduces a novel approach by combining two successive steps: multi-task feature selection and multi-modal support vector machine. This innovative method aims to jointly predict multiple regression and classification variables from multimodal data. Recognition of Complementary Information from Different Modalities: Acknowledging the	Dependency on Availability of Multi-Modal Data: The M3T method relies on the availability of multi-modal data, specifically MRI, PET, and CSF. The requirement for each subject to have corresponding modality data limits the size of the subject pool for study. Exclusion of Other Modalities (e.g., APOE): Despite the existence of other modalities such as APOE data, the study does not consider

			complementary information from various modalities, the M3T method effectively combines MRI, PET, and CSF data for joint regression and classification tasks. It outperforms individual-modality-based methods, showcasing the advantage of leveraging multi-modal information. Robustness to Feature Selection Variability: The M3T method incorporates feature selection techniques like MTFS and Lasso, ensuring adaptability to different subsets of selected features in cross-validation trials. Crucial features, such as hippocampal regions, consistently contribute to the model. Consideration of Multi-Modal Regression: In contrast to existing works that focus on multi-modal classification, the M3T method extends its application to multi-modal regression. It demonstrates that combining MRI, PET, and CSF data enhances the performance of regression models. Use of Multi-Modal SVM for Both Regression and Classification: The current model employs multi-modal SVM for both regression and classification tasks. The linear kernel,	them due to data limitations. The exclusion of certain modalities may impact the comprehensiveness of the analysis. Limited Sample Size for Comprehensive Study: The study acknowledges limitations in sample size, especially concerning subjects with all baseline MRI, PET, and CSF data. This limitation may affect the generalizability of the findings. Exclusion of Some Clinical Variables: Due to data availability constraints, the study does not investigate certain clinical variables present in the ADNI database. While including more clinical variables could enhance performance, this was not explored in the current study.

							with normalized feature vectors, proves effective and requires no additional parameter tuning.	
8	Sparse Learning and Stability Selection for Predicting MCI to AD Conversion Using Baseline ADNI Data	BMC Neurology	2012	Sparse Learning and Stability Selection	Alzheimer's Disease Neuroimaging Initiative (ADNI)	MCI: 319	Large and Unbiased MCI Cohort: The study benefits from a large cohort of Mild Cognitive Impairment (MCI) samples, ensuring statistical robustness. The crucial aspect of this advantage is that the cohort is unbiased concerning age or education status, minimizing confounding variables that could affect the results. Integration of Various Baseline Data: Unlike some other studies, this research integrates and tests a diverse range of baseline data available in the Alzheimer's Disease Neuroimaging Initiative (ADNI). This comprehensive approach includes data from MRI scans, demographic information, genetic factors (APOE genotyping), and cognitive measures. The inclusion of multiple types of data enhances the depth and richness of the analysis. Application of Sparse Logistic Regression with Stability Selection: The study employs advanced statistical techniques by applying sparse logistic regression with stability selection to ADNI data. This methodology ensures robust feature	Cerebellar Atrophy Association: The study notes a surprising correlation between cerebellar atrophy and AD. While this association has been detected in other studies, the specific role of the cerebellum in AD remains unclear. This unexpected finding highlights the need for further investigation to understand the significance of cerebellar atrophy in the context of MCI-to-AD conversion. Cingulate Cortex Atrophy: The study identifies the atrophy of the rostral anterior cingulate cortex as predictive of conversion to AD. While this aligns with previous studies, the specific implications and functional consequences of cingulate cortex atrophy in early AD stages warrant further exploration. CSF Biomarkers Lack Specificity: The results suggest that cerebrospinal fluid (CSF) biomarkers, while showing an aberrant signature in MCI Converters, lack the specificity to discriminate between MCI to AD Converters and Non-converters. This limitation

							selection, enabling the identification of the most relevant variables for predicting the outcome. Sparse logistic regression helps prevent overfitting and enhances the interpretability of the selected features. Four-Year Follow-Up Period: The evaluation considers a 4-year follow-up period, providing a longitudinal perspective on the progression of MCI. This extended timeframe allows for a more comprehensive understanding of the factors influencing the conversion from MCI to other conditions.	emphasizes the need for additional and more specific biomarkers for accurate predictions. Feature Interpretation and Redundancy: While the study effectively identifies a set of 15 features (Biosignature-15) with high predictive power, some of these features are known to be important in characterizing Alzheimer's Disease (AD). This raises a challenge in distinguishing whether the identified features genuinely contribute to prediction or if they are redundant with existing knowledge. The reliance on features closely associated with AD may limit the novelty of the findings.
9	Multivariate and Univariate Neuroimaging Markers of Alzheimer's Disease	NeuroImage	2012	Univariate and Multivariate Discriminant Analysis of FDG–PET	Self Built Dataset using Patients from University of Michigan and Technical University of Munich	Michigan AD: 17 HC: 33 Munich AD:102 HC: 20	Multivariate Analysis Sensitivity: The multivariate analytic method shows higher sensitivity (range [0.85,1]) compared to the univariate method, making it potentially more effective in identifying early Alzheimer's disease. Robustness of Multivariate Marker: The multivariate marker appears to be more robust, as demonstrated by the unchanged ROC characteristics even when specific brain regions are removed. This suggests that multivariate analysis, using the entire spatial covariance structure of the	Non-Random Recruitment Bias: The study acknowledges potential biases in non-random recruitment that may limit the generalization of findings to the broader population. This raises concerns about the representativeness of the study sample. Small Sample Size Concerns: The study recognizes the limitations of small sample sizes, especially when trying to map the neural correlates of AD progression. The regional composition of disease markers varied substantially based

		data, may capture more comprehensive information. Prospective Validation: The study includes a prospective validation approach, applying identified patterns to new data samples. This enhances the credibility of the findings and supports the potential generalizability of the identified markers. Realistic Context: The data is divided into derivation and replication samples, providing a more realistic context for evaluating the effectiveness of the analytic methods. This helps in understanding how well the findings generalize to different datasets. Systems-Level Biomarker Potential: The study suggests that FDG–PET imaging combined with multivariate analysis has promise as a systems-level biomarker for Alzheimer's disease. This could enable early detection before clear clinical symptoms manifest.	on the selected AD sample, suggesting the need for larger subject numbers for more reliable insights. Specificity Confirmation Pending: The study doesn't claim confirmation of the specificity of the AD-related covariance pattern to Alzheimer's disease. Further research is needed to validate the specificity of the identified patterns with respect to other neurodegenerative diseases. Lack of Mechanistic Insight: The study emphasizes that the identified FDG–PET patterns are downstream effects of the disease and do not provide mechanistic insight into the etiology of Alzheimer's disease. The focus is on diagnostic efficacy rather than understanding the underlying mechanisms. Trade-off in Biomarker Development: The study discusses the trade-off between reductionist strategies targeting specific diseases with molecular compounds and a more widely applicable approach using standard imaging and analytic technology. While the latter is cost-effective, specificity to Alzheimer's disease needs further confirmation.

10	Alzheimer's Disease Pattern Of Brain Atrophy Predicts Cognitive Decline In Parkinson's Disease	Brain	2012	Voxel-Based Morphometry Analysis	University of Pennsylvania Center of Excellence for Research on Neurodegenerative Diseases (CERND)	84 participants	Validation of Neurodegeneration Patterns: By applying a validated Alzheimer's disease-pattern of brain atrophy to MRI scans, the study offers a robust foundation for the examination of neurodegenerative patterns in Parkinson's disease. This validation enhances the reliability and interpretability of the findings. Prediction of Cognitive Decline in Non-Demented Patients: Notably, the Alzheimer's disease pattern of atrophy predicts cognitive decline even in non-demented patients with Parkinson's disease. This finding suggests the potential utility of neurodegeneration patterns as preclinical biomarkers for cognitive decline, providing an opportunity for early intervention.	Lack of Formal Diagnostic Criteria: The absence of formal diagnostic criteria for Mild Cognitive Impairment (MCI) and dementia is a limitation. The use of standardized criteria would enhance the precision of cognitive status classification and contribute to the validity of the findings. Sample Characteristics: The study primarily includes patients with mild to moderate stage Parkinson's disease. This sample characteristic might limit the generalizability of the findings to individuals with more advanced stages of the disease. Enrolling a broader range of disease stages would provide a more comprehensive understanding. Limited Follow-up Period: The 2-year follow-up period may be relatively short, especially considering the progressive nature of neurodegenerative diseases. Longer follow-up durations would offer a more thorough exploration of cognitive decline trajectories. Psychotic Symptom Assessment: Lack of a validated rating scale for assessing psychotic symptoms is acknowledged as a limitation. The

								sensitivity of the assessment tool used may impact the detection of psychotic symptoms, potentially influencing the study's findings. Need for Larger Samples: The study acknowledges the need for larger sample sizes to enhance statistical power and generalizability. Enrolling larger cohorts would strengthen the robustness of the results and allow for more reliable conclusions. Cautions in Extrapolation: While the study suggests an overlap in neurodegenerative regions between Alzheimer's disease and Parkinson's disease, caution is warranted in extrapolating these findings without direct clinicopathological correlation. The specific contributions of each pathology to cognitive decline need further investigation.
11	Sparse Learning and Stability Selection for Predicting MCI to AD Conversion Using Baseline ADNI Data	BMC Neurology	2012	Sparse Learning and Stability Selection	Alzheimer's Disease Neuroimaging Initiative (ADNI)	MCI: 319	Large and Unbiased MCI Cohort: The study benefits from a large cohort of Mild Cognitive Impairment (MCI) samples, ensuring statistical robustness. The crucial aspect of this advantage is that the cohort is unbiased concerning age or education status, minimizing confounding variables that could affect the results.	Cerebellar Atrophy Association: The study notes a surprising correlation between cerebellar atrophy and AD. While this association has been detected in other studies, the specific role of the cerebellum in AD remains unclear. This unexpected finding highlights the need for further investigation to understand the significance of cerebellar

		Integration of Various Baseline Data: Unlike some other studies, this research integrates and tests a diverse range of baseline data available in the Alzheimer's Disease Neuroimaging Initiative (ADNI). This comprehensive approach includes data from MRI scans, demographic information, genetic factors (APOE genotyping), and cognitive measures. The inclusion of multiple types of data enhances the depth and richness of the analysis. Application of Sparse Logistic Regression with Stability Selection: The study employs advanced statistical techniques by applying sparse logistic regression with stability selection to ADNI data. This methodology ensures robust feature selection, enabling the identification of the most relevant variables for predicting the outcome. Sparse logistic regression helps prevent overfitting and enhances the interpretability of the selected features. Four-Year Follow-Up Period: The evaluation considers a 4-year follow-up period, providing a longitudinal perspective on the progression of MCI. This extended timeframe allows for a more comprehensive	atrophy in the context of MCI-to-AD conversion. Cingulate Cortex Atrophy: The study identifies the atrophy of the rostral anterior cingulate cortex as predictive of conversion to AD. While this aligns with previous studies, the specific implications and functional consequences of cingulate cortex atrophy in early AD stages warrant further exploration. CSF Biomarkers Lack Specificity: The results suggest that cerebrospinal fluid (CSF) biomarkers, while showing an aberrant signature in MCI Converters, lack the specificity to discriminate between MCI to AD Converters and Non-converters. This limitation emphasizes the need for additional and more specific biomarkers for accurate predictions. Feature Interpretation and Redundancy: While the study effectively identifies a set of 15 features (Biosignature-15) with high predictive power, some of these features are known to be important in characterizing Alzheimer's Disease (AD). This raises a challenge in distinguishing whether the

							understanding of the factors influencing the conversion from MCI to other conditions.	identified features genuinely contribute to prediction or if they are redundant with existing knowledge. The reliance on features closely associated with AD may limit the novelty of the findings.
12	A semi-quantitative method for correlating brain disease groups with normal controls using SPECT: Alzheimer's disease versus vascular dementia	Computerized Medical Imaging and Graphics	2012	Semi-Quantitative Circumferential-Profile Analysis of Regional Cerebral Blood Flow (RCBF) SPECT in Alzheimer's Disease (AD) Versus White Matter Vascular Dementia (WM-VAD)	Patients referred over a five-year period from the UAB Memory Clinic and Alzheimer's Disease Centers to the Division of Nuclear Medicine of University of Alabama at Birmingham Medical Center (UAB) for Tc99m HMPAO brain SPECT	MCI: 86 NC: 17	Novel Approach to Dementia Differentiation: The study proposes a novel approach to differentiate between dementia subtypes, particularly Alzheimer's Disease (AD) and White Matter Vascular Dementia (WM-VaD), by assessing regional cerebral blood flow (rCBF) patterns. Clear Differentiation of rCBF Patterns: The study demonstrates clear and significant differences in rCBF among controls, AD, and WM-VaD patients, offering a distinct visualization of the cerebral impairment pattern associated with each condition. Clinical Relevance: The findings have potential clinical relevance, suggesting that WM-VaD may play a more prominent role in dementia than previously suspected. This could impact diagnostic and treatment strategies. Utilization of SPECT for Differential Diagnosis: The study supports the	Semi-Quantitative Scoring System: The scoring system used for quantifying the severity of White Matter Hyperintensities (WMH) is relatively simple and involves four stages in volume. This simplicity may limit the precision of the regional analysis. Referral Bias and Sample Composition: The high incidence of WM-VaD patients in the dementia evaluation sample may reflect referral bias, potentially impacting the generalizability of the findings to broader populations. Limited Regional and Volumetric Data: The study acknowledges that further subdivision of the WM-VaD group based on more precise regional and volumetric data could offer additional insights. This limitation implies that the current study provides a broad overview. Absence of Automated Computational Quantification: The study

			clinical utility of Single Photon Emission Computed Tomography (SPECT) in diagnosing and differentiating between dementia types, providing additional evidence for the role of SPECT in dementia evaluation. Focus on Frontal Cortical Involvement: The identification of a primarily frontal cortical involvement in WM-VaD patients contributes to a better understanding of the neurobehavioral effects associated with this type of dementia.	suggests that a more complex automated computational quantification of WMH might help reduce inter-rater subjectivity. The absence of such an approach limits the study's objectivity. Causality vs. Association: The study acknowledges that the relationship between WMH and clinical symptoms is not necessarily causal. Understanding the precise relationship between WMH and dementia requires further investigation. Difficulty in Differential Diagnosis: The study notes that distinguishing between typical AD and typical WM-VaD may be clinically challenging, especially in cases involving later stages of AD, frontal degenerative processes, and non-frontal VaD. Limited Validation of SPECT Approach: While the study supports the clinically feasible approach of analyzing brain SPECT for differentiating WM-VaD from AD, the validation of this approach needs further confirmation through larger studies and comparison with gold standard diagnostic methods.

13	Partial Least Squares For Discrimination in fMRI Data	Magnetic Resonance Imaging	2012	Linear Discriminant Analysis (LDA), Principal Component Analysis (PCA), Partial Least Squares (PLS), Orthogonal Partial Least Squares (OrPLS)	13 women with high Alzheimer's disease (AD) risk and 11 with low risk based on family history and apolipoprotein-E4 status	Not Specified	Focused Dimension Reduction with OrPLS: The use of Orthogonal Partial Least Squares (OrPLS) is advocated as an alternative to PCA. OrPLS is suggested to be more effective for dimension reduction in the context of discrimination among groups of subjects. The study implies that OrPLS may provide a more suitable approach by incorporating information on class structure. Preserving Discriminative Information: The study suggests that methods like OrPLS may better preserve the discriminative information present in functional neuroimaging data. It emphasizes the need for techniques that can capture relevant patterns in distributed brain networks for accurate group discrimination. Superior Performance of OrPLS: The text asserts the superior performance of OrPLS over PCA in the context of Linear Discriminant Analysis (LDA) for identifying brain functional networks. It indicates that OrPLS achieves a higher classification accuracy compared to PCA-based approaches. Stability in	Parcellation Scheme Impact: The use of a parcellation scheme fixed to the Talairach atlas is acknowledged to have potential drawbacks. It is noted that this method may dilute activation and reduce sensitivity and specificity. A suggestion is made to consider an intersection of parcellation-based and functionally defined ROI methods to mitigate this effect. Dependency on Thresholds: The proposed approach suggests including only those voxels with significant activation in the calculation of subject-specific ROI mean values. This introduces a dependency on the choice of thresholds for defining significant activation, which could impact the results. Small Sample Size for Longitudinal Data: While the study employs longitudinal data, it mentions that further validation with an independent sample will be conducted in the future. The small sample size in longitudinal validation raises concerns about the generalizability of the findings. Computational Challenges with Voxel-Wise Data: The study recognizes the increased complexity of

							Discriminant Patterns: The study references Habeck et al.'s findings regarding the stability of the spatial pattern of weights in OrPLS. It suggests that the discriminant patterns obtained through OrPLS are more stable, which is crucial for reliable discrimination among subjects. Longitudinal Validation: The study presents the use of follow-up fMRI test data from the same subjects acquired four years later. It claims that the OrPLS method demonstrated low misclassification rates in this longitudinal validation, supporting its potential for robust discrimination over time. Identification of Relevant Brain Networks: OrPLS is credited with identifying brain networks that include specific regions such as the ventral temporal lobe, Brodmann areas 19 and 37, the praecuneus, and cingulate gyrus. This suggests that OrPLS not only discriminates effectively but also provides insights into the neuroanatomical basis of the discrimination.	applying OrPLS to voxel-wise neuroimaging data. Computational approaches to deal with the inversion of very large covariance matrices will need to be developed, indicating potential challenges in scalability. Comparison with Other Discrimination Techniques: The study primarily compares OrPLS with PCA-LDA approaches. However, a broader comparison with other state-of-the-art discrimination techniques commonly used in neuroimaging studies would enhance the understanding of the method's relative strengths and weaknesses.
14	Prediction of Alzheimer's Disease in Subjects with Mild Cognitive Impairment from the	NeuroImage	2013	Cortical Thickness Measurement, Machine Learning	Alzheimer's Disease Neuroimaging Initiative (ADNI)	AD: 194 CN: 226	Utilization of cortical thinning patterns for Alzheimer's disease prediction in individuals with mild cognitive impairment: The study leverages cortical thickness	Possible confounding factors not fully addressed: While the study focuses on cortical thinning patterns, it may not fully address all potential confounding factors that

	ADNI Cohort Using Patterns of Cortical Thinning						measurements as potential biomarkers for predicting the progression from mild cognitive impairment (MCI) to Alzheimer's disease. By focusing on changes in cortical thickness, the research highlights the potential of structural neuroimaging as an effective tool for early detection and prediction of Alzheimer's disease. Leave-One-Out Validation Technique to prevent bias: A leave-one-out (LOO) validation strategy was employed, wherein, for each comparison, all subjects, with the exception of one, were employed to select features and construct a classification model. Subsequently, the excluded subject was utilized for testing. This process was iterated for every subject within the two compared groups, thereby validating the method across all subjects. By excluding the test subject from feature selection and classifier construction, any potential bias or "double dipping" in predictive efforts for converters was effectively avoided.	could influence cortical thickness measurements. Factors such as age-related cortical changes or comorbid conditions could potentially impact the accuracy of the predictions, highlighting the need for comprehensive consideration of various contributing factors in the analysis.
15	Deep Sparse Multi-Task Learning for Feature Selection in Alzheimer's Disease Diagnosis	Brain Structure and Function	2013	Deep Sparse Multi-Task Learning, Feature Selection	Alzheimer's Disease Neuroimaging Initiative (ADNI)	AD: 198 pMCI: 167 sMCI: 236 NC: 229	Addressing 'High-Dimension and Small Sample' Problem: The method tackles the challenging issue of 'high-dimension and small sample' in neuroimaging-based Alzheimer's Disease/Mild Cognitive Impairment	Sensitivity to Hyperparameters: Deep sparse multi-task learning methods often rely on the selection of appropriate hyperparameters to achieve optimal performance. The sensitivity of these methods to the

		(AD/MCI) diagnosis, a common problem in this field. Sparse Multi-Task Learning for Feature Selection: The proposed method employs sparse multi-task learning for feature selection, effectively reducing dimensionality and addressing the challenge of noise interference with informative features during optimization. Incorporation of Subclass Labeling Scheme: The method incorporates a subclass labeling scheme, reflecting the complex distributional characteristics in each class, which contributes to a more nuanced and accurate representation in AD/MCI diagnosis. Iterative Filtering in Hierarchical Fashion: Instead of selecting informative features in a single hierarchy, the method iteratively filters out uninformative features in a hierarchical fashion, preventing the underestimation of informative features and overestimation of uninformative features. Utilization of Regression Coefficients as Context Information: At different hierarchies, the method utilizes regression coefficients	choice of hyperparameters can pose a significant challenge, as the effectiveness of the feature selection process and the overall diagnostic accuracy may heavily depend on the specific settings of these parameters. Inaccurate or suboptimal hyperparameter choices may lead to subpar feature selection results and affect the overall diagnostic performance, highlighting the importance of careful parameter tuning and optimization. Interpretability of Feature Selection Results: While deep learning-based feature selection methods can effectively identify relevant biomarkers and imaging features, interpreting the specific rationale behind the selection of these features can be challenging. The lack of interpretability in the feature selection process may hinder the understanding of the underlying biological or pathological significance of the selected features. Interpreting the selected features in the context of Alzheimer's disease pathology is crucial for gaining insights into the disease mechanisms and improving the clinical interpretability and translational value of the feature selection results.

							optimized in the lower hierarchy as context information, enhancing the determination of informative features for classification. Equal Consideration of Feature Importance from Different Modalities: While acknowledging the potential impact of different modalities, the method treats features from different modalities equally. However, it suggests the possibility of adapting to modality-specific importance through the use of multi-kernel SVM, as demonstrated by prior studies. Consideration of Subjective Cognitive Complaints as a Genetic Risk Factor: The method recognizes the significance of subjective cognitive complaints as an important genetic risk factor, contributing to the identification of individuals in the 'pre-MCI' stage, an aspect often underestimated in the field.	Data Dependency and Generalizability: The performance of deep sparse multi-task learning methods for feature selection can be heavily influenced by the characteristics of the specific datasets used for training and evaluation. Variations in data distributions, imaging protocols, and patient demographics across different datasets can impact the effectiveness and generalizability of the feature selection results. Ensuring the robustness and generalizability of the feature selection methodology across diverse datasets is essential for establishing the broader applicability and reliability of the proposed approach in different research and clinical settings.
16	Statistical Analysis of Longitudinal Neuroimage Data with Linear Mixed Effects Models	NeuroImage	2013	Linear Mixed Effects Models	Alzheimer's Disease Neuroimaging Initiative (ADNI)	AD: 188 cMCI: 166 sMCI: 227 HC: 210	Powerful and Versatile Framework: Linear Mixed Effects (LME) models offer a powerful and versatile framework for analyzing longitudinal data, particularly suitable for handling unbalanced data with variable missing rates across timepoints and imperfect timing.	Complexity of Analysis: The comprehensive nature of LME models, while advantageous, introduces complexity in managing and interpreting large datasets. This complexity may pose challenges, especially for researchers unfamiliar with the intricacies of LME models.

							Effective Handling of Unbalanced Data: LME models elegantly handle unbalanced data, accommodating subjects with a single time-point to characterize inter-subject variation, providing a robust representation of group mean trajectories and covariance structures between serial measurements. Enhanced Sensitivity in Realistic Settings: The study demonstrates that LME models provide enhanced sensitivity in realistic LNI settings compared to alternative methods like repeated measures ANOVA or the analysis of summary metrics. This is crucial for accurately detecting effects in complex datasets.	Focused on Univariate Analysis: The study primarily focuses on univariate analysis, where the correction for "multiple comparisons" is not addressed. Future work is suggested to extend the LME framework to the mass-univariate setting, acknowledging the need for more complex analyses involving a large number of pixels/voxels. Subject to Model Fit Improvement: While the study highlights the potential improvement in model fit afforded by including subjects with a single time-point, this may introduce biases or assumptions that need to be carefully considered. The generalizability of this improvement to other datasets should be explored in future research.
17	Prediction of Alzheimer's Disease in Subjects with Mild Cognitive Impairment Using Patterns of Cortical Thinning	NeuroImage	2013	Cortical Thickness Measurement, Machine Learning	Alzheimer's Disease Neuroimaging Initiative (ADNI)	AD: 194 CN: 226	Utilization of cortical thinning patterns for Alzheimer's disease prediction in individuals with mild cognitive impairment: The study leverages cortical thickness measurements as potential biomarkers for predicting the progression from mild cognitive impairment (MCI) to Alzheimer's disease. By focusing on changes in cortical thickness, the research highlights the potential of	Possible confounding factors not fully addressed: While the study focuses on cortical thinning patterns, it may not fully address all potential confounding factors that could influence cortical thickness measurements. Factors such as age-related cortical changes or comorbid conditions could potentially impact the accuracy of the predictions, highlighting the need

							structural neuroimaging as an effective tool for early detection and prediction of Alzheimer's disease. Leave-One-Out Validation Technique to prevent bias: A leave-one-out (LOO) validation strategy was employed, wherein, for each comparison, all subjects, with the exception of one, were employed to select features and construct a classification model. Subsequently, the excluded subject was utilized for testing. This process was iterated for every subject within the two compared groups, thereby validating the method across all subjects. By excluding the test subject from feature selection and classifier construction, any potential bias or "double dipping" in predictive efforts for converters was effectively avoided.	for comprehensive consideration of various contributing factors in the analysis. Potential Overestimation of Prediction Accuracies: The study highlights the risk of artificially inflating prediction accuracies by including the subject under analysis in the generation of discriminant features. This emphasizes the importance of avoiding "double dipping" in the estimation of classifications and predictions based on cortical thickness statistical maps. Trade-off Between Sensitivity and Specificity: While the specificity of predicting MCI to AD conversion within 3 years is relatively high (84%), the sensitivity is relatively low (64%). This trade-off suggests challenges in achieving both high sensitivity and specificity, which are essential for clinically applicable predictions.
18	Hierarchical Feature Representation and Multimodal Fusion with Deep Learning for AD/MCI Diagnosis	NeuroImage	2014	Deep Boltzmann Machine Model	Alzheimer's Disease Neuroimaging Initiative (ADNI)	AD: 93 NC: 101 MCI: 204	Feature Extraction: A patch-based approach is employed, positioned as an intermediary level between a voxel-based approach and an ROI-based approach. This strategy proves efficient in addressing concerns related to high feature dimensions and sensitivity to subtle changes. Also patch-based	Visualization of Trained Weights: From a clinical standpoint, interpreting resulting feature representations, particularly in investigating neurodegenerative diseases like AD or MCI, is challenging. The method lacks utility in providing clinically relevant information. Exploring the extension of the

		approaches adeptly manages region-wide pathologies, extending beyond specific ROIs. This aligns with the perspective of neurologists or radiologists, who analyze images by investigating local patterns and subsequently integrating distributed local information across the entire brain to formulate clinical decisions. Multi Modal Data Fusion: The paper explores the fusion of multiple modalities like MRI and PET. This is done through analysis of their inherent shared features. Use of DBM: Deep Boltzmann Machine model can hierarchically find feature representations in a probabilistic manner. Rather than using the noisy voxel intensities as features the high-level representation obtained via DBM is more robust to noises and thus helps enhance diagnostic performances. Meanwhile, from a multimodal data fusion perspective, unlike the conventional multimodal feature combination methods that first extract modality-specific features and then fuse their complementary information during classifier learning, the proposed multimodal DBM fuses the complementary information from different modalities during a feature	proposed method to identify brain abnormalities in terms of regions or areas is suggested for easier comprehension by clinicians. Small Dataset for Deep Learning Model: In experiments, the number of hidden units in each layer was manually determined, and relatively small data samples (93 AD, 76 MCI-C, 128 MCI-NC, and 101 NC) were used. Consequently, the network structures employed for discovering high-level feature representations may not be optimal. Emphasizing the need for more intensive studies, such as learning optimal network structures from larger datasets, is highlighted for practical implementation of deep learning in clinical settings. Fusion of Different Modalities: The current method solely considers the bi-modalities of MRI and PET. Acknowledging the benefits of combining multiple modalities for richer information, a more systematic model is deemed necessary to efficiently find and utilize complementary information from genetics, proteomics, imaging, cognition, disease status, and other phenotypic modalities.

							representation step. Also in such a multimodal data fusion method, the methodological characteristic of the DBM, allows the bidirectional information flow from one modality (e.g., MRI) to the other modality (e.g., PET) and vice versa. Therefore, feature representations can be distributed over different layers in the path between modalities and thus efficiently discover a shared representation while still utilizing the full information in the observations.	Dataset Limitations: Recent studies indicate subjective cognitive complaints as a significant genetic risk factor for progression to MCI or AD. In the ADNI dataset, however, relevant information is lacking. Consequently, in experiments, the NC group may include both genuine controls and those with subjective cognitive complaints.
19	Inter-modality Relationship Constrained Multi-modality Multi-task Feature Selection for Alzheimer's Disease and Mild Cognitive Impairment Identification	NeuroImage	2014	Multi-modality Multi-task Feature Selection, Inter-modality Relationship Constraint	Alzheimer's Disease Neuroimaging Initiative (ADNI)	AD: 51 MCI: 99 NC: 52	Preservation of Inter-Modality Relationships: The method introduces a novel multi-task feature selection approach that considers feature selection from each modality as a separate task. Importantly, it imposes a constraint to preserve the inter-modality relationship, ensuring that different yet complementary information from various modalities is not overlooked. This enables the model to capture a more comprehensive and synergistic representation of the data. Enhanced Classification Performance: The proposed method achieves superior performance compared to state-of-the-art classification methods. The accuracy rates of 94.37% and 78.80% for AD and MCI identification, respectively, along	Requirement for Equal Feature Numbers Across Modalities: The proposed feature selection method necessitates that each modality provides the same number of features. This poses a limitation, particularly when dealing with modalities, such as CSF and genetic data, in the ADNI database, which may have different feature counts. This limitation restricts the method's immediate applicability to datasets with varying feature dimensions. Potential Inclusion of Additional Modalities in Future Work: The study acknowledges the absence of certain modalities, such as CSF and genetic data, in the current method. While the intention is to extend the method to include more

							with high area under the ROC curve (AUC) values, highlight the effectiveness of the approach in accurately classifying Alzheimer's Disease and mild cognitive impairment. Applicability to MCI Conversion Prediction: The method extends its utility to predicting the conversion of Mild Cognitive Impairment (MCI) to Alzheimer's Disease. With an accuracy of 67.83% and an AUC of 0.6957 for distinguishing between MCI converters and non-converters, the proposed method addresses a clinically significant aspect, showcasing its versatility in handling different diagnostic tasks related to AD.	modalities in future work, the current limitation may result in an incomplete representation of pathological information available in these additional data sources. Need for Testing on Completely Independent Datasets: Despite using cross-validation to evaluate generalizability, the study acknowledges the importance of testing on a completely independent dataset. The lack of such testing introduces a potential limitation in establishing the method's performance and reliability across diverse datasets and real-world applications.
20	View-Centralized Multi-Atlas Classification for Alzheimer's Disease Diagnosis	Human Brain Mapping	2014	View-Centralized Multi-Atlas Classification	Alzheimer's Disease Neuroimaging Initiative (ADNI)	AD: 97 pMCI: 117 sMCI: 117 NC: 128	Ensemble Classification Strategy: The study proposes an ensemble classification method by combining results from multiple classifiers corresponding to multiple atlases. This ensemble strategy, including PC, Lasso, and the proposed VCMA method, consistently performs better than other methods, highlighting the effectiveness of ensemble approaches in boosting classification results based on multiatlas data. Robustness to Parameter Variations: The VCMA method demonstrates robustness to parameter variations,	High Computational Cost: The use of multiple atlases for image registration contributes to a high computational cost in the VCMA method. This limitation should be considered, especially in scenarios where computational efficiency is crucial. Limited Feature Representation: The study only extracts regional features for feature representation, neglecting other morphometric features such as Jacobian determinants. Incorporating a broader set

							indicating that its performance is not highly sensitive to the selection of parameter values. This enhances its applicability and ease of use. Effective Feature Selection: The VCMA method introduces a feature selection approach that focuses on one atlas at a time, addressing redundancy in features extracted from multiple atlases. This effective feature selection contributes to improved diagnostic power, surpassing the performance of compared methods.	of features could enhance the method's capability. Single-Modality Data: The VCMA method relies solely on MRI data for learning the classification model, overlooking the potential benefits of incorporating other biomarkers such as FDG-PET. The inclusion of multiple biomarkers could potentially enhance the overall learning performance. Baseline Data Only: The experiments utilize only the MRI baseline data from the ADNI dataset. Future work could involve incorporating both baseline and longitudinal data to capture spatiotemporal development patterns of brain atrophy, improving the diagnosis and prediction of brain diseases.
21	Machine Learning Framework for Early MRI-based Alzheimer's Conversion Prediction in MCI Subjects	NeuroImage	2015	Semi-Supervised Learning, Novel Random Forest-Based Data Integration Scheme	Alzheimer's Disease Neuroimaging Initiative (ADNI)	AD: 200 NC: 231 pMCI: 164 sMCI: 100 uMCI: 130	Use of Semi Supervised Learning: Most of the earlier studies were based on supervised learning methods, where only labeled data samples are used for learning the model. Semi-supervised learning (SSL) approaches are able to use unlabeled data in conjunction with labeled data in a learning procedure for improving the classification performance.	Lack of Validation on External Datasets: The study primarily focuses on internal validation or cross-validation within the same dataset. It would be beneficial to validate the model's performance on external datasets to assess its generalizability to different populations and imaging protocols. Clinical Interpretability: While the paper emphasizes the accuracy of

							Feature Selection: A feature selection was conducted on MRI data derived from both AD subjects and normal controls, excluding data from MCI subjects. This was achieved through regularized logistic regression. Removal of Aging Effects: The aging effects within the MRI data were eliminated before the training of the classifier to avoid potential confounding arising from age-related atrophies. Incorporation of Cognitive Measurements with MRI Scans: An aggregate biomarker was formulated by initially deriving a distinct MRI biomarker and subsequently integrating it with age and cognitive measures pertaining to MCI subjects at the baseline, accomplished through the application of a random forest classifier. The added value of these innovative features in predicting the conversion from MCI to AD was empirically demonstrated using data acquired from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database.	the prediction model, it does not delve into the clinical interpretability of the features or factors contributing to the predictions. Understanding the biological or anatomical basis of the predictions can enhance the clinical utility of the model. Longitudinal Data: The paper does not discuss the utilization of longitudinal MRI data. Cognitive decline and Alzheimer's disease progression are dynamic processes, and incorporating longitudinal information may improve prediction accuracy. Feature Importance and Explainability: The paper does not provide insights into which features or regions of interest contributed most to the predictions. Understanding feature importance could aid in the clinical interpretation of the results.
22	Early Diagnosis of Alzheimer's Disease with Deep Learning	IEEE Journal of Bio-Medical and Health Informatics	2015	Stacked Sparse Auto-Encoders, Softmax Regression	Alzheimer's Disease Neuroimaging Initiative (ADNI)	AD: 65 MCI: 169 NC: 77	Semi Supervised Non-Linear Model: SVM Based Classification reduces Alzheimer’s prediction to binary classification problem. Similarly,	Longitudinal Analysis: Alzheimer's disease is a progressive condition, and the paper does not discuss whether the deep learning

		some methods embed prior knowledge into the data but the dependence of prior knowledge may be also sensitive to the changes of the dataset and hard to configure. Targeting at the constraints in previous studies, we believe the existing workflows can be efficiently optimized. This paper proposed a novel early diagnosis method for AD based on a deep learning architecture, consisting of stacked sparse auto-encoders and a soft-max regression layer. The proposed method has a multi-class nature and could reduce the reliance on prior knowledge about the data. Furthermore, this method is semi-supervised that can be extended to use unlabelled training samples, which are easier and cheaper to obtain. Lucky Trail Avoidance: To maximally avoid the 'lucky trails', the training and testing instances from each class, were randomly sampled to ensure they have similar distributions as the original dataset. For all methods in each fold of cross validation, about 90% subjects were used for training (including the pre-training of the deep neural nets) and the rest subjects were used for testing.	model can effectively track disease progression over time. Data Variability: Deep learning models are highly data-dependent, and the effectiveness of the approach may vary depending on the diversity and quality of the dataset. The paper does not discuss how the model performs on different datasets or the impact of data variability.

							Reservation of Synergies between different Modalities: The proposed method performs dimensionality reduction and data fusion at the same time to reserve the synergy between data modalities.	
23	Fully Convolutional Networks for Semantic Segmentation	Proceedings of the IEEE conference on Computer Vision and Pattern Recognition (CVPR)	2015	Fully Convolutional Networks (FCNs)	PASCAL VOC 2011, 2012	Not Specified	FCNs enable pixel-wise semantic segmentation: Unlike traditional methods that classify images into pre-defined categories, FCNs provide the ability to assign a label to each pixel in the image. This capability allows for detailed understanding and analysis of the image content, leading to more precise segmentation results. End-to-end learning for dense prediction tasks: FCNs facilitate end-to-end learning, enabling the network to directly optimize the segmentation task. This streamlined learning process avoids the need for separate feature extraction and classification steps, leading to more efficient and accurate predictions. Utilizes deconvolutional layers for Upsampling: FCNs incorporate deconvolutional layers that allow the network to learn to increase the spatial resolution of the feature maps. By using these layers, FCNs can	Memory-intensive; may require substantial GPU resources: FCNs can be computationally demanding and memory-intensive, particularly when processing high-resolution images. This characteristic necessitates significant computational resources, particularly high-end GPUs, which can increase the cost and hardware requirements for implementing FCNs in real-world applications. Limited context for large objects in early layers: FCNs may encounter challenges in capturing the complete context of large objects, particularly in the early layers of the network. As a result, the segmentation of such objects may not be as accurate or detailed as desired, potentially leading to misclassifications or incomplete segmentation results. Fine-grained details may still be challenging: Although FCNs excel

							efficiently produce segmentation maps with pixel-level accuracy, enabling better visualization and understanding of the intricate details within the image.	at capturing the overall semantic information in an image, they may encounter difficulties in accurately capturing fine-grained details within the image. This limitation can impact the precise delineation of intricate structures or objects with subtle visual variations, potentially leading to inaccuracies or loss of important information during the segmentation process.
24	Detection of Subjects and Brain Regions Related to Alzheimer's Disease Using 3D MRI Scans Based on Eigenbrain and Machine Learning	Frontiers in Computational Neuroscience	2015	Eigenbrain Analysis, Machine Learning	Open Access Series of Imaging Studies (OASIS)	126 subjects (98 NCs and 28 ADs)	Utilization of Eigenbrain and machine learning for the detection of subjects and brain regions associated with Alzheimer's disease: The study leverages the potential of Eigenbrain analysis and machine learning techniques to identify specific subjects and brain regions that may be indicative of Alzheimer's disease. By employing advanced analytical methodologies, the research can potentially detect subtle patterns and variations in the brain's structural characteristics, aiding in the early identification and diagnosis of individuals at risk of developing Alzheimer's disease. Provides insights into the application of 3D MRI scans and advanced analytical techniques for early identification of Alzheimer's disease: By emphasizing the use of 3D MRI scans in conjunction with advanced	Two-Dimensional Behaviour of Eigenbrain: Eigenbrain is essentially two-dimensional, which does not reduce the redundancy along the slice direction. Computationally Intensive: There is a need of preprocessing for spatial registration, which costs large amount of computation resources.

							analytical techniques, the research offers valuable insights into the potential application of sophisticated imaging modalities for the early detection of Alzheimer's disease. This approach facilitates a more comprehensive and detailed analysis of the brain's structural features, enabling the identification of specific biomarkers and patterns associated with the onset and progression of Alzheimer's disease. Advantages of Eigenbrains: The advantages of eigenbrain are threefold: (i) it reaches very high classification accuracy, which was better than or competitive with state-of-the-art methods; (ii) it can directly find discriminant voxels/regions within the whole brain; (iii) it can be combined with other features, in order to increase the classification performance.	
25	Gaussian process classification of Alzheimer's disease and mild cognitive impairment from resting-state fMRI	NeuroImage	2015	Bayesian Gaussian Process Logistic Regression (GP-LR)	Independent Study	AD: 27 aMCI: 50 NC: 39	Innovative Approach: Introduces a fresh method for patient stratification from resting-state fMRI scans, targeting the early phases of Alzheimer's disease (AD). Diverse Methodology: Incorporates the Gaussian process logistic regression (GP-LR) model alongside support vector machines (SVMs),	Sample Size and Data Balance: Small sample size in relation to tested features increases the risk of overfitting. Non-uniform distribution of educational backgrounds and gender ratios among groups introduces potential confounding variables. Motion Artifacts and Local Atrophy Effects: Potential impact of

							providing a comprehensive approach for neuroimaging studies. Principled Predictions and Customizability: Offers principled estimates of predicted class membership and customizable classification thresholds, enhancing adaptability for clinical decision-making. Insightful Analysis: Provides valuable insights into the key data features crucial for understanding AD diagnosis.	motion artifacts and local gray matter loss on functional connectivity requires further investigation for improved preprocessing methods. Diagnostic Challenges and Misclassifications: Lack of post-mortem confirmation and potential overlap with other conditions like major depression pose challenges in accurate differentiation between various neurodegenerative disorders.
26	DeepAD: Alzheimer's Disease Classification via Deep Convolutional Neural Networks Using MRI and fMRI	BioRxiv	2016	Deep Convolutional Neural Networks, MRI, fMRI	Alzheimer's Disease Neuroimaging Initiative (ADNI)	AD: 263 NC: 183	Robust Pipelines: The method involves the development of robust pipelines that incorporate extensive preprocessing modules and deep learning-based classifiers. Structural and Functional MRI Data Integration: The approach utilizes both structural and functional MRI data, allowing for a comprehensive analysis of brain abnormalities associated with Alzheimer's disease. Scale and Shift Invariant Features: The method employs a convolutional neural network architecture to extract scale and shift invariant low- to high-level features from a substantial volume of whole-brain	Practical Implementation Challenges: Implementing deep learning models, especially those involving complex neuroimaging data like MRI and fMRI, in clinical settings can present practical challenges. The computational requirements for training and deploying deep learning models can be significant, necessitating powerful computing resources. The hardware and infrastructure demands could potentially limit the widespread adoption of the proposed methodology in real-world clinical environments with limited computational resources. High Computation Requirements: The pipelines used to process the

			data, enhancing the model's robustness. Highly Accurate Predictive Model: The extracted features contribute to the creation of a highly accurate and reproducible predictive model for distinguishing Alzheimer's-affected brains from normal healthy brains in older adults. Superior Performance: The accuracy rates achieved for both MRI and fMRI modalities, as well as the use of state-of-the-art architectures like LeNet and GoogleNet, surpass the performance of all previous methods used for the same purpose. Incorporation of fMRI Data: The study pioneers the use of fMRI data to train a deep learning-based pipeline, extending the applicability of the method to new and valuable sources of information. Characterization of Multimodal MRI Biomarkers: The study showcases that the developed pipelines serve as effective algorithms for characterizing multimodal MRI biomarkers, contributing to a more comprehensive understanding of brain conditions.	modal data were executed on a GPU-based high performance computing platform.

							Potential for Disease Progression Prediction: The proposed methods exhibit strong potential for predicting the stages of Alzheimer's disease progression, offering valuable insights into the temporal evolution of the condition. Classification of Aging Effects: The methods also show promise in classifying the effects of aging in the normal brain, contributing to a better understanding of age-related changes in brain structure and function.	
27	Alzheimer's Disease Diagnostics by a Deeply Supervised Adaptable 3D Convolutional Network	arXiv Preprint	2016	Feature Extraction: 3D Convolutional Autoencoder Task Specific Classification: Deeply supervised target-domain-adaptable 3D-CNN	CADDementia Dataset	Not Specified	Domain Adaptability: The deep 3DCNN is designed to learn generic and transferable features across different domains. It effectively detects and extracts characteristic AD biomarkers in one domain (source) and performs task-specific classification in another domain (target). Combination of Networks: The network combines a generic feature-extracting stacked 3D-CAE, pre-trained in the source domain, with upper task-specific fully-connected layers. The lower layers capture generic features, while the upper layers are fine-tuned for domain-specific tasks in the target domain.	Single Imaging Modality: The proposed DSA-3D-CNN relies solely on a single imaging modality, namely structural Magnetic Resonance Imaging (sMRI). Limiting the method to a single modality might result in a lack of comprehensive information that could be obtained from the integration of multiple modalities, such as functional MRI (fMRI), Positron Emission Tomography (PET), or other imaging techniques. Omission of Skull-Stripping: The method does not perform skull-stripping as a preprocessing step. Skull-stripping is a common step in neuroimaging to remove non-

							Feature Extraction Capability: The 3D-CAE addresses feature extraction limitations of conventional approaches by automatically learning and extracting discriminative AD features that capture anatomical variations associated with AD. Adaptation to Different Datasets: Pre-trained convolutional filters of the 3D-CAE are adapted to another domain dataset, such as the Alzheimer's Disease Neuroimaging Initiative (ADNI), after initial pre-training on a different dataset (CADDementia). This adaptation allows the network to leverage pre-learned generic features for improved performance in specific tasks in the target domain. Deep Supervision for Adaptability: The incorporation of deep supervision allows for the effective adaptation of pre-learned generic features to specific tasks.	brain tissues and artifacts, and its omission may lead to the inclusion of irrelevant information or noise in the input data.
28	Topographical Information-Based High-Order Functional Connectivity and Its Application in Abnormality Detection for Mild	Journal of Alzheimer's Disease	2016	High-Order Functional Connectivity (HOFC)	Independent Study	MCI: 80 NC: 90	Increased Sensitivity to Group Differences: HOFC demonstrated higher sensitivity to detect group differences compared to conventional FC, providing more biologically meaningful results.	Dependency on Pearson's Correlation: The use of Pearson's correlation in HOFC has inherent limitations, such as neglecting time-domain information (e.g., phase synchrony, dynamic properties) and the inability to measure complex inter-regional interactions (e.g.,

	Cognitive Impairment	Individual Variability: HOFC better captured individual variability than FC, allowing for potential applications in individualized classification. Enhanced Modular Structure: HOFC-based network analysis revealed a more prominent modular structure compared to FC, aiding in better analysis of network organization. Prominent Group Differences: Group differences in modularity were more pronounced in the HOFC network, indicating its ability to detect subtle alterations in brain functional organization. Correlation with Behavioral Data: HOFC network properties showed correlations with behavioral data, suggesting its potential biological relevance. Olfactory Cortex and Frontal Cortex Connectivity Changes: HOFC revealed interesting connectivity changes between the olfactory cortex and the frontal cortex in MCI, indicating early pathology changes associated with neurodegeneration in AD.	modulation effects, partial correlation, mutual information). Future extensions could explore more sophisticated metrics to address these drawbacks. Two-Level Correlation Computation: The study only implemented a two-level correlation computation for HOFC. Future enhancements could involve adding more levels within the computational framework, incorporating time-varying, multi-frequency, and multimodal information for higher-order FC. Limited to fMRI Information: The HOFC analysis in this study solely relied on functional connectivity (FC) information extracted from fMRI. Future developments could naturally extend HOFC by integrating more features into the profile vector, allowing for the fusion of multi-channel information from various modalities, such as diffusion tensor imaging-based structural connectivity. Coarse Brain Parcellation: The study utilized a coarse brain parcellation atlas (AAL). Adopting a finer parcellation scheme could improve the accuracy of FC

							Modularity Configuration: MCI showed a higher diversity in modularity configuration, suggesting a more functionally segregated brain and potential compensatory effects in response to pathology.	estimation and extend the FC profile, enhancing the benefits of HOFC calculation. Integration of Additional Features: While the study focused on FC information, the integration of additional features into the profile vector for HOFC calculation was mentioned as a future consideration. This could involve incorporating information from diverse modalities, contributing to a more comprehensive understanding of high-level brain functional organization.
29	Multi-modal Classification of Alzheimer's Disease using Nonlinear Graph Fusion	Pattern Recognition	2017	Nonlinear Graph Fusion	Alzheimer's Disease Neuroimaging Initiative (ADNI)	AD: 37 MCI: 75 NC: 35	Non -Linear Fusion Method for Combining Multiple Modalities: State-of-the-art studies use linear methods to fuse the information from multiple modalities, which is not optimal for exploiting the complimentary information across modalities. The complementary information from multi-modal data is not necessarily linearly related. The authors presented a novel framework for multi-modality classification of AD using a nonlinear graph fusion method. Data Imputation to Expand Sample Size: The imputation approaches can fill the missing data of the excluded subjects so that it is likely to	Lack of Demographic Data: The data used in the paper lacks demographic information about the subjects. However, the demographic information could potentially provide complementary information to boost the classification performance. For example, it is reported that older subjects are more likely to develop AD than younger subjects. This means that age is an important predictor for classification of AD. Therefore, the classification performance could potentially be improved by adding age as an additional feature. Dataset Limitations: A subset of 147 subjects from ADNI were

						use as many samples as possible in the evaluation.	included in this study. The sample size is limited by the fact that four modalities were used in the work, which requires that data from four modalities were available for each subject. Many subjects were excluded due to the missing of their data from one modality or more. However, the excluded subjects could be potentially useful and provide additional information for classifier training. Lack of focus on Longitudinal Data: This paper only focuses on cross-sectional data. Interesting insights can be found on using longitudinal data with graph fusion.
30	Association of Elevated Amyloid Levels with Cognition in Preclinical Alzheimer's Disease	JAMA Neurology	2017	Exploratory Analysis	Alzheimer's Disease Neuroimaging Initiative (ADNI)	Identification of Preclinical AD: The study successfully Identification of Preclinical AD: The study successfully identifies a larger proportion of cognitively normal individuals with elevated brain amyloid at baseline who later developed cognitive symptoms. Dichotomizing participants into elevated vs. normal amyloid groups effectively separates those with progressive cognitive decline from those without, suggesting that preclinical Alzheimer's Disease (AD) may manifest in clinically normal individuals with elevated brain amyloid.	Infrequent Use of Antidementia Medications: The study acknowledges the infrequent but greater use of antidementia medications in the group with elevated amyloid during follow-up. This introduces a potential confounding factor, as these medications may have influenced the progression of cognitive decline, potentially impacting the observed differences between groups. Uncertain Clinical Importance of Group Differences: The study notes that group differences and changes on continuous measures

		Longitudinal Assessment: The study benefits from a long-term follow-up (up to 10 years) of the ADNI cohort, providing insights into the natural history of cognitive decline in relation to amyloid status. This extended follow-up allows for the observation of changes over an extended period and enhances the understanding of the trajectory of cognitive decline. Use of Composite Cognitive Measures: The study utilizes a modified version of the Preclinical Alzheimer Cognitive Composite (PACC), a cognitive composite designed for preclinical AD trials. The inclusion of PACC, MMSE, and Logical Memory tests contributes to a comprehensive assessment of cognitive function, enhancing the reliability of the findings. Biomarker Data Analysis: The study analyzes biomarker data, including CSF tau, pTau, and Aβ42, providing a comprehensive understanding of their associations with elevated brain amyloid. The longitudinal analysis of biomarkers reveals their sensitivity to elevated amyloid but suggests that they may not reflect cognitive and clinical decline once amyloidosis is established.	are of uncertain clinical importance. While statistical significance may be observed, the clinical relevance of these differences remains unclear. This limitation highlights the need for additional studies to establish the practical implications of the findings. Limited Number of Observations and High Loss to Follow-up: The study expresses concern about the limited number of observations at the latest time points and a high rate of loss to follow-up. This raises questions about the reliability of conclusions drawn from these latest time points and the possibility of unsupported extrapolations from earlier trends. However, sensitivity analyses with models imposing no assumptions about mean trajectory shape yielded similar conclusions. Need for Randomized Trials: The study recognizes that randomized trials would be necessary to assess whether interventions based on the findings affect the course of the disease. This limitation highlights the observational nature of the study, and the need for interventional studies to establish causal relationships.

		Association with Genetic Risk (APOE Genotype): The study explores the association between APOE genotype, amyloid accumulation, and cognitive decline. The presence of an APOEε4 allele is found to be associated with substantially increased cognitive decline, emphasizing the importance of genetic risk factors in preclinical AD. Support for Amyloid as a Critical Factor: The results support previous findings pointing to the critical role of amyloid in the neurobiology of AD. The study strengthens the link between elevated amyloid and primary manifestations of AD-related cognitive dysfunction.	Lack of Tau PET Imaging and Limited CSF Tau Data: The absence of tau PET imaging and limited collection of cerebrospinal fluid (CSF) tau data are acknowledged as limitations. Only 83% of participants had lumbar punctures at baseline, limiting the utility of CSF tau in the analysis. While ventricular volume was used as a covariate, the absence of direct tau measurements is a constraint in understanding the full spectrum of neurodegeneration. Absence of Baseline Cognition as a Covariate: The study notes that baseline cognition was not included as a covariate in the models. Instead, it was modelled as an outcome variable to illustrate the degree of separation at baseline. This approach may introduce complexities in fully accounting for baseline cognitive differences between groups. Exploratory Nature of Analyses: The study emphasizes the exploratory nature of analyses, highlighting that the analyses were not specified prior to data collection and the large number of comparisons carried out. This underscores

								the need for cautious interpretation of results and encourages further confirmatory studies.
31	Classifying MCI Subtypes in Community-Dwelling Elderly Using Cross-Sectional and Longitudinal MRI-Based Biomarkers	Frontiers in Aging Neuroscience	2017	Ensemble Voting Classifier made of SVM (rbf kernel), Logistic Regression and Random Forest.	Self Built Dataset using Patients from Sydney Memory and Aging Study	HC: 56 aMCI: 28 naMCI: 7	Classification Framework: The study employs a comprehensive classification framework for Mild Cognitive Impairment (MCI) subtypes using both cross-sectional and longitudinal MRI measurements. This approach allows for a more nuanced understanding of the dynamics of brain changes associated with MCI. Data-Resampling: The study addresses the challenge of class-imbalance through a data-resampling step in the classification framework. This enhances the reliability of the classification results, particularly when dealing with varying sample sizes across different cognitive states. Effective Differentiation: The study successfully demonstrates that individuals with amnestic MCI (aMCI) can be differentiated from cognitively normal (CN) and non-amnestic MCI (naMCI) using MRI-based biomarkers. The achieved classification accuracy, sensitivity, specificity, and AUC are reported to be superior to previous studies.	Limited Sample Size: The study acknowledges a limitation related to the sample size, particularly in the context of longitudinal data requirements. The restricted availability of subjects with MRI scans at both time points might impact the generalizability of findings. Population-Based Sample: The study is conducted on a population-based sample, which may introduce biases as it consists of more cognitively normal individuals than those with MCI. The difference in sample sizes between aMCI and naMCI is also noted as a potential limitation. Potential Double-Dipping Risk: The study acknowledges the risk of "double-dipping" when using the same dataset for both feature selection and classification. Careful separation of training and test datasets through cross-validation is implemented to mitigate this risk.

32	Residual And Plain Convolutional Neural Networks For 3D Brain MRI Classification	arXiv	2017	Residual And Plain 3D Convolutional Neural Network	Alzheimer's Disease Neuroimaging Initiative (ADNI)	AD: 50 LMCI: 43 EMCI: 77 HC: 61	End-to-End Models: The proposed deep learning algorithms for brain MRI classification offer end-to-end models, eliminating the need for complex multistep pipelines and handcrafted feature generation. Small Dataset Handling: Neuroimaging datasets are often small, posing a challenge for traditional neural network training. The convolutional neural networks (CNNs) introduced in this study can learn features efficiently even with limited data. Advanced Techniques: Leveraging modern advancements in deep learning, such as batch normalization and residual network architectures, mitigates issues associated with small training datasets while facilitating automatic feature generation. Applicability to 3D MRI Images: The proposed models can be directly applied to 3D MRI images without the need for intermediate handcrafted feature extraction.	Limited Exploration of Augmentation Techniques: Although suggesting data augmentation as a future research avenue, the study does not experiment with or provide details on specific augmentation techniques to enhance model robustness. Scalability Concerns: The study mentions a potential future goal of achieving similar or better results for unprocessed images. However, the scalability and computational efficiency of such models for large-scale deployment are not extensively discussed.
33	Association of Elevated Amyloid Levels with Cognition in Preclinical Alzheimer's Disease	JAMA Neurology	2017	Exploratory Analysis	Alzheimer's Disease Neuroimaging Initiative (ADNI)		Identification of Preclinical AD: The study successfully Identification of Preclinical AD: The study successfully identifies a larger proportion of cognitively normal individuals with elevated brain amyloid at	Infrequent Use of Antidementia Medications: The study acknowledges the infrequent but greater use of antidementia medications in the group with elevated amyloid during follow-up. This introduces

			baseline who later developed cognitive symptoms. Dichotomizing participants into elevated vs. normal amyloid groups effectively separates those with progressive cognitive decline from those without, suggesting that preclinical Alzheimer's Disease (AD) may manifest in clinically normal individuals with elevated brain amyloid. Longitudinal Assessment: The study benefits from a long-term follow-up (up to 10 years) of the ADNI cohort, providing insights into the natural history of cognitive decline in relation to amyloid status. This extended follow-up allows for the observation of changes over an extended period and enhances the understanding of the trajectory of cognitive decline. Use of Composite Cognitive Measures: The study utilizes a modified version of the Preclinical Alzheimer Cognitive Composite (PACC), a cognitive composite designed for preclinical AD trials. The inclusion of PACC, MMSE, and Logical Memory tests contributes to a comprehensive assessment of cognitive function, enhancing the reliability of the findings.	a potential confounding factor, as these medications may have influenced the progression of cognitive decline, potentially impacting the observed differences between groups. Uncertain Clinical Importance of Group Differences: The study notes that group differences and changes on continuous measures are of uncertain clinical importance. While statistical significance may be observed, the clinical relevance of these differences remains unclear. This limitation highlights the need for additional studies to establish the practical implications of the findings. Limited Number of Observations and High Loss to Follow-up: The study expresses concern about the limited number of observations at the latest time points and a high rate of loss to follow-up. This raises questions about the reliability of conclusions drawn from these latest time points and the possibility of unsupported extrapolations from earlier trends. However, sensitivity analyses with models imposing no assumptions about mean trajectory shape yielded similar conclusions.

	Biomarker Data Analysis: The study analyzes biomarker data, including CSF tau, pTau, and Aβ42, providing a comprehensive understanding of their associations with elevated brain amyloid. The longitudinal analysis of biomarkers reveals their sensitivity to elevated amyloid but suggests that they may not reflect cognitive and clinical decline once amyloidosis is established. Association with Genetic Risk (APOE Genotype): The study explores the association between APOE genotype, amyloid accumulation, and cognitive decline. The presence of an APOEε4 allele is found to be associated with substantially increased cognitive decline, emphasizing the importance of genetic risk factors in preclinical AD. Support for Amyloid as a Critical Factor: The results support previous findings pointing to the critical role of amyloid in the neurobiology of AD. The study strengthens the link between elevated amyloid and primary manifestations of AD-related cognitive dysfunction.	Need for Randomized Trials: The study recognizes that randomized trials would be necessary to assess whether interventions based on the findings affect the course of the disease. This limitation highlights the observational nature of the study, and the need for interventional studies to establish causal relationships. Lack of Tau PET Imaging and Limited CSF Tau Data: The absence of tau PET imaging and limited collection of cerebrospinal fluid (CSF) tau data are acknowledged as limitations. Only 83% of participants had lumbar punctures at baseline, limiting the utility of CSF tau in the analysis. While ventricular volume was used as a covariate, the absence of direct tau measurements is a constraint in understanding the full spectrum of neurodegeneration. Absence of Baseline Cognition as a Covariate: The study notes that baseline cognition was not included as a covariate in the models. Instead, it was modelled as an outcome variable to illustrate the degree of separation at baseline. This approach may introduce

								complexities in fully accounting for baseline cognitive differences between groups. Exploratory Nature of Analyses: The study emphasizes the exploratory nature of analyses, highlighting that the analyses were not specified prior to data collection and the large number of comparisons carried out. This underscores the need for cautious interpretation of results and encourages further confirmatory studies.
35	Classification of Alzheimer's disease and prediction of mild cognitive impairment-to-Alzheimer's conversion from structural magnetic resource imaging using feature ranking and a genetic algorithm	Computers in Biology and Medicine	2017	Feature Ranking and Genetic Alogorithm	Alzheimer's Disease Neuroimaging Initiative (ADNI)	AD: 60 MCI : 136 NC: 65	Advanced Feature Selection Method: The proposed method introduces an automatic feature-selection technique based on feature ranking and a Genetic Algorithm (GA), which aims to select the most discriminative features with minimal dimensionality. Effective High-Dimensional Pattern Recognition: The feature-selection approach is specifically designed for high-dimensional pattern analysis, making it suitable for complex neuroimaging studies with intricate spatial patterns in brain structure. Utilization of Fisher Criterion: The incorporation of the Fisher criterion in the Genetic Algorithm enhances the ability to find an optimal subset	Potential Algorithm Selection Bias: The study employs a Genetic Algorithm for feature selection, and while this is shown to be effective, the choice of meta-heuristic optimization algorithms could introduce bias. Consideration of other algorithms such as simulated annealing, particle swarm optimization, or ant colony optimization is suggested for future studies. Semi-Quantitative GM Atrophy Pattern: The study uses a semi-quantitative cortical circumferential system for GM atrophy pattern analysis, which may introduce subjectivity. More complex automated computational quantification methods could be explored to enhance objectivity.

							of features, ensuring maximum separation between different groups. Integration of Voxel-Based Morphometry (VBM) Analysis: The method incorporates VBM analysis to define a mask based on regions of gray matter (GM) atrophy from Alzheimer's Disease (AD) and Healthy Control (HC) subjects. This integration enriches feature extraction for Mild Cognitive Impairment (MCI) prediction. Consideration of Heterogeneity in MCI: Acknowledging the heterogeneity in Mild Cognitive Impairment (MCI), the study includes both progressive (pMCI) and stable (sMCI) MCI patients, recognizing the differences in health status between the two subgroups.	Limited Exploration of Other Modalities: The study primarily focuses on MRI data, and future studies could benefit from exploring the inclusion of other modalities such as positron emission tomography (PET), cerebrospinal fluid (CSF), and genetic information for a more comprehensive analysis. Absence of Longitudinal Data: The study does not involve longitudinal data, limiting the ability to capture changes over time. Including longitudinal data in future studies could enhance understanding of disease progression. Evaluation Only on Binary Classification: The study mainly focuses on binary classification tasks (AD vs. HC and MCI conversion prediction), and the extension to multiclass classification scenarios could provide a more comprehensive assessment.
36	Hippocampus and Amygdala Volume Estimation in Magnetic Resonance Images Using Deep Learning	Alzheimer's & Dementia	2018	Deep Learning	Alzheimer's Disease Neuroimaging Initiative (ADNI)	Not specified	Accurate estimation of hippocampus and amygdala volumes through deep learning techniques. Efficient analysis of large-scale MRI datasets for automated	Lack of information regarding specific deep learning architecture and training procedures.

							hippocampus and amygdala volume measurement.	
37	Multimodal Neuroimaging Feature Learning With Multimodal Stacked Deep Polynomial Networks for Diagnosis of Alzheimer's Disease	IEEE Journal of Biomedical and Health Informatics	2018	Multimodal Stacked Deep Polynomial Networks (MMSDPN)	Alzheimer's Disease Neuroimaging Initiative (ADNI)	AD: 51 MCI: 99 NC: 52	Effectiveness for Small Datasets: The proposed MM-SDPN algorithm demonstrates effectiveness, particularly for small datasets. This suggests potential utility in scenarios where limited data is available, showcasing adaptability to situations common in medical imaging studies. Fast Processing for Large-Scale Data: Due to the absence of forward and backward feedbacks between successive basic DPNs, SDPN, and MM-SDPN, these algorithms are relatively simple and fast. This simplicity positions them as promising candidates for handling large-scale data, ensuring efficiency in processing. Versatility in Learning Feature Representation: The planned application of MM-SDPN to learn feature representation directly from local patches of MRI and PET demonstrates the versatility of the algorithm. This adaptability to diverse data sources signifies its potential in handling multimodal neuroimaging information.	Limited Theoretical Foundation: DPN is a new DL algorithm with limited theoretical foundation and algorithmic development. This raises concerns about its robustness and generalizability, emphasizing the need for further theoretical advancements to establish its credibility. Effectiveness Primarily Demonstrated for Small Datasets: While MM-SDPN exhibits effectiveness for small datasets, its performance on larger datasets is assumed based on the simplicity and speed of the algorithm. This assumption requires empirical validation to ensure consistent efficacy across different data scales. Dependency on Future Algorithmic Improvements: The mention of future work to improve the DPN algorithm implies a dependency on algorithmic advancements for the method's overall efficacy. The success of the MM-SDPN and related frameworks is contingent on continuous improvements in the underlying DPN algorithm.

							Future Exploration of Semi-Supervised Learning: The intention to explore semi-supervised MM-SDPN acknowledges the practicality of acquiring unlabeled medical images, offering a pathway to enhance representation learning. This strategic approach aligns with the recognition that unlabeled data can contribute to performance improvement. Integration with MKL for Enhanced Classification: The proposed integration of Multiple Kernel Learning (MKL) with MM-SDPN for Alzheimer's Disease (AD) classification represents a sophisticated strategy. Leveraging the learned features from SDPN and individual features from MRI and PET enhances the classification framework, with the potential to further improve performance.	
38	Automated Detection of Amnestic Mild Cognitive Impairment in Community-Dwelling Elderly Adults: A Combined Spatial Atrophy and White Matter Alteration Approach	NeuroImage	2018	Combined Spatial Atrophy and White Matter Alteration Approach	A total of 1037 participants were drawn from the Sydney Memory and Aging Study (MAS), a longitudinal study of non-demented, community dwelling individuals aged 70–90 years old at baseline	79 patients with a clinical diagnosis of aMCI and 204 who were cognitively normal	Innovative Multimodal Approach: The method introduces a novel approach by incorporating measures of both spatial atrophy from T1-weighted images and white matter alterations assessed through Diffusion Tensor Imaging (DTI) tract-based spatial statistics (TBSS). This multimodal strategy enhances the comprehensiveness of the neuroimaging analysis.	Sample Selection Bias: The method's reliance on participants from the Sydney Memory and Aging Study (MAS) introduces a potential selection bias. The study population consists of non-demented, community-dwelling individuals aged 70–90 years at baseline, recruited randomly from specific areas in Eastern Sydney. This may limit the generalizability of findings to broader demographics.

		Advanced Feature Extraction Techniques: Subcortical volumetric features are extracted using a sophisticated FreeSurfer-initialized Large Deformation Diffeomorphic Metric Mapping (FS+LDDMM) segmentation approach. Additionally, fractional anisotropy (FA) values are obtained for white matter regions of interest. These advanced techniques contribute to a more nuanced and detailed representation of brain structure. Optimized Feature Selection with SVM: The method employs a Support Vector Machine (SVM) for feature selection, identifying an optimal subset of features ranked by their discriminative ability between individuals with amnestic Mild Cognitive Impairment (aMCI) and those with normal cognition. This ensures an efficient and focused feature set for training SVM classifiers, enhancing classification accuracy. Consideration of Potentially Confounding Factors: The study goes beyond neuroimaging features and identifies various sociodemographic, lifestyle, health, and other factors that may impact the classification of individuals. This comprehensive consideration adds a layer	Limited Generalization to Non-Scanner Subgroup: The subgroup with both T1-weighted and DTI scans (283 individuals) may not fully represent the entire cohort (1037 MAS participants). Findings from the neuroimaging analysis may not be directly generalizable to the larger group without imaging data, potentially limiting the broader applicability of the method's outcomes. Statistical Testing Challenges: Conducting multiple tests (ANOVA or chi-square) simultaneously for numerous factors raises concerns about inflated Type I error rates. The liberal significance level (pb0.10) for pair-wise comparisons may increase the risk of identifying statistically significant results by chance, impacting the reliability of the findings. Backward Stepwise Logistic Regression: The use of backward stepwise logistic regression introduces the risk of overfitting and may result in a model that fits the specific dataset too closely. This could limit the generalizability of the model to broader populations or datasets.

							of contextual understanding to the classification schema, addressing potential confounding variables.	
39	Aberrant Connectivity in Mild Cognitive Impairment and Alzheimer Disease Revealed by Multimodal Neuroimaging Data	Neuro-degenerative Diseases	2018	Multimodal SICE	Alzheimer's Disease Neuroimaging Initiative (ADNI)	AD: 116 MCI: 116 NC: 116	Identification of Neurobiological Changes: The study effectively identified the neurobiological changes between MCI and NC based on multimodal SICE, providing crucial insights into the progression of cognitive decline. Insight into Brain Network Connectivity: By analyzing connectivity patterns, the research highlighted the progressive weakening of connectivity in key brain regions, notably the temporal, temporal-parietal, and occipital-parietal lobes, contributing to a comprehensive understanding of the underlying mechanisms of MCI. Detection of Impaired Cognitive Regions: The study successfully detected significant declines in connectivity within the temporal lobe, highlighting the correlation between temporal lobe connectivity loss and cognitive decline, a key feature in the trajectory of MCI progression. Evidence of Compensatory Mechanisms: The research indicated potential compensatory mechanisms in MCI patients, particularly observed	Limited Neuroimaging Modalities: The study only considered a subset of neuroimaging modalities, including sMRI, FDG-PET, and florbetapir PET. Other potentially relevant data sources for distinguishing Alzheimer's disease (AD) or Mild Cognitive Impairment (MCI), such as cerebrospinal fluid (CSF), cognitive measures, and genomics, were not included. This limits the comprehensiveness of the analysis. Incomprehensive Multimodal Integration: The approach used to integrate multimodal imaging data may not be optimal. The study acknowledges that the method employed for integrating these modalities may have limitations, and there is a recognition that alternative techniques, such as weighted combination methods, could be explored for more effective multimodal integration. Lack of Intrasubject Linkage: The study did not consider intrasubject linkage of the imaging data. This means that the connection or relationship between different modalities within the same subject was

							in the increased connectivity between certain brain regions, offering valuable insights into the brain's adaptive processes in the face of cognitive challenges. Enhanced Diagnostic Capabilities: By integrating multiple neuroimaging modalities such as sMRI, FDG-PET, and florbetapir PET, the study provided more comprehensive and accurate imaging-based biomarkers, enabling improved differentiation between NC, MCI, and AD patients, leading to earlier and more accurate diagnoses.	not explicitly taken into account. In future studies, addressing this limitation and exploring methods to link intrasubject multimodal data could enhance the accuracy and reliability of classification. Potential Modality Weighting Issues: The weighted combination method used for multimodal integration may have its limitations. The study acknowledges that there may be alternative ways to combine multimodal data, and the chosen method may not be the most effective for the specific dataset. Exploring and comparing different weighting strategies could provide insights into the robustness of the results.
40	Learning Brain Connectivity Sub-networks by Group-Constrained Sparse Inverse Covariance Estimation for Alzheimer's Disease Classification	Frontiers in Neuro-informatics	2018	Leave-one-out Cross Validation, SICE	Independent Study Conducted in Xuanwu Hospital, Capital Medical University, Beijing, China	AD: 30 NC: 32	Enhanced Classification Accuracy: The proposed classification framework using fMRI time series significantly improved the diagnosis accuracy for distinguishing AD patients from NC, indicating its superiority over other methods. It showcases an improvement of at least 7.27% in diagnosis accuracy, highlighting the effectiveness of the sparse-based method in constructing brain networks compared to traditional fully-connected correlation-based networks.	Parameter Sensitivity: The study's reliance on a fixed SICE tuning parameter (λ) for different subjects may impact the classification performance due to varying optimal parameters across individuals. To address this, the study proposes subject-specific λ optimization using the BIC method, enabling the construction of optimal connectivity networks for each subject in future research. Atlas Relevance: Limitations arise from using a Chinese brain atlas

		Effective Biomarker Identification: The top 5 brain connections identified by the proposed method serve as promising connectivity-based biomarkers for the diagnosis of AD. These connections correspond to brain regions known to be significantly associated with AD pathology, such as the Cingulum_Post and regions within the Default Mode Network (DMN), aligning with existing research findings. This strengthens the potential applicability of the proposed method in the diagnosis of MCI patients, showcasing its robustness and versatility. Improved Network Construction: The method's integration of the Group-constrained topology structure detection algorithm with SICE aids in improving classification performance. By encouraging a consistent network topology across subjects, the method minimizes inter-subject variability issues, thereby enhancing the generalization performance of trained classifiers. This emphasizes the efficacy of the method in constructing efficient functional brain sub-networks, crucial for accurate classification. Reinforcement of Disconnection Hypothesis: The findings support the	for MRI analysis in a study involving Chinese participants, rather than the Caucasian atlas (SPM8). Incorporating a culturally specific brain atlas during image segmentation and registration is suggested to enhance the extraction of precise MRI features, improving the accuracy of diagnosing AD patients.

							disconnection hypothesis of AD, demonstrating the significance of the identified brain regions in DMN. The reported decline in synaptic numbers in regions such as the Cingulum_Post aligns with previous studies, reinforcing the proposed method's ability to detect both impairments and compensatory mechanisms within DMN. This not only adds to the existing body of knowledge but also establishes the potential of the method for diagnosing MCI patients in the future. Robustness in Detecting Pathological Changes: By highlighting the connectivity changes within DMN, the method proves its robustness in detecting pathological changes, even at the MCI stage of AD. This further underscores its relevance and potential for early diagnosis and intervention, establishing it as a valuable tool in the assessment and management of cognitive impairments.	
41	Dual-Model Radiomic Biomarkers Predict Development of Mild Cognitive Impairment Progression to Alzheimer's Disease	Frontiers in Neuroscience	2019	Dual-Model Radiomic Analysis with Multivariate Cox Proportional Hazards Regression Model	Alzheimer's Disease Neuroimaging Initiative (ADNI)	cMCI: 131 ncMCI: 132	Prediction Enhancement: Radiomic analysis, combined with Cox models, enhances the prediction of MCI (Mild Cognitive Impairment) conversion to AD (Alzheimer's Disease), offering valuable prognostic insights.	Short Follow-Up Period: The study's relatively short 3-year follow-up for individuals with Mild Cognitive Impairment (MCI) may limit the ability to capture long-term changes and progression to Alzheimer's Disease (AD),

		Comprehensive Topography: Identification of MCI conversion-related regions, integrating structural atrophy and metabolic abnormalities, aligns with existing literature, providing a comprehensive topographical understanding. APOE ε4 Gene as a Predictor: The identification of APOE ε4 gene as a risk predictor in the clinical Cox model adds genetic information, contributing to a more comprehensive understanding of MCI conversion risk. Prognostic Stability: The fusion-modality Cox model exhibits higher Harrell's C and more stable relative risk, indicating increased stability and reliability in predicting the risk of MCI progression. Applicability to Single Cases: Radiomics analysis's ability to be applied on a single-case basis enhances its clinical utility, allowing for personalized predictions and facilitating individualized patient care. Quantitative Feature Set: The study's extraction of 172 radiomic features, including intensity, texture, and wavelet features, provides	potentially affecting the generalizability of findings. Incomplete Baseline Data: A subset of MCI images lacked baseline data for both MRI and FDG PET imaging, introducing potential biases and limiting the comprehensiveness of the multimodal analysis. Future studies with complete baseline data are warranted for improved model performance. Smoothing Step Impact: The preprocessing smoothing step may introduce variability in the calculation of features and the definition of ROIs. While common in Alzheimer's Disease (AD) studies, the impact of smoothing on the accuracy of features and ROIs should be considered, and future research may explore alternative, more accurate segmentation methods. Target Region Extraction Method: The study utilized a routine smoothing step in preprocessing, which may impact the extraction of target regions. Future research could benefit from more accurate methods, such as manual segmentation, especially in the context of oncological radiomics studies, to enhance precision.

							a robust quantitative foundation for analysis and prediction. Texture Analysis Value: The significance of top quantitative features, such as entropy, complexity, and coarseness, underscores the value of texture analysis in capturing complex brain structural alterations associated with MCI conversion.	Impact of APOE ε4 Gene: Although the study identifies the APOE ε4 gene as a risk predictor, the specific impact and interactions with other predictors are not thoroughly explored. Future research could delve into the nuanced relationships between genetic factors and radiomic features in the context of prediction models.
42	Early diagnosis of Alzheimer's disease using combined features from voxel-based morphometry and cortical, subcortical, and hippocampus regions of MRI T1 brain images	PLoS ONE	2019	Feature Extraction: Voxel-Based Morphometry; Classification: SVM, Random Forest, KNN	NRCD dataset (private dataset which was generated in Chosun University hospitals)	163 (sMRI)	Improved Classification Performance: Combining various structural MRI features enhanced the classification accuracy compared to using individual features. The method achieved good AUC and ACC values, indicating robust performance. Powerful and Steady Classifier: The combination of VBM, CSC, and HV features resulted in a more powerful and steady classifier than using a single feature. Optimized Hyperparameter Tuning: The study employed a rigorous approach to hyperparameter tuning using a grid search and five-fold stratified cross-validation, ensuring unbiased estimates of performance. High Agreement Levels:	Preliminary Nature: The study is acknowledged as a preliminary proof-of-concept, indicating that further replication and validation are needed to solidify its findings. Single Modality: The proposed method relies on structural MRI (sMRI) modality, and the effectiveness with other imaging modalities like PET and functional MRI remains unexplored.

							The proposed model demonstrated high agreement levels between different classification groups, as indicated by Cohen's kappa values. Novel Feature Fusion Technique: The introduction of a novel feature fusion technique, combining morphometric features with cortical and hippocampal volume features, contributed to improved classification accuracy.	
43	Alzheimer's Disease Diagnosis Based on Cortical and Subcortical Features	J Healthc Eng	2019	Softmax Classifier, SVM, KNN, and naïve Bayes	NRCD Dataset and OASIS	326	Use of Combined Features: The combination of cortical thickness and subcortical volume features proved effective, showcasing the importance of using multiple features for robust classification. Multiple Classifiers Tested: The study employed four different classifiers (softmax, SVM, KNN, naïve Bayes) for comprehensive evaluation, providing insights into the strengths of each in different classification scenarios. Applicability to Tertiary Group Classification: The proposed technique demonstrated success in classifying a tertiary group (AD vs HC vs mAD), showcasing its potential for handling more complex classification scenarios.	Limited Feature Set: The study relied only on cortical thickness and subcortical volume features, potentially limiting the diversity and richness of information for classification. Need for Longitudinal Data: The study's focus on cross-sectional data may limit its ability to capture temporal changes over time. Future work using longitudinal datasets could enhance understanding of disease progression. Classifier Dependency: While the RBF-SVM classifier performed well in several cases, the choice of classifier might be dataset-dependent, and the robustness across different datasets and scenarios should be explored.

							Effective on External Datasets: The model's performance was not only validated on the NRCD dataset but also demonstrated effectiveness when applied to the OASIS dataset, suggesting potential generalizability.	No Exploration of Hyperparameter Tuning: The study does not provide details on hyperparameter tuning for classifiers, and optimal parameter settings could significantly impact performance. Assumption of Homogeneous Data: The study assumes homogeneity within each diagnostic group, which may not fully represent the heterogeneity present in real-world clinical populations.
44	Decoding Brain Functional Connectivity Implicated in AD and MCI	Medical Image Computing and Computer Assisted Intervention	2019	Feed-Forward Deep Neural Network	Alzheimer's Disease Neuroimaging Initiative (ADNI)	Not Specified	Efficient Feature Reduction: The recursive elimination of low-relevance features effectively addresses the challenge of handling fMRI data with a high feature-to-instance ratio. Leaner DNN Design: The approach results in a leaner Deep Neural Network (DNN), optimizing the model for fMRI classification tasks. State-of-the-Art Performance: Achieving state-of-the-art classification accuracy for MCI/AD and CN/AD, and comparable accuracy for CN/MCI classification highlights the effectiveness of the proposed method. Biological Interpretability: The identified important brain regions align with previous studies, enhancing	Model Complexity: The complexity of the 5-layer feedforward DNN might limit interpretability and generalization, and alternative architectures could be explored. Dependency on Feature Relevance Scores: The effectiveness relies on the accuracy of feature relevance scores, which could be influenced by the choice of the reference-based decoder. Clinical Validation: While the method shows promise for biomarker detection, its clinical utility should be validated through further studies and real-world applications. Interpretability Challenges: The interpretability of DNNs remains a

							the biological interpretability of the results. Potential for Biomarker Detection: The method holds promise for biomarker detection in various neurological ailments, supporting computer-aided detection applications.	challenge, and understanding the clinical implications of identified features requires additional research.
45	Early Detection of Alzheimer's Disease Using Magnetic Resonance Imaging: A Novel Approach Combining Convolutional Neural Networks and Ensemble Learning	Frontiers in Neuroscience	2020	Convolutional Neural Networks, Ensemble Learning	Alzheimer's Disease Neuroimaging Initiative (ADNI)	509 subjects AD: 137 NC: 162	Application of Data Augmentation Techniques to Enhance the Dataset: To address the potential over-fitting issue in training resilient CNN models and to integrate potential image disparities, augmented images were created from the original slices using six operations: rotation, translation, gamma correction, random noise addition, scaling, and random affine transformation. These augmented data were then incorporated into the initial training dataset to ensure an adequately large sample size. The utilization of data augmentation also served to alleviate the initially imbalanced dataset (for instance, there were more subjects with MCInc than those with MCIc). The predefined number of augmented slices to be generated varied from class to class based on the specific dataset imbalance. Integration of Convolutional Neural Networks and Ensemble Learning for early detection of Alzheimer's	Lack of detailed information on the specific architecture and parameters of the utilized Convolutional Neural Networks: While the study integrates Convolutional Neural Networks, it does not provide comprehensive insights into the specific architecture, hyperparameters, or training procedures employed. This limitation may hinder the reproducibility and further optimization of the proposed methodology by other researchers, potentially limiting the broader application of the approach in different research settings.

			disease using MRI data: The study's innovative approach combines the strengths of Convolutional Neural Networks (CNNs) and Ensemble Learning techniques to enhance the accuracy and efficacy of early detection of Alzheimer's disease based on Magnetic Resonance Imaging (MRI) data. This integrated methodology allows for more comprehensive analysis and identification of complex patterns associated with the disease, potentially leading to improved diagnostic accuracy and earlier intervention. Automatic Selection of ROIs: The discussed method does not require manual selection of ROIs, but automatically extracts the discriminable features from the MR images using a CNN-based adaptive representation learning method in a data-driven way. The proposed method also employs a two-stage EL scheme to improve generalization and robustness. Ability to Detect Other Neurological Problems: The advocated method may be useful for identifying additional candidate neuroimaging biomarkers for AD as well as for other brain diseases such as Parkinson's disease, autism, schizophrenia and

							severe depression, especially for identifying candidate neuroimaging biomarkers for other little-known brain disorders, in a data-driven way.	
47	Preclinical Detection of Alzheimer's Disease Using FDG-PET, with or without Amyloid Imaging	Journal of Alzheimer's Disease	2020	FDG-PET, with or without Amyloid Imaging	Review Paper	-	High Sensitivity for Disease Discrimination: FDG-PET exhibits high sensitivity in distinguishing Alzheimer's Disease (AD) from both healthy controls and other neurodegenerative diseases. It serves as a valuable tool for identifying individuals at higher risk for AD. Quantitative and Topographical Correlation: The method offers good quantitative and topographical correlation with clinical progression. This strength enhances its utility in tracking disease-related changes and understanding the spatial distribution of metabolic abnormalities. Potential for Risk Stratification: FDG-PET's ability to differentiate individuals at higher versus lower AD risk enhances its role in stratifying risk levels, aiding in early identification and intervention.	Absence of Postmortem Data: A major limitation lies in the absence of postmortem data in most FDG-PET studies. This hinders the confirmation of clinical symptoms and reductions in cerebral metabolic rate of glucose (CMRglc) as solely attributable to AD pathology, raising uncertainties about the specificity of findings. Reliance on Clinical Diagnosis: The use of clinical diagnosis as the gold standard introduces a potential limitation, as it may result in the inclusion of patients with a dementia other than AD in the AD group and vice versa. This reliance on clinical diagnosis raises the risk of misclassification. Hypometabolism Not Exclusive to AD: In asymptomatic subjects with hypometabolism, CMRglc deficits may arise from causes other than AD pathology. Additionally, not all individuals with hypometabolism may necessarily progress to AD, introducing ambiguity in the

								interpretation of FDG-PET findings. Need for Imaging of AD Pathology: The authors emphasize the essential role of imaging AD pathology in resolving uncertainties. This indicates a dependence on complementary imaging modalities to provide a more definitive understanding of the underlying pathology contributing to hypometabolism observed in FDG-PET studies.
48	Preclinical Detection of Alzheimer's Disease Using FDG-PET, with or without Amyloid Imaging	Journal of Alzheimer's Disease	2020	FDG-PET, with or without Amyloid Imaging	Review Paper	-	High Sensitivity for Disease Discrimination: FDG-PET exhibits high sensitivity in distinguishing Alzheimer's Disease (AD) from both healthy controls and other neurodegenerative diseases. It serves as a valuable tool for identifying individuals at higher risk for AD. Quantitative and Topographical Correlation: The method offers good quantitative and topographical correlation with clinical progression. This strength enhances its utility in tracking disease-related changes and understanding the spatial distribution of metabolic abnormalities. Potential for Risk Stratification: FDG-PET's ability to differentiate individuals at higher versus lower AD risk enhances its role in stratifying	Absence of Postmortem Data: A major limitation lies in the absence of postmortem data in most FDG-PET studies. This hinders the confirmation of clinical symptoms and reductions in cerebral metabolic rate of glucose (CMRglc) as solely attributable to AD pathology, raising uncertainties about the specificity of findings. Reliance on Clinical Diagnosis: The use of clinical diagnosis as the gold standard introduces a potential limitation, as it may result in the inclusion of patients with a dementia other than AD in the AD group and vice versa. This reliance on clinical diagnosis raises the risk of misclassification.

							risk levels, aiding in early identification and intervention.	Hypometabolism Not Exclusive to AD: In asymptomatic subjects with hypometabolism, CMRglc deficits may arise from causes other than AD pathology. Additionally, not all individuals with hypometabolism may necessarily progress to AD, introducing ambiguity in the interpretation of FDG-PET findings. Need for Imaging of AD Pathology: The authors emphasize the essential role of imaging AD pathology in resolving uncertainties. This indicates a dependence on complementary imaging modalities to provide a more definitive understanding of the underlying pathology contributing to hypometabolism observed in FDG-PET studies.
49	Computer aided Alzheimer's disease diagnosis by an unsupervised deep learning technology	Neurocomputing	2020	PCANet	Alzheimer's Disease Neuroimaging Initiative (ADNI)	AD: 243 MCI: 525 NC: 307	High Prediction Accuracy: The proposed method achieves high prediction accuracy, particularly for AD vs. MCI (97.01%) and AD vs. NC (89.15%), demonstrating its effectiveness in Alzheimer’s disease (AD) classification. Utilization of TOP Slices: Incorporating the TOP slices of MRI images significantly enhances classification accuracy (92.5%) compared to using a single slice, emphasizing the	Cluster Size Bias: The k-means clustering tendency to produce equal-sized clusters might lead to suboptimal results, especially for the AD vs. NC group, where the distributions are ellipse-shaped, indicating a limitation in handling certain data distributions. Dependency on Specific Views: The method's reliance on specific views (TOP slices) for improved accuracy might limit its generalizability to diverse datasets,

							importance of capturing anatomical structures for accurate predictions. Effective Feature Extraction: The use of PCANet for feature extraction contributes to the success of the method, showcasing its capability to learn discriminative features from different views of MRI images. Comparative Performance: Despite slight performance differences with some state-of-the-art methods, the proposed method demonstrates competitive accuracy on a larger dataset without data selection.	potentially hindering its applicability to different imaging protocols or populations. Comparison to Selective Databases: While the method achieves competitive results on the ADNI database, direct comparisons with some state-of-the-art methods involve selected databases, potentially affecting the generalizability of performance comparisons. Limited Modalities: The method utilizes only one modality (MRI), neglecting potential benefits from combining multiple modalities, as seen in some state-of-the-art approaches, limiting the method's scope in capturing diverse information sources.
50	Ambivert degree identifies crucial brain functional hubs and improves detection of Alzheimer's Disease and Autism Spectrum Disorder	NeuroImage: Clinical	2020	Using ambivert degree as input features, deep neural networks detect AD and ASD from healthy controls	Alzheimer's Disease Neuroimaging Initiative (ADNI) and Autism Brain Imaging Data Exchange (ABIDE)	ADNI AD: 29 NC: 49 ABIDE ASD: 73 NC: 88	Innovative Measures: Introduces novel measures like ambivert degree and gateway coefficient, accounting for the strength of connections and the uniqueness of intermodular connections. Preservation of Modular Structure: Addresses the impact of sparsification on network modularity, utilizing a thresholding scheme to maintain the modular structure during network analysis.	Dependence on Segmentation Maps: Acknowledges a dependence on the quality of segmentation maps for aggregating patch-based grading and estimating abnormality, which may introduce variability. Complexity of Interpretation: Introduces new measures, which, while innovative, might add complexity to the interpretation of results and require a thorough understanding of their implications.

							Clinical Relevance: Applies the method to investigate the disruption of brain hubs in Alzheimer's disease (AD) and Autism Spectrum Disorder (ASD), highlighting potential clinical applications. Informative Features for Classification: Demonstrates the effectiveness of hub scores as features for classifying AD and ASD subjects, offering insights into the diagnostic potential of hub disruptions.	Generalization: The study focuses on specific disorders (AD and ASD) and large datasets, and the generalizability of the proposed measures and methods to other disorders or smaller datasets remains to be explored. Impact of Lesions: While the study considers the effect of inducing artificial lesions in brain functional networks, the broader implications of this manipulation on the clinical relevance of hub disruptions need further exploration. Algorithm Sensitivity: The method's sensitivity to parameters and the potential need for optimization could impact its robustness across different datasets and experimental conditions.
51	Variationally Regularized Graph-based Representation Learning for Electronic Health Records	arXiv	2021	Encoder-Decoder Graph Neural Network	eICU Cohort, MIMIC-III Cohort, AD-EHR (Alzheimer’s Disease Prediction) Cohort obtained from inpatient and outpatient EHR data from NYU Langone Health.	Not Specified	Variational Regularization: Introduction of variational regularization for node representation learning addresses limitations of self-attention in graph-based models. This addresses challenges in constructing knowledge graphs manually from real-world noisy data, enhancing the model's adaptability. Improved Predictive Performance: The method's innovative design,	Potential Overfitting: The use of adaptive learning of connections and variational regularization may lead to a risk of overfitting to the specific characteristics of the training data. The method's performance on unseen data or different populations is not discussed. Exploration of Self-Supervised Learning:

							incorporating variational regularization on node representations in Graph Neural Networks (GNN), leads to superior performance compared to previous graph representation learning methods in health predictive tasks. This is demonstrated through evaluations on clinical EHR data and two public EHR datasets.	The text indicates that future studies will explore self-supervised learning to improve generalization. This implies that the current method might have limitations in terms of generalization, and additional techniques are being considered to address this.
52	Brain Age Estimation from MRI Images using 2D-CNN instead of 3D-CNN	Acta Infologica	2021	2D CNN	IXI Dataset	563 T1-Weighted NC	Efficient Plane-Based Approach: The study efficiently employs axial, coronal, and sagittal planes of brain scans, eliminating the need for complex 3D models and reducing computational demands. Optimized Model Utilization: Leveraging pre-trained DenseNet121 model weights mitigates the impact of a small dataset, enhancing model performance without intensive computational requirements. Reduced Training Time: Achieves a low Mean Absolute Error (MAE) of 6.3 in estimating brain age, with a notably short training time of 5.35 minutes, demonstrating efficiency in model training. Applicability for Neurodegenerative Disease Detection: The proposed method, focused on Brain Age Estimation (BAE), holds promise for	Limited Dataset Size: The study acknowledges a constraint in dataset size, potentially affecting the model's generalization to diverse populations or specific demographic groups. Dependency on Pre-trained Weights: Utilizing pre-trained model weights may introduce biases from the original dataset, limiting adaptability to the unique features of the brain scans in the current study. Simplification of Brain Structure: By focusing on specific planes, the study may overlook nuanced three-dimensional interactions in brain structures, potentially affecting the accuracy of age estimation. Task-Specific Application: While excelling in Brain Age Estimation, the method's applicability may be constrained to tasks directly

							detecting neurodegenerative diseases like Alzheimer's and Parkinson's. Comparable Performance: Despite its efficiency, the method yields results comparable to similar studies, emphasizing its effectiveness in Brain Age Estimation tasks.	related to estimating age from brain MR images. Optimization Dependency: The reported performance metrics are tied to specific configurations (sagittal planes, Adamax optimizer), and generalization to alternative setups requires validation.
53	Multi-scale graph-based grading for Alzheimer's disease prediction	Medical Image Analysis	2021	Multi Scale Graph-Based Grading Framework	Alzheimer's Disease Neuroimaging Initiative (ADNI)	AD: 130 MCI: 216 NC: 213	Efficient Combination of Variability and Similarity: The proposed method excels in efficiently combining intra-subject variability and inter-subject similarity within a common model. This ability enhances the adaptability of the model across different anatomical scales. Multi-Scale Graph-Based Grading: The utilization of a multi-scale graph-based grading framework is a significant strength. This approach allows for a comprehensive analysis that considers variations at different anatomical scales, providing a more nuanced understanding.	Dependency on Segmentation Quality: The major limitation stems from the method's dependence on the quality of segmentation maps. The accuracy and reliability of the proposed approach are directly influenced by the precision of segmentation, and any inaccuracies in segmentation maps can compromise the results. Sensitivity to Segmentation Errors: In cases where segmentation maps contain errors or inaccuracies, the proposed framework is likely to be sensitive to these issues. This sensitivity could lead to misinterpretations or misgradings, impacting the overall robustness of the method. Subject-Specific Variation: The method's performance may be affected by subject-specific variations in segmentation quality. If there are considerable variations in

			the accuracy of segmentation across different subjects, it could introduce inconsistencies in the grading framework. Applicability Across Diverse Populations: The method's generalizability might be limited when applied to diverse populations with varying anatomical characteristics. Anatomical variations between different groups may challenge the universality of the proposed framework. Complexity and Computational Demands: Depending on the computational demands of the multi-scale graph-based grading framework, there might be challenges related to computational resources. The complexity of the approach could limit its practical applicability, especially in resource-constrained environments. Interpretability of Multi-Scale Results: The interpretability of results obtained from a multi-scale analysis might be challenging. Understanding and extracting meaningful insights from variations at different anatomical scales could be complex and may require advanced expertise.

								Need for High-Resolution Imaging: The proposed method may necessitate high-resolution imaging for accurate segmentation. In scenarios where high-resolution data is not available, the effectiveness of the method could be compromised.
54	Episodic Memory in Amnestic Mild Cognitive Impairment (aMCI) and Alzheimer's Disease Dementia (ADD): Using the "Doors and People" Tool to Differentiate between Early aMCI—Late aMCI—Mild ADD Diagnostic Groups	Diagnostics	2022	Doors and People	90 patients from Greek Association of Alzheimer's Disease and Related Disorders	-	Multifaceted Evaluation: The Doors and People tool used in the study evaluates both visual and verbal aspects of episodic memory, providing a comprehensive assessment and enhancing ecological validity. Confirmation of Hypotheses: The study confirms three hypotheses related to the discriminative power of episodic memory, particularly in distinguishing between early and late aMCI stages, and between early aMCI and mild Alzheimer's Disease (ADD) patients. Support from Neuroimaging Studies: Findings align with neuroimaging studies, such as MRI and Voxel-based morphometry, supporting the significance of episodic memory in predicting progression from MCI to ADD. Longitudinal Predictive Validity: The study contributes to the existing	Uncertainty in Individual Prognostication: While episodic memory measures can provide individual prognostic information, the study acknowledges the considerable scatter in regression plots, indicating uncertainty in individual predictions. Age Discrepancy: Participants in each group have significantly different ages, which may introduce a confounding factor. A longitudinal study with reevaluation after several years is suggested to address this limitation. Limited Use of Biomarkers: The study does not incorporate neuroimaging methods, cortical thickness measurements, Voxel-based morphometry, or CSF biomarkers. Integrating these measures could enhance the robustness of the findings.

							literature by demonstrating the longitudinal predictive validity of episodic memory measures, supporting their role in identifying individuals at risk of developing dementia up to 10 years prior to diagnosis.	Task-Specific Findings: The discriminant potential varies among different tasks within the Doors and People tool. While certain tasks show excellent discriminant potential, others exhibit only fair or poor potential, limiting the tool's uniform efficacy across all its components. Generalizability: The sample selection based on participants willing to undergo specific assessments may limit the generalizability of the findings to a broader population.
55	A Computational Monte Carlo Simulation Strategy to Determine the Temporal Ordering of Abnormal Age Onset Among Biomarkers of Alzheimer's Disease	IEEE/ACM Trans Comput Biol Bioinform	2022	Computational Monte Carlo Simulation Strategy	Not Specified	MCI: 382	Quantitative Temporal Ordering: The computational Monte-Carlo simulation (CMCS) provides a quantitative approach to determine the temporal ordering of abnormal age onsets (AAO) for various Alzheimer's disease biomarkers. Statistical Examination: CMCS employs statistical simulations to assess the ordering of AAO pairs and overall AAO, contributing to a robust understanding of Alzheimer's disease progression. Multimodal Biomarker Assessment: The study incorporates diverse biomarkers, including hippocampus volume, glucose hypometabolic	Data Specificity: The findings are based on data from 382 mild cognitive impairment converters and non-converters, potentially limiting generalizability to broader populations or diverse cohorts. Sensitivity to Biomarker Selection: The observed type-I error differences are specific to the selected biomarkers (V_HC, AVLT_STM, AVLT_LTM, HCI, MMSE, CDR-SOB, NfL), and the results may vary with alternative biomarker combinations. Simulation Assumptions: The accuracy of CMCS depends on the assumptions made during the

							convergence index, plasma neurofilament light, and cognitive assessments, offering a comprehensive view of disease onset. Identification of Significant Differences: The CMCS identifies significant differences in the AAO of biomarkers, revealing insights into the sequence of abnormalities in Alzheimer's disease progression.	simulation of longitudinal data, introducing inherent uncertainties. Clinical Application: While the CMCS provides statistical inferences, translating these findings into direct clinical applications may require further validation and integration into diagnostic frameworks.
56	Consistent connectome landscape mining for cross-site brain disease identification using functional MRI	Medical Image Analysis	2022	Connectome Landscape Modeling Method, Alternating Direction Method of Multipliers	-	-	Cross-Site Consistency: Addresses the challenge of inconsistent findings in brain disorder studies across different sites by mining cross-site consistent connectome landscapes. Data-Driven Representation: Utilizes data-driven representation of functional connectivity networks, avoiding reliance on human-engineered features and potentially improving discriminative power. Connectome Landscape Learning: Learns a weight matrix for joint cross-site connectome landscape learning, network feature extraction, and disease identification, providing a comprehensive approach. Norm Penalties: Incorporates row–column overlap norm penalty for capturing consistent connectome landscapes across multiple sites,	Dependence on Functional Connectivity Networks: The effectiveness of CLM is contingent on the accuracy and relevance of the functional connectivity networks used as input. Inaccuracies or biases in these networks may impact the reliability of the results. Sensitivity to Parameter Choices: The performance of CLM may be sensitive to the choice of parameters, such as the regularization parameters for norm penalties. Suboptimal parameter selection could affect the method's robustness and generalizability. Complexity and Interpretability: The complexity introduced by the joint learning of connectome landscapes and disease identification may make the model challenging to interpret. Understanding the

							enhancing reliability. Also, introduces an -norm penalty for capturing site-specific patterns. Efficient Algorithm: Employs the Alternating Direction Method of Multipliers (ADMM) for efficient and effective solution to the proposed objective function.	specific contributions of different features or patterns may be non-trivial. Applicability to Specific Disorders: The method's suitability for different brain disorders may vary. Its generalizability across diverse neurological conditions and its ability to capture nuanced differences between disorders need careful consideration. Computational Intensity: The efficient algorithm used (ADMM) may still be computationally intensive, especially when dealing with large-scale datasets or high-dimensional connectome representations, limiting its scalability. Validation and Generalization: While demonstrating potential in real-world datasets, the generalizability and validation of CLM across a broader range of datasets and disorders need further exploration to establish its robustness.
57	Integrating Different Data Modalities for the Classification of Alzheimer's Disease Stages	SN Computer Science	2023	ML-based omics imaging approach	ANMerge	AD: 42 MCI: 428 NC: 250	Improved Performance through Integration: The method demonstrates enhanced performance by integrating omics and imaging features, surpassing the individual contributions of these features when considered separately.	Potential Bias from Clinical Features: The utilization of clinical features, particularly cognitive test scores, is identified as a potential limitation. These features, being part of the clinician diagnosis process, may introduce a positive bias

							Consistent Performance Across AD Classification Problems: The observed improvement holds true for various binary AD classification problems, indicating the robustness and generalizability of the approach. Superior Performance in Challenging MCI vs. CN Patient Distinction: The method excels in the challenging task of distinguishing Mild Cognitive Impairment (MCI) vs. Cognitively Normal (CN) patients, showcasing its efficacy even in complex diagnostic scenarios.	in the results, hindering a fair evaluation of the method. Dependency on Specific Dataset (ANMerge): The method's applicability and performance are contingent on the characteristics of the ANMerge dataset. The generalizability to other datasets or real-world scenarios may need further validation. Limited Comparison with Existing Methods: As the ANMerge dataset is relatively new, there is a limitation in the ability to compare the method with existing approaches, potentially limiting the contextual understanding of its performance.
58	Classification and prediction of cognitive performance differences in older age based on brain network patterns using a machine learning approach	Netw Neurosci	2023	Support vector machine (SVM), K-nearest while (KNN), decision tree (DT), naïve Bayes (NB) and linear discriminant analysis (LDA)	1000BRAINS project			
59	Baseline structural MRI and plasma biomarkers predict longitudinal structural atrophy and cognitive decline in early Alzheimer's disease	Alz Res Therapy	2023	Linear Mixed Effect Modeling	Alzheimer's Disease Neuroimaging Initiative (ADNI)	MCI: 439 CN: 286	Comprehensive Analysis: The study includes a thorough analysis involving both structural MRI and plasma measurements, providing a comprehensive understanding of potential biomarkers for disease progression.	Limited Plasma Measurements: The study only incorporated two plasma measurements (plasma p-tau181 and NfL), potentially limiting the comprehensiveness of the identified biomarkers. The exclusion of other promising plasma measures, such as plasma p-tau217

		Diagnostic Specificity: By conducting analyses within specific diagnostic groups (CN and MCI) and Aβ+/Aβ− subgroups, the study recognizes the heterogeneity within the cohort and tailors the analysis to each subgroup's characteristics. Stepwise Modeling: The stepwise linear mixed effect modeling approach allows the systematic identification of the subset of baseline measurements that optimally predict longitudinal changes. The use of Akaike Information Criterion (AIC) for model selection adds rigor by favoring models that balance goodness of fit with simplicity. Iterative Feature Selection: The iterative process of adding baseline measurements to the model based on AIC improvement enhances the precision of the final models, ensuring that only the most informative variables are included. Covariate Consideration: The inclusion of relevant covariates such as age, sex, education, APOE ε4 status, and intracranial volume (ICV) in the initial model controls for potential confounding effects and strengthens the validity of the	and glial fibrillary acidic protein, might impact the overall predictive power of combined biomarkers. Potential Plasma Biomarker Omissions: The omission of potentially more sensitive plasma biomarkers, like plasma p-tau217, which is suggested to be more sensitive to early Alzheimer's disease pathology than p-tau181, could impact the study's ability to capture early disease stages. Variability in Predictive MRI Measures: The structural MRI models demonstrated variability in the most predictive measures across different iterations (e.g., BA35 thickness, posterior hippocampal volume, anterior hippocampal volume, ERC thickness, BA36 thickness). This variability challenges the specificity of the identified effects in the medial temporal lobe (MTL), making it challenging to interpret the consistent impact of MTL subregions on disease progression. Need for MTL Summary Measurement: While structural MRI measurements were consistently included, the study suggests that

		results. Longitudinal Analysis: The use of longitudinal measurements allows the examination of changes over time, providing insights into disease progression dynamics. Subgroup Analysis for Progression Prediction: The logistic regression analyses for discriminating fast and slow progressors add a predictive element to the study, enhancing its relevance for identifying markers associated with different rates of disease progression. Model Comparisons: Comparing different models, including base models with only covariates and models with selected baseline plasma or MRI measures, enables a nuanced understanding of the contribution of each type of measurement to the predictive accuracy of the models. Supplementary Analysis: The inclusion of supplementary analyses, such as univariate analysis, provides additional insights into the predictive value of individual baseline measurements, contributing to a more comprehensive interpretation of the results.	using a summary value derived from all MTL subregional measurements through event-based modeling may offer a more consistent and sensitive measurement. This approach is proposed for future investigation to enhance the specificity and reliability of MTL effects. Clinical Trial Stratification Focus: The study highlights the potential utility of plasma and structural MRI biomarkers for clinical trial stratification and prognosis. However, the focus on these specific applications might limit the broader implications of the findings for other clinical contexts or research objectives.

60	An Optimized Deep Learning Model for Predicting Mild Cognitive Impairment Using Structural MRI	Sensors	2023	VGG16, InceptionV3, and ResNet50	Alzheimer's Disease Neuroimaging Initiative (ADNI)	MCI: 337 NC: 442	Early Diagnosis Potential: Utilizing the entorhinal cortex (EC) as a biomarker enables early detection of Mild Cognitive Impairment (MCI) since changes in this area precede those in the hippocampus. Innovative Approach: The study pioneers the use of EC, an often overlooked biomarker due to its size, in predicting MCI, providing a unique perspective. Efficient Classification: Through experiments on brain slices, feature extraction, and classifier optimization, the study achieves an efficient classification system for distinguishing between MCI and normal cognition (NC) samples.	Data Scope Restriction: The model relies solely on MRI data, neglecting other potentially valuable data types such as clinical, genetic, and genomics, limiting the comprehensiveness of the predictions. Small EC Size Challenge: Detecting changes in the EC, being smaller compared to the hippocampus, poses a challenge, potentially limiting the precision of the predictions. Parameter Dependency: The performance improvement of the convolutional neural network (CNN) classifier is contingent on tuning parameters for specific pre-trained models, potentially limiting generalizability.

3. Data Types and Sources

The intricacies surrounding data play a pivotal role in the realm of deep learning, where the nature and abundance of data wield a direct influence on model performance and its capacity for generalization. The reviewed studies have extensively drawn from a diverse array of data, originating from multiple sources. In the subsequent section, we encapsulate the primary types of data at play and shed light on the diverse origins of these crucial datasets.

3.1 Data Types

The available data encompasses two primary categories: longitudinal data and cross-sectional data. Longitudinal data captures the progression of a subject's disease over time, while cross-sectional data represents single-instance, time-independent information. Longitudinal data can take on various forms, serving as independent data, time-series data, or comparative data.

Demographics, often referred to as meta-data, accompany other examinations, offering essential insights into age, gender, and education. Neuroimaging data, a cornerstone, is collected in various modalities, including PET, MRI, and CT for diagnostic purposes, and 3D-MRI, MRI, and SPECT for research endeavors. Cognitive assessments (CA) contribute significantly, featuring MMSE, CDR, ADAS-Cog, logical memory tests, and postural kinematic analysis. CSF, blood plasma biomarkers, and genetic data are sourced from diverse outlets.

Less common yet impactful data types encompass electroencephalography (EEG) for monitoring brain activity, mass spectra data derived from surface-enhanced laser desorption and ionization assays of saliva, and retinal imaging for abnormalities. Electronic health records are explored for dementia and AD screening, while alternative data types like speech, activity

pattern monitoring, and eye-tracking undergo investigation with a deep learning approach. Limited comparative studies within the realm of deep learning exist, particularly between different data modalities and less common data types.

3.2 Data Sources

Over the past two decades, the establishment of several open libraries has significantly facilitated researchers' access to data related to Alzheimer's Disease (AD) and its counterparts. A prominent example is the Alzheimer's Disease Neuroimaging Initiative (ADNI), a comprehensive longitudinal study focused on developing innovative biomarkers for AD detection and progression monitoring. The original ADNI cohort, gathered from 2004 to 2010, comprises T1-weighted MRI, FDG-PET, blood, and CSF biomarkers from 800 subjects. Subsequent cohorts, ADNI-Go and ADNI-2, extended the study while incorporating a broader range of AD stages and adding 200 new subjects with early Mild Cognitive Impairment (MCI). A fourth cohort, ADNI3, initiated in 2016, delves into additional modalities targeting tau protein tangles and is set to conclude in 2022.

ADNI stands out as a widely used open library for neuroimaging data. Another prevalent resource is the Open Access Series of Imaging Studies (OASIS), featuring cross-sectional (OASIS-1) and longitudinal (OASIS-2) cohorts, providing MRI data for demented or non-demented subjects. A supplementary longitudinal cohort (OASIS-3) furnishes MRI and PET data for subjects with normal cognition or AD. While ADNI includes genomic data, OASIS focuses solely on neuroimaging and neuropsychology data.

Other valuable open libraries include the Harvard Aging Brain Study (HABS) and Minimal Interval Resonance Imaging in Alzheimer's Disease (MIRIAD). These repositories play a pivotal role in advancing machine learning and deep learning studies in AD research. Additionally, various local studies, such as Japan ADNI (J-ADNI), the Hong Kong Alzheimer's Disease Study, and the Australian Imaging Biomarkers and Lifestyle Study of Ageing (AIBL), model their approaches akin to ADNI for data compatibility.

Numerous institutes have established platforms, including NeuGRID and the Global Alzheimer's Association Interactive Network, to provide information and efficient access to available databases and libraries. A shortlist of selected data sources is presented in Table 2.

Table 3. Datasets Available for AD Prediction

Library	Number of Subjects	Modalities	Link
LiADNI	2750	MRI, PET, CSF, Genetic	http://adni.loni.usc.edu/
OASIS	1300+	MRI, PET	https://oasis-brains.org/
AIBL	1100+	MRI, PET, CSF, Genetic	https://aibl.csiro.au/
NACC	47,000+	Neuropathology, Genetic	https://www.alz.washington.edu/
EDSD	471	MRI, DTI, Genetic	https://www.neugrid2.eu/
ARWIBO	2700+	MRI, PET, Genetic	http://www.arwibo.it/
HABS	290	MRI, PET, Genetic	https://habs.mgh.harvard.edu/
KLOSCAD	6818	MRI, QOL, Behavioral	http://kloscad.com/
VITA	606	MRI, Genetic	https://www.neugrid2.eu/

For machine learning practitioners and researchers, alternative data sources include challenges hosted by ADNI and other institutions, such as CADDementia, TADPOLE, DREAM, and the Kaggle international challenge for automated prediction of MCI from MRI data. These challenges often offer pre-selected or preprocessed data, alleviating the need for extensive expert knowledge. Some studies propose utilizing brain age as a surrogate measure of cognitive decline, drawing on databases of cognitively normal individuals, including UKBioBank,

NKI, IXI, LifespanCN, and the Cambridge dataset. Other potential data sources encompass the International Genomics of Alzheimer's Project (IGAP), the Korean Longitudinal Study on Cognitive Aging and Dementia (KLOSCAD), the INSIGHT-preAD study, the Imaging Dementia—Evidence for Amyloid Scanning (IDEA) study, and the European version of ADNI—AddNeuroMed. Institutes holding private data collections include the National Alzheimer's Coordination Center, the Biobank of Beaumont Reference Laboratory, and IRCCS. Given the magnitude of the health crisis posed by AD, numerous studies have amassed data, with the aforementioned sources representing the most commonly utilized in reviewed literature, along with examples of alternative sources.

3.3 Image Modalities

In the pursuit of advancing early detection and predictive modelling for Alzheimer's Disease (AD), the integration of diverse medical image modalities has become paramount. This sub section explores the intricate landscape of imaging technologies, delving into their collective potential to unveil subtle neurological changes associated with AD. From magnetic resonance imaging (MRI) and positron emission tomography (PET) to emerging techniques, each modality contributes unique insights, fostering a comprehensive understanding of the disease's progression. This paper navigates the spectrum of medical imaging modalities, unraveling their significance in refining AD prediction models and paving the way for more effective diagnostic strategies.

3.3.1 Structural Magnetic Resonance Imaging

MRI stands as a secure and non-invasive medical imaging technique, leveraging a potent combination of magnetic fields, radio waves, and computational power to produce high-quality images with exceptional spatial resolution. Within the realm of brain exploration, two distinct MRI techniques—structural MRI (sMRI) and functional MRI (fMRI)—unfold, each offering unique perspectives. sMRI, a non-invasive brain imaging method, probes alterations in brain structure, including atrophy, tissue loss, and morphological changes, particularly crucial in understanding cognitive decline.

Navigating the complexities of MRI machines, the inherent challenges such as B1-field inhomogeneity, bias field artifacts, and gradient non-linearity necessitate meticulous preprocessing steps. Bias field correction, often employing B1-scans, addresses non-uniformity, while gradient non-linearity is rectified using methods like Gradwarp. Intensity normalization, a pivotal step, mitigates discrepancies across various MRI machines, crucial for large-scale multi-center studies. Motion correction techniques, addressing subject motion artifacts during scans, further enhance data quality.

Brain extraction emerges as a pivotal preprocessing component, involving the removal of non-brain elements. Techniques like skull-stripping, cerebellum removal, and neck removal contribute to isolating crucial brain components. Subsequent spatial normalization, achieved through registration to anatomical templates like MNI-152, aids in aligning images and enabling comparative analyses. An alternative challenge, age misalignment between control and AD subjects, prompts the creation of study-specific template spaces based on training data.

Segmentation, the division of brain MRI into anatomical regions, proves essential for isolating structures pertinent to AD. Manual segmentation, while accurate, is time-intensive; hence, automated algorithms like FSL FIRST and FreeSurfer gain prominence for large datasets. Neural networks focused on segmentation, particularly of the hippocampus, have become a recurring theme in AD-related studies, showcasing the application of deep learning in this domain.

Post preprocessing, downsampling is often employed to reduce dimensionality, ensuring uniform input dimensions for neural networks. Smoothing techniques, enhancing signal-to-noise ratios, are commonplace but entail trade-offs in amplitude and peak bandwidth. Age correction strategies, acknowledging the impact of normal brain atrophy with age, employ voxel-wise linear regression models post-registration, contributing to overall model performance. In the intricate journey through MRI preprocessing, each step plays a pivotal role in refining data quality and paving the way for robust neurodegenerative studies.

3.3.2 Positron Emission Tomography

PET, employing a radioactive tracer, offers insights into cellular and tissue activity within the body. In the context of neurological disorders, this tracer selectively binds to disease-associated proteins like amyloid beta, a hallmark of AD, and tau in AD cases. PET also aids in detecting alterations in glucose metabolism, a key aspect of the brains of Alzheimer's patients. The preprocessing of PET images mirrors that of structural MRI.

In AD-related investigations, PET data frequently accompany MRI data, often collected jointly in prominent studies such as ADNI. The initial preprocessing steps, including image registration and segmentation, are executed on the MRI data. Subsequently, PET images align rigidly with their corresponding MRI counterparts. Post-segmentation, downsampling, and smoothing procedures are akin to those applied to MRI images. In instances where MRI data are absent, independent studies either adopt simplified preprocessing akin to MRI methods or opt for minimal preprocessing. This strategic integration of PET and MRI data fortifies the robustness of analyses in unraveling the complexities of Alzheimer's disease.

3.3.3 Functional Magnetic Resonance Imaging

Functional MRI (fMRI) is a dynamic magnetic resonance imaging technique designed to gauge brain activity by monitoring cerebral blood flow. Unlike static structural MRI, fMRI captures a temporal series of images, enabling the study of disease-related changes in brain function. These alterations may involve shifts in connectivity between distinct brain regions and variations in how the brain responds to stimuli. Notably, fMRI is instrumental in investigating changes in memory and attention linked to cognitive impairment in conditions like MCI and AD. Both structural MRI (sMRI) and fMRI serve as valuable tools for tracking disease progression by detecting changes in specific brain regions over time.

In addition to the preprocessing procedures outlined for structural MRI in sub-section 3.3.1, specific steps are essential for fMRI data. Slice time correction ensures accurate timing of the time-series data, correcting for temporal offsets between scan instances. Prolonged fMRI scanning and the collection of multiple images in a single session elevate the risk of head motion artifacts, necessitating additional filtering or motion correction. This correction is typically achieved through spatial alignment to the first or chosen scan before spatial normalization. Introduction of high-pass and low-pass filters to the temporal domain controls the frequency and period of fMRI data, enhancing data quality.

Various preprocessing tools, such as the SPM REST Toolkit, DPABI, or FreeSurfer, automate fMRI data preprocessing. Methods for reducing data redundancy in fMRI are often categorized as common spatial pattern (CSP) or brain functional network (BFN) based. CSP methods produce spatial filters optimizing one group's variance while minimizing another. BFN-based methods leverage ROI segmentation to construct a brain network, with ROI features as vertices and functional connections as edges. Deep learning approaches have recently been applied to create weighted correlation kernels within neural network architectures, extracting dynamic functional connectivity networks. These advancements underscore the evolving landscape of fMRI preprocessing techniques, enhancing the utility of fMRI in unraveling intricate aspects of brain function associated with neurodegenerative diseases.

4. Data Processing

Data processing is crucial in deep learning, influencing how models are built and perform. Unlike traditional machine learning that focuses on extracting features, deep learning prepares input data for neural networks, avoiding the need for predefined representations. The goal is to highlight important information in the data and make it consistent for the model across different samples and types of data. This processing for AD prediction is grouped into different types of inputs, as shown in Figure 2.

4.1 Feature Based

Feature-based approaches focus on individual features in the provided data, particularly in neuroimaging, where it's known as the voxel-based approach. This method is applied to individual image voxels after spatial normalization, ensuring alignment across images. Grey matter probability maps are often segmented to reduce input information, and machine learning can extract texture, shape, or other features, forming a hybrid ML-DL approach. While voxel-based methods retain global 2D or 3D information, they may overlook local details. For data types like cognitive assessments, CSF, serum, and genetic biomarkers, feature-based approaches are common. Handling longitudinal data, such as EEG or speech, involves additional processing for completeness, like imputing missing data and aligning timestamps.

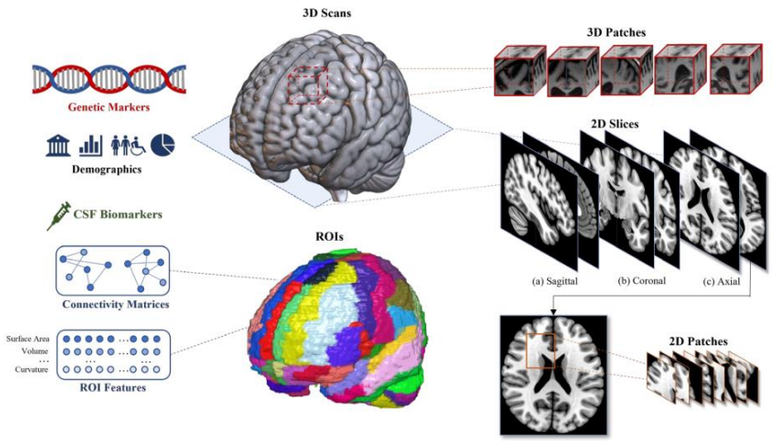

Figure 2. Different Data Processing Methods Illustrated over SMRI

4.2 Patch Based

Instead of utilizing all features or 2D slices, the patch-based approach involves using predefined regions of specific sizes as input for the model, which can be either 2D or 3D based on model requirements. Some studies, like that of Suk HI et al. [18], position patch-based extraction as an intermediary level between a voxel-based approach and an ROI-based approach. This strategy proves efficient in addressing concerns related to high feature dimensions and sensitivity to subtle changes. Also patch-based approaches adeptly manage region-wide pathologies, extending beyond specific ROIs. This aligns with the perspective of neurologists or radiologists, who analyze images by investigating local patterns and subsequently integrating distributed local information across the entire brain to formulate clinical decisions. Individual patches have a smaller memory footprint and lower input dimensions, reducing computational resources needed for training. However, reconstructing sample-level results during testing and application requires additional resources, affecting efficiency.

The challenge with patch-based approaches lies in selecting the most informative regions, considering factors like patch size, overlap between patches, and the choice of essential patches. Studies use methods like statistical significance of voxels or landmark-based approaches around anatomically significant points to determine patching regions.

4.3 Slice Based

Slice-based approaches rely on 2D images or data, assuming that 2D information is sufficient for representing the required information when dealing with 3D data. In practical clinical diagnosis, a limited number of 2D slices are often used instead of a complete 3D image. Some studies extract single or multiple 2D slices from the sagittal, axial, or coronal planes of the 3D scan. While axial plane slices are commonly extracted, the coronal view may contain crucial AD-related regions. The selection of slices from 3D scans often focuses on specific brain dissections and their anatomical components, such as the sagittal slices of the hippocampus, a known region of interest. Some studies use sorting procedures to identify the most valuable slices, like entropy sorting with greyscale histograms. Slice-based approaches are generally less computationally expensive than feature-based approaches, but they come with the drawback of losing global and 3D geometric structures. This limitation can be addressed by using multiple slices from various views, for example, slices from three projections showcasing the hippocampal region, and combining slices from different modalities like MRI and PET.

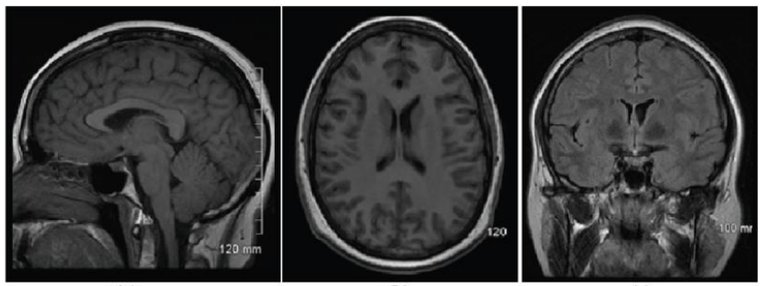

Figure 3. Brain MRI in (a) Sagittal Plane (b) Axial Plane and (c) Coronal Plane

4.4 Region of Interest Based

Patch-based methods involve predefined regions with rigid extraction sizes, while ROI-based methods concentrate on anatomical regions of interest (ROIs) within the brain. These functionally relevant ROIs are carefully chosen during the preprocessing stage of atlas-based brain registration. The Automated Anatomical Labeling (AAL) atlas, with 93 ROIs, is a common choice among the reviewed studies, along with other atlases like the Kabani reference work and the Harvard-Oxford cortical and subcortical structural atlases. Elastic registration techniques, such as HAMMER, exhibit higher registration performance.

Following ROI extraction, reviewed studies commonly use gray matter (GM) tissue mean intensities or volumes from brain ROIs as features, derived from various modalities such as PET, MRI, and fMRI. Additional measures include subcortical volumes, gray matter densities, cortical thickness, brain glucose metabolism, cerebral amyloid-β accumulation, and average regional CMRGlc for PET. The hippocampus, a key focus in the papers, is studied using 3D data and morphological measurements, including cortical thickness, curvature, surface area, and volume. Some studies explore relationships between ROIs, using correlation to generate connectivity matrices that are often categorized into cortical and subcortical regions.

ROI-based methods closely align with anatomical regions, offering high interpretability and clinical implementability. However, their strong connection to prior knowledge limits their potential in explorative studies. While the computational cost falls between voxel and slice-based approaches, ROI-based methods can preserve local 3D geometric information. Proposed hierarchical neural network frameworks feature sub-networks at each representation level, incorporating effective network pruning to retain complete information.

4.5 Voxel Based

Voxel-based approaches are feature-based methods that analyze individual voxels, the three-dimensional pixels composing a medical image. These voxels, representing discrete brain locations, can be adjusted in size and number to balance computational efficiency and spatial resolution. Unlike slice-based approaches, voxel-based methods capture the three-dimensional structure of the brain, revealing changes not evident in two-dimensional slices. Due to the brain's complexity and inter-subject differences, spatial co-alignment (registration) is crucial.

Registration aligns image scans to an anatomical reference space, involving the alignment of MRI images to a standardized template representing a common anatomical space. Many studies segment aligned images into different tissue types, like grey matter, white matter, and cerebrospinal fluid, before applying the model. Comparing gray and white matter across groups or time points can sensitively detect subtle changes in brain structure. However, voxel-based approaches have limitations, notably the need for high spatial resolution. One paper utilizes functional network topologies to depict neurodegeneration in a low-dimensional form, expressing these topologies in a low-dimensional manifold, and representing brain state configurations in a relatively compact space.

5. Unsupervised Learning Techniques Employed in AD Prediction

Unsupervised learning extracts insights without the predefined categorization of data samples or labels, whereas supervised approaches rely on labelled data samples. In deep learning, architectures aren't strictly categorized as supervised or unsupervised when broken down into their core components, such as feature extraction and classification in convolutional neural networks. In this survey, the categorization is based on the relationship between the optimization target of the main neural network or framework and ground truth labels. This section focuses on summarizing unsupervised learning methods, while Section 6 provides a summary of supervised learning methods.

5.1 Autoencoders (AE)ni

Autoencoders stand as a pivotal paradigm in artificial neural networks, engineered to glean efficient data representations. This unsupervised learning method comprises two integral components: the encoder, denoted as f_e , and the decoder, represented by f_d . In this classical setup, the encoder serves as a neural network meticulously crafted to navigate from the input domain to a latent feature representation. On the flip side, the decoder mirrors the encoder, diligently working to reconstruct the original input from the compressed representation.

The crux of autoencoder functionality lies in its ability to learn a compressed, information-rich representation of input data, encapsulating essential features. This compressed representation, often referred to as the latent space, captures intricate patterns within the data. The encoder-decoder architecture ensures that the reconstructed output closely approximates the initial input, reflecting the model's capacity to distil critical information while maintaining fidelity.

The application of autoencoders extends across various domains, including medical imaging and neuroimaging, where preserving relevant features and reducing dimensionality are paramount. Studies such as Liu et al. [22] have successfully employed autoencoders in unsupervised scenarios for effective feature learning. Hosseini-Asl et al. [27] delved into the nuanced architectural choices influencing autoencoder performance, shedding light on optimal configurations.

In essence, autoencoders emerge as versatile tools in unsupervised learning, offering a pathway to unravel efficient data representations. Their application holds promise across diverse fields, with implications for improved feature extraction, dimensionality reduction, and nuanced data understanding.

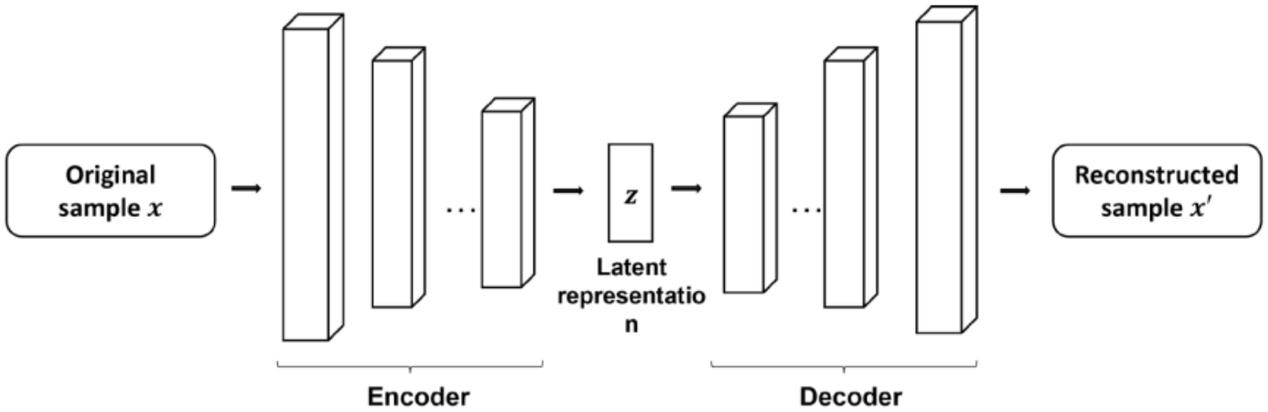

Figure 4. 2D Stacked Autoencoder

5.2 Generative Adversarial Networks (GANs)

Generative methods, a subset of unsupervised learning, embark on the compelling quest of supplementing existing data distributions by creating novel data instances. Within this realm, both Variational Autoencoders (VAEs) and Restricted Boltzmann Machines (RBMs), as previously discussed, function as generative models. Another formidable contender in this landscape is the Generative Adversarial Network (GAN), an innovative approach where neural networks engage in a zero-sum game, fostering an intriguing competition.

In the classical GAN architecture, two pivotal neural networks come into play: the generative network (G) and the discriminator network (D). The generative network is tasked with crafting synthetic data, generating dummy images (x') from noise (e), belonging to the generated data distribution $x' \in p_g$. On the other front, the discriminator network strives to distinguish between the synthetic images (x') and real images $x' \in p_r(x)$. The interplay between these networks is encapsulated in a loss function, manifesting a dynamic zero-sum game. The objective is to simultaneously minimize the generative network (G) and maximize the discriminator network (D).

The applications of GANs in the medical imaging domain are vast, spanning synthesis, reconstruction, segmentation, and classification. A notable instance by Islam and Zhang leveraged a convolutional GAN to synthesize PET images for Alzheimer's Disease (AD), Normal Control (NC), and Mild Cognitive Impairment (MCI). Achieving a mean Peak Signal-to-Noise Ratio (PSNR) of 32.83 and a mean Structural Similarity Index (SSIM) of 77.48, the generated images were subsequently classified using a 2D Convolutional Neural Network (CNN) with an accuracy of 71.45%. However, this performance variation underscores the intricate challenges in generating high-quality synthetic images for training purposes.

Building upon the GAN framework, transfer learning played a pivotal role in achieving promising results. In a similar vein, Roychowdhury and Roychowdhury introduced a conditional GAN, where both the discriminator and generator were conditioned by labels (y), exemplifying the multifaceted potential of GANs in fostering creativity within the unsupervised learning paradigm.

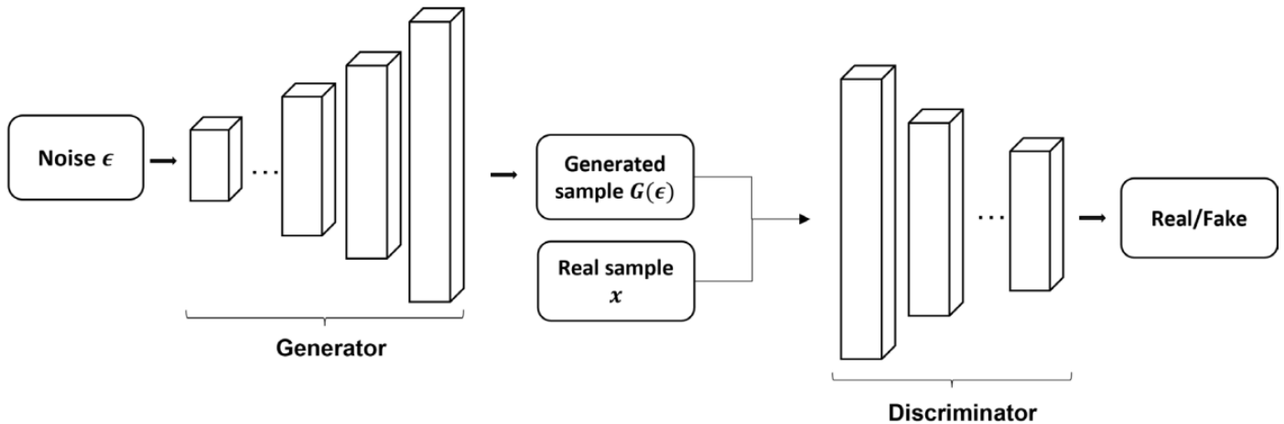

Figure 5. Structure of GAN

5.3 Restricted Boltzmann Machine and Other Unsupervised Models

Various unsupervised methods beyond GANs and Autoencoders have found application in Alzheimer's Disease (AD) and related research. Among these, the restricted Boltzmann machine (RBM) stands out. Operating as a generative network, RBM utilizes a bipartite graph to extract probability distributions of input data. With symmetrically-linked layers housing visible and hidden units, RBM encodes input data during the forward pass and reconstructs it during the backward pass, aided by two sets of biases. Like autoencoders, RBMs also serve as feature extraction tools. Li et al. [79] employed multiple RBMs to initialize hidden layers sequentially, while Suk et al. [80] integrated RBM with the autoencoder learning module, employing layer-wise learning with greedy optimization. Conditional RBM has proven effective in unsupervised progression forecasting of Mild Cognitive Impairment (MCI), showcasing predictive performance for ADAS-Cog 13 compared to supervised methods [81].

Deep Belief Networks (DBNs), composed of stacked RBMs, offer a robust architecture with a backward pass of generative weights, enhancing noise resilience. Despite the computational expense of layer-by-layer learning, Suk, Lee, Shen, and Initiative [82] demonstrated the effectiveness of combining Multi-Layer Perceptrons (MLP) and DBM for feature extraction across multiple modalities.

Recent advances in unsupervised learning exhibit diversity in approaches. Razavi et al. [83] utilized sparse filtering as a pre-training strategy for a 2D CNN, effectively minimizing feature sparsity. Bi et al. [84] innovatively merged a CNN with Principal Component Analysis (PCA)-generated filters and k-means clustering for a fully unsupervised framework in clustering MRI data of AD, MCI, and Normal Controls (NC). Wang, Xin, Wang, Gu, Zhao, and Qian [85] introduced hierarchical extreme learning machines for unsupervised feature representation extraction, a variant of feedforward neural networks using the Moore-Penrose generalized inverse. Majumdar and Singhal [86] explored deep dictionary input, employing denoising autoencoders for categorical classification with noisy inputs. Cheng et al. [87] employed a U-net-based CNN with rigid alignment for cortical surface registration in MRI images.

6. Supervised and Semi-Supervised Learning

6.1 Convolutional Neural Networks

Convolutional Neural Networks (CNNs) have revolutionized neural networks, particularly with the introduction of AlexNet by Krizhevsky, showcasing the practicality of neural networks as universal approximators. CNNs excel in hierarchical feature extraction, offering parameter efficiency and translational invariance. The architecture's ability to retain spatial information makes it particularly well-suited for handling neuroimaging data. This effectiveness is apparent in diverse applications, whether as standalone models or integrated components within larger networks.

Typically, a CNN consists of multiple convolutional layers succeeded by non-linear activation functions. Commonly used activation functions include Rectified Linear Unit (ReLU), hyperbolic tangent, and sigmoid functions, with newer alternatives like leaky-ReLU and parametric-ReLU gaining traction in recent literature [87,88]. Pooling, achieved through average or maximum filter downsampling, is a common practice in CNNs. Batch normalization, a process of standardizing each mini-batch of data, is often applied after convolution. These individual procedures form a convolution block, and a standard CNN comprises multiple such blocks. Following these blocks are usually a few fully connected layers and a Softmax activation for classification or a linear activation for regression. While the theoretical foundations of CNN involve the decomposition of tensors [89], this paper focuses on the practical applications of CNNs in AD-related tasks.

6.1.1 2D CNN

The original Convolutional Neural Network (CNN) was initially designed for pattern recognition in 2D computer vision applications, making it well-suited for processing 2D neuroimaging data. A fundamental 2D-CNN structure is depicted in Figure 6. Aderghal, Benois-Pineau, and Afdel [90] employed a two-layer CNN featuring Rectified Linear Unit (ReLU) activation and max-pooling of 2D+ images. These images were generated by projecting slices from sagittal, coronal, and axial planes into a three-channel 2D image. Alternatively, when multiple 2D slices are available from various planes of a 3D image, individual 2D-CNNs can be applied to each image, and their results can be combined (ensembling). Notably, an increase in neural network depth is often associated with improved performance. For instance, Wang, Phillips, Sui, Liu, Yang, and Cheng [91] proposed a deeper eight-layer CNN with leaky rectified linear units to classify single-slice MRI images. Similar architectures were employed for Florbetaben-18 PET image classification [87].

Tang et al. [92] demonstrated the application of a CNN model to identify amyloid plaques in Alzheimer's disease histology slides. This CNN, akin to the 2D CNNs mentioned earlier, comprises alternating layers of 2D convolution and max-pooling, followed by fully connected layers featuring ReLU activation. The Softmax activation is then applied to produce classification outputs. This CNN model exhibited outstanding performance in amyloid plaque classification, achieving an impressive Area Under the Curve (AUC) of 0.993. While state-of-the-art 2D CNN models have primarily been developed for natural image classification, their adaptability makes them easily applicable to 2D AD-related data.

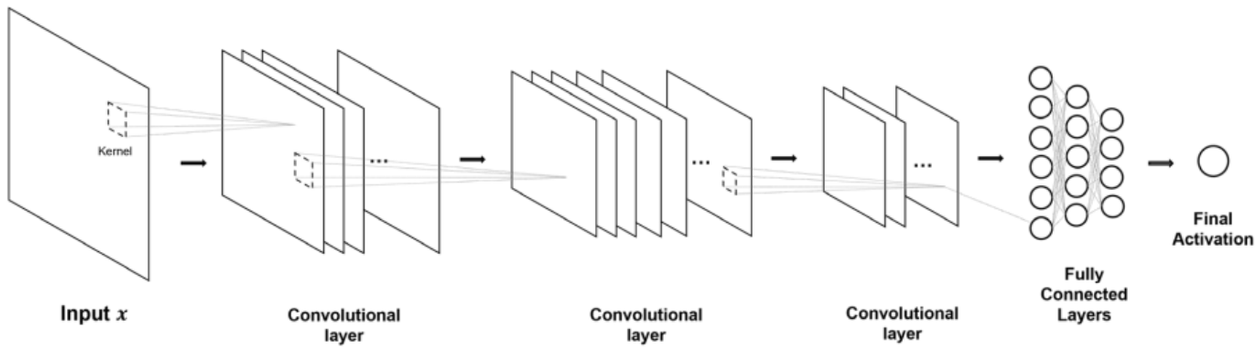

Figure 6. 2D CNN Architecture

Due to the inherent limitation of two dimensions, data with multiple slices are typically treated as independent or similar in 2D-CNN applications. Interestingly, 2D-CNNs can be effectively employed with 1D data, such as transforming cognitive assessment data into 2D using the Hilbert space-filling curve [93] or representing the time-series data of multi-channel EEG as a 2D matrix [94]. Despite the constraint of dimensionality, 2D-CNNs offer practicality in real-world applications and deployments. This is particularly relevant in clinical practice, where the data used is often 2D or lacks sufficient slices to construct high-dimensional 3D T1-weighted MRI, which is predominantly used in medical research.

For instance, 3D neuroimaging data available in open libraries like ADNI and OASIS are often processed to obtain 2D slices or patches, as discussed in Section 4. To preserve 3D spatial information, 2D slices or patches are commonly extracted from sagittal, coronal, and axial views for multi-view networks [95]. The lower dimensionality of 2D-CNNs also makes them suitable for adaptation to 1D data. Alavi et al. [96], for example, utilized the triplet architecture of face recognition and the Siamese one-shot learning model for automated live comparative analysis of RNA-seq data from GEC.

6.1.2 3D CNN

A 3D-CNN shares fundamental similarities with a 2D-CNN but includes an additional dimension in all components, most notably in the convolutional kernel. This additional dimension enhances the spatial information captured by 3D-CNN compared to its 2D counterpart. Unlike 2D-CNN, which is constrained by kernel dimensionality, 3D-CNN efficiently captures spatial information across slices. A basic representation of a 3D-CNN is illustrated in Figure 7. In line with foundational 2D-CNN models, Islam and Zhang [97] implemented a 3D-CNN comprising four 3D convolutional layers with PCL and Softmax using T1-weighted MRI. Similarly, Duc et al. [98] utilized a comparable CNN architecture with rs-fMRI functional networks. Basaia et al. [99] demonstrated the effectiveness of a simple two-block 3D-CNN, showcasing either comparable or superior performance to 2D-CNN in binary classification tasks involving AD, NC, and various MCI subtypes.

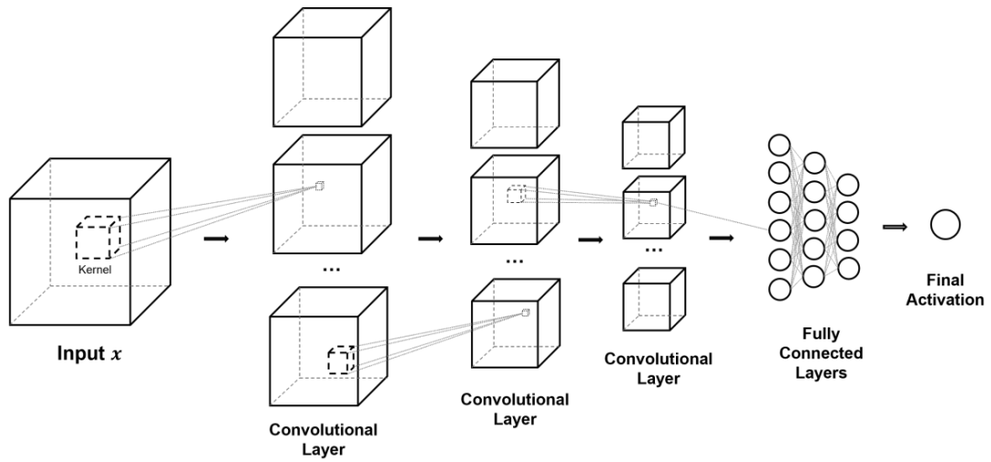

Figure 7. 3D CNN Architecture

The versatility of 3D-CNNs lies in their adaptability from high-performing 2D architectures. For instance, Basaia et al. [99] and Qiu et al. [100] transitioned from successful 2D versions to 3D adaptations. Choi et al. [101] leveraged an all convolutional CNN for MCI conversion prediction. Unsupervised pre-training with 3D convolutional autoencoders was explored by both Hosseini-Asl, Keynton, and El-Baz [238] and Martinez-Murcia et al. [102]. Additionally, 3D-CNN features have been used as inputs for sparse autoencoders [103], and Ge et al. [104] combined a U-Net-structured 3D-CNN with XG-Boost feature selection. State-of-the-art 2D-CNN architectures, such as Inception-v4 [105], have successfully transitioned to 3D, exemplified by Liu et al.'s [106] use of 3D-AlexNet and 3D-ResNet. Wang et al. [108] introduced a probability-based ensemble of densely connected neural networks with 3D kernels, showcasing the potential of ensemble learning for enhanced performance. 3D-CNN is not limited to spatial domain input; it can handle dynamic functional connectivity networks (FCNs) represented in 2D with an additional temporal dimension. This allows the network to integrate temporal and spatial connectivity, enhancing its ability to characterize time-dependent interactions [112]. Despite its advantages, the additional dimension in 3D-CNN results in more parameters and increased computational cost. Strategies like parameter-efficient 3D separable convolution [109] and the consideration of normalization techniques [107] have been proposed to address these challenges. Liu, Cheng, Wang, Wang, and Initiative [113] proposed a combined use of an ensemble 3D-CNN and 2D-CNN sequentially. The 3D-CNN captures spatial correlations in the 3D input, and an ensemble of cascading 3D-CNN-generated feature maps serves as input for 2D-CNNs. While many studies focus on classification and disease progression, there are applications in segmentation and image processing. Yang et al. [110] introduced a 3D-CNN with residual learning for efficient hippocampal segmentation. Pang et al. [111] combined a semi-supervised autoencoder with local linear mapping. The ongoing development of more powerful hardware has contributed to the growing popularity of 3D-CNN applications in Alzheimer's and related diseases within the reviewed literature.

6.2 Recurrent Neural Networks

Longitudinal data in Alzheimer's Disease (AD) research offer a unique perspective by providing multiple instances of data for each subject over time. While traditional Deep Neural Networks (DNN) and Convolutional Neural Networks (CNN) might not fully exploit the temporal nature of such data, Recurrent Neural Networks (RNN)

emerge as a valuable solution. RNNs, designed to handle sequential data, introduce a temporal dimension through time-varying activation and a synapse-like sequential structure.

Long Short-Term Memory (LSTM), a type of RNN, has been successfully applied to brain network graph matrices derived from functional Magnetic Resonance Imaging (fMRI) data. The combination of LSTM and Extreme Learning Machine (ELM) exhibited a slight improvement over a CNN-ELM model in classification tasks [185]. Another variant, the Gated Recurrent Unit (GRU), possesses a structure similar to LSTM but lacks a forget gate, resulting in fewer parameters. GRU has been employed for classifying temporal clustering patterns in actigraphy time-series monitoring, showcasing its efficacy in capturing long-term temporal patterns [114].

Bi-directional GRU (BGRU), capable of processing input both forwards and backward, has been used in various studies as a classification component, replacing traditional Multi-Layer Perceptron (MLP) or machine learning classifiers [115,116]. RNNs can also be applied to structural data, as demonstrated by a study combining CNN and RNN. In this approach, CNN captures features within single slices, while BGRU structures process a time-series of CNN-extracted features to extract inter-slice features, eventually used as input for an MLP classifier component [117]. Modifications of LSTM to accommodate 3D structural data, such as the 3D Convolutional LSTM, have also been explored [118].

Addressing challenges related to time-series or sequential data, where datasets often have missing or delayed collection time points, is crucial. Nguyen et al. [119] proposed a minimal RNN model for imputing missing data, achieving notable success in the TADPOLE longitudinal challenge for predicting ADAS-Cog13 and ventricular volume over a 6-year period. These studies collectively demonstrate the effectiveness of RNNs in temporal modeling for AD and related diseases. As longitudinal data collection continues to grow across various projects, the impact of RNNs on the trajectory of deep learning approaches is expected to be significant.

6.3 Other Methods

Reinforcement learning, a distinct branch of artificial intelligence research, diverges from supervised and unsupervised learning by focusing on agents' actions within an environment. Tang, Uchendu, Wang, Dodge, and Zhou [120] applied reinforcement learning in conjunction with natural language processing techniques to develop an MCI screening dialogue agent. In this setup, the reinforcement learning environment employed the Actor-Critic method, where a user simulator neural network generated new dialogue data. This approach shares similarities with Generative Adversarial Networks (GAN), but in GAN, the actor cannot influence the reward of the critic function [121].

While perceptron units in neural networks aim to simulate the fundamental function of human brain neurons, this representation is considered oversimplified. Recent research in deep learning has ventured into creating neural networks based on more representative biological neurons, giving rise to spiking neural networks (SNN). Unlike the sequential nature of Recurrent Neural Networks (RNN), SNNs are inherently temporal by design. Capecci et al. [122] provided a proof-of-concept using an SNN architecture with EEG data for predicting MCI conversion.

Following is the summary of the literature mentioned in Section 4-6

Table 4. Summary of Literature Discussed

Study Name	Data Modalities	Number of Subjects			Classification Accuracy (%)	
		AD	NC	MCI	AD vs NC	MCI vs NC
Deep learning-based feature representation for AD/MCI classification	MRI, PET	51	52	99	95.9	85
Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis	MRI, PET	93	101	204	95.35	85.67
Multimodal neuroimaging feature learning for multiclass diagnosis of Alzheimer's disease	MRI, PET	85	109	77	82.59	82.10
A Robust Deep Model for Improved Classification of AD/MCI Patients	MRI, PET, CSF	51	99	52	91.4	77.4
Classification of sMRI for Alzheimer's disease Diagnosis with CNN: Single Siamese Networks with 2D Approach and Fusion on ADNI	MRI	188	228	399	69.53	91.41
Deep ensemble learning of sparse regression models for brain disease diagnosis	MRI	186	393	226	91.02	-
Noisy deep dictionary learning: Application to Alzheimer's Disease classification	MRI, PET, CSF	51	99	52	95.4	85.7
Longitudinal analysis for Alzheimer's disease diagnosis using RNN	MRI	188	228	399	69.53	91.41
Multimodal neuroimaging feature learning with multimodal stacked deep polynomial networks for diagnosis of Alzheimer's disease	MRI, PET, CSF	51	99	52	95.4	85.7
Improving Alzheimer's disease classification by combining multiple measures	MRI	198	229	-	89.69	-
Multiscale deep neural network based analysis of FDG-PET images for the early diagnosis of Alzheimer's disease	PET	226	304	521	93.58	-
Classifying Alzheimer's disease with brain imaging and genetic data using a neural network framework	MRI, Genetic	138	225	358	-	-

Multi-Modality Cascaded Convolutional Neural Networks for Alzheimer's Disease Diagnosis	MRI, PET	93	100	204	93.26	74.34
Multiscale Deep Convolutional Networks for Characterization and Detection of Alzheimer's Disease Using MR images	MRI	193	139	-	93.53	-
Early diagnosis of Alzheimer's disease based on resting-state brain networks and deep learning	fMRI	-	79	91	-	86.47
Joint classification and regression via deep multi-task multi-channel learning for Alzheimer's disease diagnosis	MRI	227	249	390	93.7	-
Understanding 3D CNN Behavior for Alzheimer's Disease Diagnosis from Brain PET Scan	PET	169	400	661	88.76	-
Convolutional Neural Networks for Classification of Alzheimer's Disease: Overview and Reproducible Evaluation	MRI	336	330	787	87	-
A multi-model deep convolutional neural network for automatic hippocampus segmentation and classification in Alzheimer's disease	MRI	97	119	233	88.9	76.2
Toward an interpretable Alzheimer's disease diagnostic model with regional abnormality representation via deep learning	MRI	198	229	374	92.75	89.22
Hierarchical fully convolutional network for joint atrophy localization and Alzheimer's Disease diagnosis using structural MRI	MRI	358	205	2964	89.5	-
Hippocampus Analysis by Combination of 3-D DenseNet and Shapes for Alzheimer's Disease Diagnosis	MRI	192	223	396	92.29	74.64
Studying the manifold structure of Alzheimer's Disease: A deep learning approach using convolutional autoencoders	MRI	99	168	212	84.9	-
3D-Deep Learning Based Automatic Diagnosis of Alzheimer's Disease with Joint MMSE Prediction Using Resting-State fMRI	fMRI	133	198	-	85.3	-
Slice-selective learning for Alzheimer's disease classification using a generative adversarial network: A feasibility study of external validation.	PET	212	415	-	94.82	-
Cognitive signature of brain FDG PET based on deep learning: Domain transfer from Alzheimer's disease to Parkinson's disease.	PET	243	393	666	-	-

A Novel End-to-End Hybrid Network for Alzheimer's Disease Detection Using 3D CNN and 3D CLSTM	MRI	198	299	408	94.19	79.01
A Convolutional Neural Network based self-learning approach for classifying neurodegenerative states from EEG signals in dementia	EEG	63	63	63	85.78	85.34
GAN-based synthetic brain PET image generation	PET	98	105	208	71.45	-
Development and validation of an interpretable deep learning framework for Alzheimer's disease classification	MRI, Demo, CA	488	978	-	96.8	-
MRI signatures of brain age and disease over the lifespan based on a deep brain network and 14,468 individuals worldwide	MRI	353	833	513	86	70.2
Multi-View Separable Pyramid Network for AD Prediction at MCI Stage by ^{18}F-FDG Brain PET Imaging	PET	237	242	526	93.13	-

7. Deep Learning Techniques

7.1 Transfer Learning

The widespread adoption of deep neural networks in medical diagnostic systems faces common challenges in practical applications [123]. These challenges stem from issues such as the availability of medical data and relevant labels. Unlike computer vision systems, which benefit from large annotated databases like ImageNet [124], medical images are scarce and often require expert knowledge for labeling [125–128]. One potential solution to this challenge is transfer learning, a technique that involves transferring knowledge across domains [129]. In the context of image classification, transfer learning typically involves transferring model structure, weights, or parameters for classification in different feature spaces and distributions. Neural networks with transferred parameters have shown improved convergence and reduced requirements for hyperparameter searches compared to networks with randomized parameters [130].

There are three types of transfer learning:

1. Transfer from ImageNet pre-trained models: Examples include studies that utilized pre-trained models like Inception-V3, Inception-V4, AlexNet, ResNet-152, VGG-16, and DenseNet-121 for various classification and regression tasks related to Alzheimer's disease.
2. Transfer from pre-trained networks for similar tasks: This involves using a pre-trained network from one dataset for another dataset.
3. Transfer from pre-trained networks for different tasks: This involves using a pre-trained model for a specific classification or prediction task and applying it to a different task, such as using an AD vs. NC pre-trained model for classifying different stages of mild cognitive impairment (MCI) or other classifications.

Transfer learning has been applied across diverse datasets and tasks, including brain age prediction, brain-related diseases (e.g., Alzheimer's and Parkinson's), and even non-medical domains like eye-tracking. Notable examples include transferring knowledge between different datasets for brain age prediction and transferring models used for brain age prediction to tasks involving Alzheimer's, mild cognitive impairment (MCI), and normal control (NC) classifications. These applications showcase the versatility and effectiveness of transfer learning in addressing challenges related to limited medical data.

7.2 Ensemble Learning

Ensemble learning in the context of deep learning involves combining multiple representations to achieve enhanced overall performance, mitigating errors within individual neural networks. The application of ensemble learning has proven effective in various medical image classification tasks, including those related to Alzheimer's disease (AD) and associated conditions.

Ensemble learning can be implemented at three different levels: input, feature, and output.

1. Input-level ensemble: Combining data prior to input into the neural network. For example, adjacent slices of hippocampal data can be combined to construct mimic RGB channels. Another approach involves using zero-masking for the fusion of concatenated magnetic resonance imaging (MRI) and positron emission tomography (PET) inputs.
2. Feature-level ensemble: Combining features from patch-level, region-level, and subject-level sub-networks as input features for a classification module. An example is the hierarchical sub-networks at each level, where the outputs of each level are concatenated and used as input for the next level. Feature-level ensembles have also been applied at individual feature levels, incorporating an ensemble of multi-scale patch-level sub-networks.
3. Output-level ensemble: Combining the predictions of component neural networks. This can involve majority voting of prediction results, combining outputs of multiple sparse regression models with varying regularization parameters, or utilizing a probability-based fusion of softmax outputs from an ensemble of 3D-DenseNets. Output-level ensembles can also allow for the combination of neural networks with traditional machine learning classifiers.

A sub-category of ensemble learning is multi-view learning, particularly relevant in the context of the 3D nature of neuroimaging data. In this approach, multiple views or representations are considered in the learning process. For instance, a pyramid network of multiple convolutional neural network (CNN) subnetworks with separable convolutions for each of the three views was created, and the features were added and concatenated for classification. While ensembling at all three levels is common in the reviewed literature, there are relatively fewer applications of the boosting method, which involves training individual components sequentially in an adaptive manner. Boosting is a standard technique in machine learning applications. The ensemble with multiple modalities, often referred to as multi-modal fusion, is introduced in the subsequent section.

7.3 Multi Modal Fusion

In the landscape of Alzheimer's disease (AD) research, the integration of multi-modal data has emerged as a pivotal strategy to overcome inherent limitations within individual data sources. Notably, genetic data may lack critical neuroimaging texture information, and the high soft-tissue resolution of MRI might not inherently correlate with essential factors such as Amyloid-β protein depositions. To address these constraints and foster a more holistic understanding of the disease, researchers frequently employ various multi-modal fusion techniques. These methodologies encompass diverse approaches, including feature-level ensembles where modality-dependent components are fused through concatenation or merging. Zero-masking for stacked autoencoders is utilized, reconstructing one modality with the information from another. Demographics and genetic biomarkers find integration through concatenation with neuroimaging data, emphasizing the importance of a unified framework. Direct fusion methods are applied for 1D data or engineered features, providing a straightforward combination. Multi-scale or multi-view learning techniques involve training individual neural networks for patches of varying sizes, with their outputs concatenated for classification. Additionally, intricately designed connections between 1D and 3D network structures facilitate the fusion of MRI and neuropsychological data. Extensive architectures further fuse demographic and genetic markers with inputs, such as the Jacobian of structural MRI images. While multi-modal fusion is anticipated to enhance neural network performance, challenges persist, notably regarding data availability, particularly in longitudinal studies. The diagram in Figure 8 provides a schematic representation of these common multi-modal fusion methods within the research context.

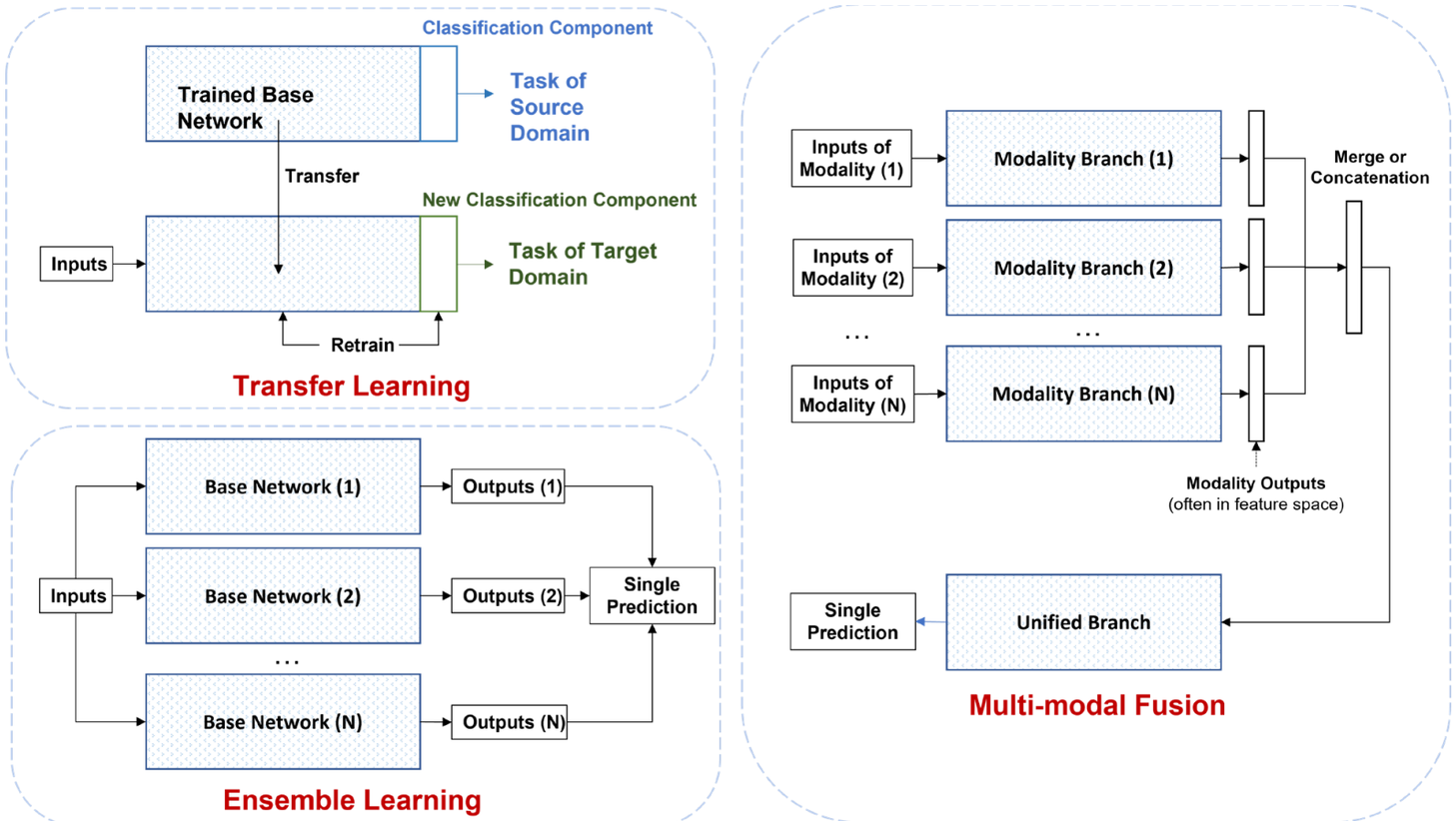

Figure 8. Overview of Different Deep Learning Techniques Employed in AD Prediction

8. Proposed Methodology

In our research, we present a pioneering methodology for Alzheimer's Disease (AD) prediction by integrating Convolutional Neural Networks (CNNs) for spatial feature extraction from neuroimaging data with Quantum Machine Learning (QML) algorithms for classification. The process begins with the preprocessing of neuroimaging datasets, followed by the utilization of a CNN architecture for hierarchical feature extraction, capturing intricate spatial patterns indicative of different cognitive states. The quantum advantage is then harnessed through algorithms such as Quantum Support Vector Machines (QSVM) or Quantum Variational Circuits (QVC) to classify the extracted features into AD, Normal Control (NC), or Mild Cognitive Impairment (MCI) categories. A crucial innovation lies in the quantum encoding of CNN-extracted features into qubits, exploiting quantum superposition to represent complex feature interactions efficiently. The quantum circuit training optimizes classification parameters, leveraging quantum parallelism for accelerated training. Ensemble learning is employed to combine outputs from multiple quantum circuits, ensuring robustness and generalization. This methodology's uniqueness lies in the synergy of classical CNNs and Quantum Machine Learning, offering enhanced feature extraction, quantum processing advantages, and the robustness of ensemble learning. Evaluation metrics such as accuracy, precision, recall, and F1 score, along

with validation on diverse datasets using cross-validation techniques, confirm the effectiveness of this novel approach, marking a significant stride towards improved accuracy and efficiency in early AD prediction models.

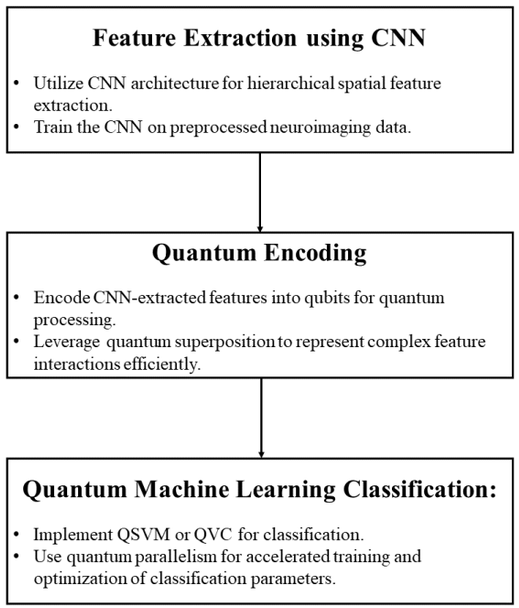

Figure 9. Proposed Methodology

9. Challenges in Generalizability of Deep Learning Models in Real World

A significant obstacle in the realm of deep learning (DL) lies in the effective generalization of models to real-world scenarios. The challenge of generalization is notably influenced by the nature of the data employed for model training. Many studies in the literature have predominantly utilized data collected under stringent acquisition protocols, specifying modalities, types, and hardware, often lacking representation of clinical settings. Although preprocessing is commonly employed to mitigate variabilities, its diverse applications and subjective nature introduce uncertainty, creating a dual-edged impact. Mårtensson et al. [131] conducted a comprehensive assessment of training and testing across different data domains. Their investigation revealed that while the recurrent CNN proposed exhibited consistency across diverse datasets, performance degradation was observed when evaluating data collected under protocols different from those used in training. To address this, incorporating a broader range of protocols during training demonstrated improved generalization performance on unseen data. While this study, centered on a single CNN-based model, does not offer conclusive findings, it provides valuable insights into potential generalization challenges and underscores the significance of data heterogeneity in addressing them. Recognizing that the amount of data utilized in training and evaluation significantly influences generalization, methods to enhance data quantity and heterogeneity include incorporating lower-dimensional data (e.g., using 2D slices of 3D scans), data augmentation, and judicious use of generative models. Alternatively, reducing the model's

training data requirement, often through semi-supervised approaches with a train-test split of 50% or lower, allows for a larger testing set and a more accurate approximation of generalizability.

The theoretical aspect of this challenge involves estimating the 'generalization gap,' defined as the difference between metrics derived from an independent test set and those in real-world scenarios. Approximate generalization bounds derived from the Hoeffding inequality [132], based on certain assumptions, offer core insights into the relationship between the testing set and the approximate generalization gap. These bounds, proportional to the inverted root of the sample size, highlight the importance of a larger sample size for improved generalizability, although practical achievement of the data required by these worst-case bounds may be challenging. Estimation of generalization also considers model complexity, typically measured through the Vapnik–Chervonenkis dimension. Alternative methods for deriving generalization bounds include using the validation set [133], measuring network smoothness [134], and comparing generalization error between deep neural networks and humans [135]. Another approach to estimating generalization addresses label inhomogeneity due to misdiagnosis. Wu et al. [136] proposed employing models for unsure data to account for discordant Mild Cognitive Impairment (MCI) samples with uncertain conversion.

Beyond the technical and theoretical aspects of generalization, the practical application in clinical settings, particularly in mass screening, is a crucial consideration. The increase in false positives generated by deep learning models, as observed in small-scale studies, has been identified to elevate radiologist workload. In large-scale screening, the impact of overdiagnosis resulting from false positives can significantly affect cost and efficiency. Therefore, close monitoring of false-positive rates alongside approximations of the generalization gap emerges as a pivotal aspect of evaluation in these scenarios.

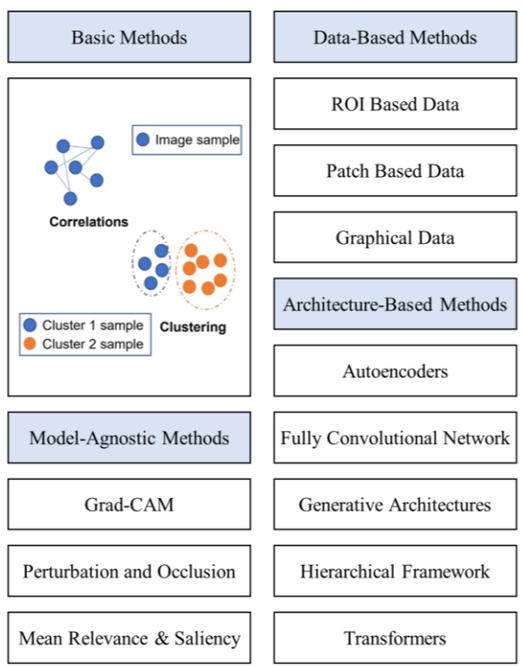

Figure 10. Overview of Common Methods found in Literature

10. Challenges

Over the past 13 years, numerous studies in deep learning have been dedicated to Alzheimer's disease (AD) and related conditions, resulting in diverse techniques, models, and protocols. This comprehensive overview summarizes the major components contributing to deep learning studies and highlights recent advancements, including recurrent neural networks, graph and geometric neural networks, and generative modeling. While these studies showcase promising outcomes across various tasks such as image processing, disease classification, and progression prediction, the lack of consistency in approaches and a scarcity of standardized benchmarks hinder effective comparisons. The majority of these studies are research-focused, with limited evaluations conducted or simulated in clinical settings, contributing to challenges in interpretation and generalization of deep learning findings. To address these issues, this review explores potential solutions for interpretation, such as visualization techniques and inherently interpretable architectures. Additionally, it offers insights into pathways for improving generalization, including considerations of data heterogeneity, data quantity, and generalization gap approximation. Beyond interpretation and generalization, there are other avenues for potential research, such as deep learning for polygenic studies and the application of transformer-based foundational models. As model architectures continue to evolve, these pathways promise to lead towards more robust and clinically applicable deep learning models for AD and related diseases.

References

1. Cortes-Briones, J.A.; Tapia-Rivas, N.I.; D'Souza, D.C.; Estevez, P.A. Going deep into schizophrenia with artificial intelligence. Schizophr. Res. 2022, 245, 122–140
2. Choo, H.; Yoo, S.Y.; Moon, S.; Park, M.; Lee, J.; Sung, K.W.; Cha, W.C.; Shin, S.-Y.; Son, M.H. Deep-learning-based personalized prediction of absolute neutrophil count recovery and comparison with clinicians for validation. J. Biomed. Inform. 2023, 137, 104268
3. Visser PJ, Verhey FR, Hofman PA, Scheltens P, Jolles J. Medial temporal lobe atrophy predicts Alzheimer's disease in patients with minor cognitive impairment. J Neurol Neurosurg Psychiatry. 2002 Apr;72(4):491-7. doi: 10.1136/jnnp.72.4.491. PMID: 11909909; PMCID: PMC1737837.
4. DeCarli C, Frisoni GB, Clark CM, et al. Qualitative Estimates of Medial Temporal Atrophy as a Predictor of Progression From Mild Cognitive Impairment to Dementia. Arch Neurol. 2007;64(1):108–115. doi:10.1001/archneur.64.1.108.
5. K. B. Walhovd et al., "Combining MR Imaging, Positron-Emission Tomography, and CSF Biomarkers in the Diagnosis and Prognosis of Alzheimer Disease," in IEEE Transactions on Medical Imaging, vol. 31, no. 2, pp. 347-354, Feb. 2010, doi: 10.3174/ajnr.A1809.
6. Huang S, Li J, Sun L, Ye J, Fleisher A, Wu T, Chen K, Reiman E; Alzheimer's Disease NeuroImaging Initiative. Learning brain connectivity of Alzheimer's disease by sparse inverse covariance estimation. Neuroimage. 2010 Apr 15;50(3):935-49. doi: 10.1016/j.neuroimage.2009.12.120. Epub 2010 Jan 14. PMID: 20079441; PMCID: PMC3068623.
7. Zhang D, Wang Y, Zhou L, Yuan H, Shen D; Alzheimer's Disease Neuroimaging Initiative. Multimodal classification of Alzheimer's disease and mild cognitive impairment. Neuroimage. 2011 Apr 1;55(3):856-67. doi: 10.1016/j.neuroimage.2011.01.008. Epub 2011 Jan 12. PMID: 21236349; PMCID: PMC3057360.
8. Wee CY, Yap PT, Denny K, Browndyke JN, Potter GG, Welsh-Bohmer KA, Wang L, Shen D. Resting-state multi-spectrum functional connectivity networks for identification of MCI patients. PLoS One. 2012;7(5):e37828. doi: 10.1371/journal.pone.0037828. Epub 2012 May 30. PMID: 22666397; PMCID: PMC3364275.
9. Zhang D, Shen D; Alzheimer's Disease Neuroimaging Initiative. Multi-modal multi-task learning for joint prediction of multiple regression and classification variables in Alzheimer's disease. Neuroimage. 2012 Jan 16;59(2):895-907. doi: 10.1016/j.neuroimage.2011.09.069. Epub 2011 Oct 4. Erratum in: Neuroimage. 2012 Sep;62(3):2179. PMID: 21992749; PMCID: PMC3230721.
10. Ye J, Farnum M, Yang E, Verbeeck R, Lobanov V, Raghavan N, Novak G, DiBernardo A, Narayan VA; Alzheimer's Disease Neuroimaging Initiative. Sparse learning and stability selection for predicting MCI to AD conversion using baseline ADNI data. BMC Neurol. 2012 Jun 25;12:46. doi: 10.1186/1471-2377-12-46. PMID: 22731740; PMCID: PMC3477025.
11. Ye J, Farnum M, Yang E, Verbeeck R, Lobanov V, Raghavan N, Novak G, DiBernardo A, Narayan VA; Alzheimer's Disease Neuroimaging Initiative. Sparse learning and stability selection for predicting MCI to AD conversion using baseline ADNI data. BMC Neurol. 2012 Jun 25;12:46. doi: 10.1186/1471-2377-12-46. PMID: 22731740; PMCID: PMC3477025.
12. Weintraub D, Dietz N, Duda JE, Wolk DA, Doshi J, Xie SX, Davatzikos C, Clark CM, Siderowf A. Alzheimer's disease pattern of brain atrophy predicts cognitive decline in Parkinson's disease. Brain. 2012 Jan;135(Pt 1):170-80. doi: 10.1093/brain/awr277. Epub 2011 Nov 21. PMID: 22108576; PMCID: PMC3316476.
13. Ye J, Farnum M, Yang E, Verbeeck R, Lobanov V, Raghavan N, Novak G, DiBernardo A, Narayan VA; Alzheimer's Disease Neuroimaging Initiative. Sparse learning and stability selection for predicting MCI to AD conversion using baseline ADNI data. BMC Neurol. 2012 Jun 25;12:46. doi: 10.1186/1471-2377-12-46. PMID: 22731740; PMCID: PMC3477025.
14. Gupta Y, Lee KH, Choi KY, Lee JJ, Kim BC, Kwon GR. Alzheimer's Disease Diagnosis Based on Cortical and Subcortical Features. J Healthc Eng. 2019 Mar 3;2019:2492719. doi: 10.1155/2019/2492719. PMID: 30944718; PMCID: PMC6421724.
15. Andersen AH, Rayens WS, Liu Y, Smith CD. Partial least squares for discrimination in fMRI data. Magn Reson Imaging. 2012 Apr;30(3):446-52. doi: 10.1016/j.mri.2011.11.001. Epub 2012 Jan 5. PMID: 22227352; PMCID: PMC3288364.
16. Eskildsen SF, Coupé P, García-Lorenzo D, Fonov V, Pruessner JC, Collins DL; Alzheimer's Disease Neuroimaging Initiative. Prediction of Alzheimer's disease in subjects with mild cognitive impairment from the ADNI cohort using patterns of cortical thinning. Neuroimage. 2013 Jan 15;65:511-21. doi: 10.1016/j.neuroimage.2012.09.058. Epub 2012 Oct 2. PMID: 23036450; PMCID: PMC4237400.

17. Suk HI, Lee SW, Shen D; Alzheimer's Disease Neuroimaging Initiative. Deep sparse multi-task learning for feature selection in Alzheimer's disease diagnosis. Brain Struct Funct. 2016 Jun;221(5):2569-87. doi: 10.1007/s00429-015-1059-y. Epub 2015 May 21. PMID: 25993900; PMCID: PMC4714963.
18. Bernal-Rusiel JL, Greve DN, Reuter M, Fischl B, Sabuncu MR; Alzheimer's Disease Neuroimaging Initiative. Statistical analysis of longitudinal neuroimage data with Linear Mixed Effects models. Neuroimage. 2013 Feb 1;66:249-60. doi: 10.1016/j.neuroimage.2012.10.065. Epub 2012 Oct 30. Erratum in: Neuroimage. 2015 Mar;108:110. PMID: 23123680; PMCID: PMC3586747.
19. Eskildsen SF, Coupé P, García-Lorenzo D, Fonov V, Pruessner JC, Collins DL; Alzheimer's Disease Neuroimaging Initiative. Prediction of Alzheimer's disease in subjects with mild cognitive impairment from the ADNI cohort using patterns of cortical thinning. Neuroimage. 2013 Jan 15;65:511-21. doi: 10.1016/j.neuroimage.2012.09.058. Epub 2012 Oct 2. PMID: 23036450; PMCID: PMC4237400.
20. Suk HI, Lee SW, Shen D; Alzheimer's Disease Neuroimaging Initiative. Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis. Neuroimage. 2014 Nov 1;101:569-82. doi: 10.1016/j.neuroimage.2014.06.077. Epub 2014 Jul 18. PMID: 25042445; PMCID: PMC4165842.
21. Liu F, Wee CY, Chen H, Shen D. Inter-modality relationship constrained multi-modality multi-task feature selection for Alzheimer's Disease and mild cognitive impairment identification. Neuroimage. 2014 Jan 1;84:466-75. doi: 10.1016/j.neuroimage.2013.09.015. Epub 2013 Sep 14. PMID: 24045077; PMCID: PMC3849328.
22. Liu M, Zhang D, Shen D; Alzheimer's Disease Neuroimaging Initiative. View-centralized multi-atlas classification for Alzheimer's disease diagnosis. Hum Brain Mapp. 2015 May;36(5):1847-65. doi: 10.1002/hbm.22741. Epub 2015 Jan 27. PMID: 25624081; PMCID: PMC6869465.
23. Moradi E, Pepe A, Gaser C, Huttunen H, Tohka J; Alzheimer's Disease Neuroimaging Initiative. Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects. Neuroimage. 2015 Jan 1;104:398-412. doi: 10.1016/j.neuroimage.2014.10.002. Epub 2014 Oct 12. PMID: 25312773; PMCID: PMC5957071.
24. H. Ji, Z. Liu, W. Q. Yan, and R. Klette, "Early Diagnosis of Alzheimer's Disease Using Deep Learning," in *Proceedings of the 2nd International Conference on Control and Computer Vision*, 2019.
25. J. Long, E. Shelhamer and T. Darrell, "Fully convolutional networks for semantic segmentation," 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 2015, pp. 3431-3440, doi: 10.1109/CVPR.2015.7298965.
26. Zhang Y, Dong Z, Phillips P, Wang S, Ji G, Yang J, Yuan TF. Detection of subjects and brain regions related to Alzheimer's disease using 3D MRI scans based on eigenbrain and machine learning. Front Comput Neurosci. 2015 Jun 2;9:66. doi: 10.3389/fncom.2015.00066. PMID: 26082713; PMCID: PMC4451357.
27. Challis E, Hurley P, Serra L, Bozzali M, Oliver S, Cercignani M. Gaussian process classification of Alzheimer's disease and mild cognitive impairment from resting-state fMRI. Neuroimage. 2015 May 15;112:232-243. doi: 10.1016/j.neuroimage.2015.02.037. Epub 2015 Feb 28. PMID: 25731993.
28. S. Sarraf, G. Tofighi, "DeepAD: Alzheimer's Disease Classification via Deep Convolutional Neural Networks using MRI and fMRI," bioRxiv, 070441, doi: 10.1101/070441.
29. Hosseini-Asl E, Ghazal M, Mahmoud A, Aslantas A, Shalaby AM, Casanova MF, Barnes GN, Gimel'farb G, Keynton R, El-Baz A. Alzheimer's disease diagnostics by a 3D deeply supervised adaptable convolutional network. Front Biosci (Landmark Ed). 2018 Jan 1;23(3):584-596. doi: 10.2741/4606. PMID: 28930562.
30. Zhang H, Chen X, Shi F, Li G, Kim M, Giannakopoulos P, Haller S, Shen D. Topographical Information-Based High-Order Functional Connectivity and Its Application in Abnormality Detection for Mild Cognitive Impairment. J Alzheimers Dis. 2016 Oct 4;54(3):1095-1112. doi: 10.3233/JAD-160092. PMID: 27567817; PMCID: PMC5437847.
31. Tong Tong, Katherine Gray, Qinquan Gao, Liang Chen, and Daniel Rueckert. 2017. Multi-modal classification of Alzheimer's disease using nonlinear graph fusion. Pattern Recogn. 63, C (Mar 2017), 171–181. https://doi.org/10.1016/j.patcog.2016.10.009
32. Petersen RC, Wiste HJ, Weigand SD, Rocca WA, Roberts RO, Mielke MM, Lowe VJ, Knopman DS, Pankratz VS, Machulda MM, Geda YE, Jack CR Jr. Association of Elevated Amyloid Levels With Cognition and Biomarkers in Cognitively Normal People From the Community. JAMA Neurol. 2016 Jan;73(1):85-92. doi: 10.1001/jamaneurol.2015.3098. Erratum in: JAMA Neurol. 2016 Apr;73(4):481. PMID: 26595683; PMCID: PMC4710552.
33. Cui Y, Wen W, Lipnicki DM, Beg MF, Jin JS, Luo S, Zhu W, Kochan NA, Reppermund S, Zhuang L, Raamana PR, Liu T, Trollor JN, Wang L, Brodaty H, Sachdev PS. Automated detection of amnestic mild cognitive impairment in community-dwelling elderly adults: a combined spatial atrophy and white matter alteration approach. Neuroimage. 2012 Jan 16;59(2):1209-17. doi: 10.1016/j.neuroimage.2011.08.013. Epub 2011 Aug 16. PMID: 21864688.
34. Ye J, Farnum M, Yang E, Verbeeck R, Lobanov V, Raghavan N, Novak G, DiBernardo A, Narayan VA; Alzheimer's Disease Neuroimaging Initiative. Sparse learning and stability selection for predicting MCI to AD conversion using baseline ADNI data. BMC Neurol. 2012 Jun 25;12:46. doi: 10.1186/1471-2377-12-46. PMID: 22731740; PMCID: PMC3477025.
35. DeCarli C, Frisoni GB, Clark CM, Harvey D, Grundman M, Petersen RC, Thal LJ, Jin S, Jack CR Jr, Scheltens P; Alzheimer's Disease Cooperative Study Group. Qualitative estimates of medial temporal atrophy as a predictor of progression from mild cognitive impairment to dementia. Arch Neurol. 2007 Jan;64(1):108-15. doi: 10.1001/archneur.64.1.108. Erratum in: Arch Neurol. 2007 Mar;64(3):459. PMID: 17210817.
36. Visser PJ, Verhey FR, Hofman PA, Scheltens P, Jolles J. Medial temporal lobe atrophy predicts Alzheimer's disease in patients with minor cognitive impairment. J Neurol Neurosurg Psychiatry. 2002 Apr;72(4):491-7. doi: 10.1136/jnnp.72.4.491. PMID: 11909909; PMCID: PMC1737837
37. Krämer C, Stumme J, da Costa Campos L, Rubbert C, Caspers J, Caspers S, Jockwitz C. Classification and prediction of cognitive performance differences in older age based on brain network patterns using a machine learning approach. Netw Neurosci. 2023 Jan 1;7(1):122-147. doi: 10.1162/netn_a_00275. PMID: 37339286; PMCID: PMC10270720.
38. Xie, L., Das, S.R., Wisse, L.E.M. et al. Baseline structural MRI and plasma biomarkers predict longitudinal structural atrophy and cognitive decline in early Alzheimer's disease. Alz Res Therapy 15, 79 (2023). https://doi.org/10.1186/s13195-023-01210-z
39. Walhovd KB, Fjell AM, Brewer J, McEvoy LK, Fennema-Notestine C, Hagler DJ Jr, Jennings RG, Karow D, Dale AM; Alzheimer's Disease Neuroimaging Initiative. Combining MR imaging, positron-emission tomography, and CSF biomarkers in the diagnosis and prognosis of Alzheimer disease. AJNR Am J Neuroradiol. 2010 Feb;31(2):347-54. doi: 10.3174/ajnr.A1809. Epub 2010 Jan 14. PMID: 20075088; PMCID: PMC2821467.
40. Mendonça AR, Loureiro LM, Nórte CE, Landeira-Fernandez J. Episodic memory training in elderly: A systematic review. Front Psychol. 2022 Jul 28;13:947519. doi: 10.3389/fpsyg.2022.947519. PMID: 35967680; PMCID: PMC9366047.
41. Zhou H, Jiang J, Lu J, Wang M, Zhang H, Zuo C; Alzheimer's Disease Neuroimaging Initiative. Dual-Model Radiomic Biomarkers Predict Development of Mild Cognitive Impairment Progression to Alzheimer's Disease. Front Neurosci. 2019 Jan 11;12:1045. doi: 10.3389/fnins.2018.01045. PMID: 30686995; PMCID: PMC6338093.
42. Jönemo J, Akbar MU, Kämpe R, Hamilton JP, Eklund A. Efficient Brain Age Prediction from 3D MRI Volumes Using 2D Projections. Brain Sci. 2023 Sep 15;13(9):1329. doi: 10.3390/brainsci13091329. PMID: 37759930; PMCID: PMC10526282.

43. Mosconi L, Berti V, Glodzik L, Pupi A, De Santi S, de Leon MJ. Pre-clinical detection of Alzheimer's disease using FDG-PET, with or without amyloid imaging. J Alzheimers Dis. 2010;20(3):843-54. doi: 10.3233/JAD-2010-091504. PMID: 20182025; PMCID: PMC3038340.
44. Alyoubi EH, Moria KM, Alghamdi JS, Tayeb HO. An Optimized Deep Learning Model for Predicting Mild Cognitive Impairment Using Structural MRI. Sensors (Basel). 2023 Jun 16;23(12):5648. doi: 10.3390/s23125648. PMID: 37420812; PMCID: PMC10302234.
45. Saleem TJ, Zahra SR, Wu F, Alwakeel A, Alwakeel M, Jeribi F, Hijji M. Deep Learning-Based Diagnosis of Alzheimer's Disease. J Pers Med. 2022 May 18;12(5):815. doi: 10.3390/jpm12050815. PMID: 35629237; PMCID: PMC9143671.
46. Guo X, Chen K, Chen Y, Xiong C, Su Y, Yao L, Reiman EM. A Computational Monte Carlo Simulation Strategy to Determine the Temporal Ordering of Abnormal Age Onset Among Biomarkers of Alzheimer's Disease. IEEE/ACM Trans Comput Biol Bioinform. 2022 Sep-Oct;19(5):2613-2622. doi: 10.1109/TCBB.2021.3106939. Epub 2022 Oct 10. PMID: 34428151; PMCID: PMC9588284.
47. Gupta Y, Lee KH, Choi KY, Lee JJ, Kim BC, Kwon GR; National Research Center for Dementia; Alzheimer's Disease Neuroimaging Initiative. Early diagnosis of Alzheimer's disease using combined features from voxel-based morphometry and cortical, subcortical, and hippocampus regions of MRI T1 brain images. PLoS One. 2019 Oct 4;14(10):e0222446. doi: 10.1371/journal.pone.0222446. PMID: 31584953; PMCID: PMC6777799.
48. Huang S, Li J, Sun L, Ye J, Fleisher A, Wu T, Chen K, Reiman E; Alzheimer's Disease NeuroImaging Initiative. Learning brain connectivity of Alzheimer's disease by sparse inverse covariance estimation. Neuroimage. 2010 Apr 15;50(3):935-49. doi: 10.1016/j.neuroimage.2009.12.120. Epub 2010 Jan 14. PMID: 20079441; PMCID: PMC3068623.
49. Chen YJ, Deutsch G, Satya R, Liu HG, Mountz JM. A semi-quantitative method for correlating brain disease groups with normal controls using SPECT: Alzheimer's disease versus vascular dementia. Comput Med Imaging Graph. 2013 Jan;37(1):40-7. doi: 10.1016/j.compmedimag.2012.11.001. Epub 2012 Dec 27. PMID: 23273615.
50. Hanseeuw BJ, Betensky RA, Jacobs HIL, Schultz AP, Sepulcre J, Becker JA, Cosio DMO, Farrell M, Quiroz YT, Mormino EC, Buckley RF, Papp KV, Amariglio RA, Dewachter I, Ivanoiu A, Huijbers W, Hedden T, Marshall GA, Chhatwal JP, Rentz DM, Sperling RA, Johnson K. Association of Amyloid and Tau With Cognition in Preclinical Alzheimer Disease: A Longitudinal Study. JAMA Neurol. 2019 Aug 1;76(8):915-924. doi: 10.1001/jamaneurol.2019.1424. Erratum in: JAMA Neurol. 2019 Aug 1;76(8):986. PMID: 31157827; PMCID: PMC6547132.
51. Beheshti I, Demirel H, Matsuda H; Alzheimer's Disease Neuroimaging Initiative. Classification of Alzheimer's disease and prediction of mild cognitive impairment-to-Alzheimer's conversion from structural magnetic resource imaging using feature ranking and a genetic algorithm. Comput Biol Med. 2017 Apr 1;83:109-119. doi: 10.1016/j.compbiomed.2017.02.011. Epub 2017 Feb 27. PMID: 28260614.
52. Gupta Y, Lee KH, Choi KY, Lee JJ, Kim BC, Kwon GR. Alzheimer's Disease Diagnosis Based on Cortical and Subcortical Features. J Healthc Eng. 2019 Mar 3;2019:2492719. doi: 10.1155/2019/2492719. PMID: 30944718; PMCID: PMC6421724.
53. Andersen AH, Rayens WS, Liu Y, Smith CD. Partial least squares for discrimination in fMRI data. Magn Reson Imaging. 2012 Apr;30(3):446-52. doi: 10.1016/j.mri.2011.11.001. Epub 2012 Jan 5. PMID: 22227352; PMCID: PMC3288364.
54. Hett K, Ta VT, Oguz I, Manjón JV, Coupé P; Alzheimer's Disease Neuroimaging Initiative. Multi-scale graph-based grading for Alzheimer's disease prediction. Med Image Anal. 2021 Jan;67:101850. doi: 10.1016/j.media.2020.101850. Epub 2020 Oct 6. PMID: 33075641; PMCID: PMC7725970.
55. Wang M, Zhang D, Huang J, Liu M, Liu Q. Consistent connectome landscape mining for cross-site brain disease identification using functional MRI. Med Image Anal. 2022 Nov;82:102591. doi: 10.1016/j.media.2022.102591. Epub 2022 Aug 29. PMID: 36070656.
56. Gupta S, Rajapakse JC, Welsch RE; Alzheimer's Disease Neuroimaging Initiative. Ambivert degree identifies crucial brain functional hubs and improves detection of Alzheimer's Disease and Autism Spectrum Disorder. Neuroimage Clin. 2020;25:102186. doi: 10.1016/j.nicl.2020.102186. Epub 2020 Jan 17. PMID: 32000101; PMCID: PMC7042673.
57. Khatri U, Kwon GR. Explainable Vision Transformer with Self-Supervised Learning to Predict Alzheimer's Disease Progression Using 18F-FDG PET. Bioengineering (Basel). 2023 Oct 20;10(10):1225. doi: 10.3390/bioengineering10101225. PMID: 37892955; PMCID: PMC10603890.
58. Huang S, Li J, Sun L, Ye J, Fleisher A, Wu T, Chen K, Reiman E; Alzheimer's Disease NeuroImaging Initiative. Learning brain connectivity of Alzheimer's disease by sparse inverse covariance estimation. Neuroimage. 2010 Apr 15;50(3):935-49. doi: 10.1016/j.neuroimage.2009.12.120. Epub 2010 Jan 14. PMID: 20079441; PMCID: PMC3068623.
59. Zhang H, Chen X, Shi F, Li G, Kim M, Giannakopoulos P, Haller S, Shen D. Topographical Information-Based High-Order Functional Connectivity and Its Application in Abnormality Detection for Mild Cognitive Impairment. J Alzheimers Dis. 2016 Oct 4;54(3):1095-1112. doi: 10.3233/JAD-160092. PMID: 27567817; PMCID: PMC5437847.
60. Li Q, Wu X, Xie F, Chen K, Yao L, Zhang J, Guo X, Li R; Alzheimer's Disease Neuroimaging Initiative. Aberrant Connectivity in Mild Cognitive Impairment and Alzheimer Disease Revealed by Multimodal Neuroimaging Data. Neurodegener Dis. 2018;18(1):5-18. doi: 10.1159/000484248. Epub 2018 Jan 13. PMID: 29334684.
61. Li Y, Liu J, Huang J, Li Z, Liang P. Learning Brain Connectivity Sub-networks by Group- Constrained Sparse Inverse Covariance Estimation for Alzheimer's Disease Classification. Front Neuroinform. 2018 Sep 7;12:58. doi: 10.3389/fninf.2018.00058. PMID: 30258358; PMCID: PMC6143825.
62. Challis E, Hurley P, Serra L, Bozzali M, Oliver S, Cercignani M. Gaussian process classification of Alzheimer's disease and mild cognitive impairment from resting-state fMRI. Neuroimage. 2015 May 15;112:232-243. doi: 10.1016/j.neuroimage.2015.02.037. Epub 2015 Feb 28. PMID: 25731993.
63. Klöppel, S.; Stonnington, C.M.; Chu, C.; Draganski, B.; Scahill, R.I.; Rohrer, J.D.; Fox, N.C.; Jack, C.R., Jr.; Ashburner, J.; Frackowiak, R.S. Automatic classification of MR scans in Alzheimer's disease. Brain 2008, 131, 681–689.
64. Janousova, E.; Vounou, M.; Wolz, R.; Gray, K.R.; Rueckert, D.; Montana, G. Biomarker discovery for sparse classification of brain images in Alzheimer's disease. Ann. BMVA 2012, 2, 1–11
65. Zhang, D.; Wang, Y.; Zhou, L.; Yuan, H.; Shen, D. Multimodal classification of Alzheimer's disease and mild cognitive impairment. Neuroimage 2011, 55, 856–867.
66. Liu, S.; Song, Y.; Cai, W.; Pujol, S.; Kikinis, R.; Wang, X.; Feng, D. Multifold Bayesian kernelization in Alzheimer's diagnosis. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Nagoya, Japan, 22–26 September 2013; Springer: Berlin/Heidelberg, Germany, 2013.
67. Zhang, Y.; Dong, Z.; Phillips, P.; Wang, S.; Ji, G.; Yang, J.; Yuan, T.-F. Detection of subjects and brain regions related to Alzheimer's disease using 3D MRI scans based on eigenbrain and machine learning. Front. Comput. Neurosci. 2015, 9, 66.
68. Feng, X.; Provenzano, F.A.; Small, S.A.; Initiative, F.T.A.D.N. A deep learning MRI approach outperforms other biomarkers of prodromal Alzheimer's disease. Alzheimer's Res. Ther. 2022, 14, 1–11.
69. Khojaste-Sarakhsi, M.; Haghighi, S.S.; Ghomi, S.; Marchiori, E. Deep learning for Alzheimer's disease diagnosis: A survey. Artif. Intell. Med. 2022, 130, 102332.
70. Reith, F.H.; Mormino, E.C.; Zaharchuk, G. Predicting future amyloid biomarkers in dementia patients with machine learning to improve clinical trial patient selection. Alzheimer's Dement. Transl. Res. Clin. Interv. 2021, 7, e12212.

71. Kim, Y.; Jiang, X.; Giancardo, L.; Pena, D.; Bukhbinder, A.S.; Amran, A.Y.; Schulz, P.E.; Initiative, A.D.N. Multimodal Phenotyping of Alzheimer's Disease with Longitudinal Magnetic Resonance Imaging and Cognitive Function Data. Sci. Rep. 2020, 10, 5527.
72. Chen, K.T.; Gong, E.; Macruz, F.B.D.C.; Xu, J.; Boumis, A.; Khalighi, M.; Poston, K.L.; Sha, S.J.; Greicius, M.D.; Mormino, E.; et al. Ultra–Low-Dose18F-Florbetaben Amyloid PET Imaging Using Deep Learning with Multi-Contrast MRI Inputs. Radiology 2019, 290, 649–656.
73. Monfared, A.A.T.; Houghton, K.; Zhang, Q.; Mauskopf, J.; Initiative, F.T.A.D.N. Staging Disease Severity Using the Alzheimer's Disease Composite Score (ADCOMS): A Retrospective Data Analysis. Neurol. Ther. 2022, 11, 413–434.
74. Sheng, J.H.; Wang, B.C.; Zhang, Q.; Zhou, R.G.; Wang, L.Y.; Xin, Y. Identifying and characterizing different stages toward Alzheimer's disease using ordered core features and machine learning. Heliyon 2021, 7, e07287.
75. Kazee, A.; Eskin, T.; Lapham, L.; Gabriel, K.; McDaniel, K.; Hamill, R. Clinicopathologic correlates in Alzheimer disease: Assessment of clinical and pathologic diagnostic criteria. Alzheimer Dis. Assoc. Disord. 1993, 7, 152–164.
76. Price, J.L.; Davis, P.; Morris, J.; White, D. The distribution of tangles, plaques and related immunohistochemical markers in healthy aging and Alzheimer's disease. Neurobiol. Aging 1991, 12, 295–312.
77. Bennett, D.A.; Schneider, J.A.; Arvanitakis, Z.; Kelly, J.F.; Aggarwal, N.T.; Shah, R.; Wilson, R.S. Neuropathology of older persons without cognitive impairment from two community-based studies. Neurology 2006, 66, 1837–1844.
78. Gopinadhan, A.; Prasanna, G.; Anbarasu, S. AD-EHS: Alzheimer's disease severity detection using efficient hybrid image segmentation. Adv. Eng. Softw. 2022, 173, 103234.
79. Li, F.; Tran, L.; Thung, K.-H.; Ji, S.; Shen, D.; Li, J. A Robust Deep Model for Improved Classification of AD/MCI Patients. IEEE J. Biomed. Health Inform. 2015, 19, 1610–1616.
80. Suk, H.-I.; Wee, C.-Y.; Lee, S.-W.; Shen, D. State-space model with deep learning for functional dynamics estimation in resting-state fMRI. Neuroimage 2016, 129, 292–307.
81. Fisher, C.K.; Smith, A.M.; Walsh, J.R. Machine learning for comprehensive forecasting of Alzheimer's Disease progression. Sci. Rep. 2019, 9, 13622.
82. Suk, H.-I.; Lee, S.-W.; Shen, D. Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis. Neuroimage 2014, 101, 569–582.
83. Razavi, F.; Tarokh, M.J.; Alborzi, M. An intelligent Alzheimer's disease diagnosis method using unsupervised feature learning. J. Big Data 2019, 6, 32.
84. Baumgartner, C.F.; Koch, L.M.; Can Tezcan, K.; Xi Ang, J.; Konukoglu, E. Computer aided Alzheimer's disease diagnosis by an unsupervised deep learning technology. Neurocomputing 2020, 392, 296–304.
85. Wang, Z.; Xin, J.; Wang, Z.; Gu, H.; Zhao, Y.; Qian, W. Computer-Aided Dementia Diagnosis Based on Hierarchical Extreme Learning Machine. Cogn. Comput. 2020, 13, 34–48.
86. Majumdar, A.; Singhal, V. Noisy deep dictionary learning: Application to Alzheimer's Disease classification. In Proceedings of the 2017 International Joint Conference on Neural Networks (IJCNN), Anchorage, AK, USA, 14–19 May 2017
87. Cheng, J.; Dalca, A.V.; Fischl, B.; Zöllei, L. Cortical surface registration using unsupervised learning. arXiv 2020, arXiv:2004.04617.
88. Son, H.J.; Oh, J.S.; Oh, M.; Kim, S.J.; Lee, J.-H.; Roh, J.H.; Kim, J.S. The clinical feasibility of deep learning-based classification of amyloid PET images in visually equivocal cases. Eur. J. Nucl. Med. 2019, 47, 332–341.
89. Wang, S.-H.; Phillips, P.; Sui, Y.; Liu, B.; Yang, M.; Cheng, H. Classification of Alzheimer's Disease Based on Eight-Layer Convolutional Neural Network with Leaky Rectified Linear Unit and Max Pooling. J. Med. Syst. 2018, 42, 85. Oseledets, I.V. Tensor-train decomposition. SIAM J. Sci. Comput. 2011, 33, 2295–2317
90. Aderghal, K.; Benois-Pineau, J.; Afdel, K. Classification of sMRI for Alzheimer's disease Diagnosis with CNN: Single Siamese Networks with 2D+? Approach and Fusion on ADNI. In Proceedings of the 2017 ACM on International Conference on Multimedia Retrieval, Bucharest, Romania, 6–9 June 2017.
91. Wang, S.-H.; Phillips, P.; Sui, Y.; Liu, B.; Yang, M.; Cheng, H. Classification of Alzheimer's Disease Based on Eight-Layer Convolutional Neural Network with Leaky Rectified Linear Unit and Max Pooling. J. Med. Syst. 2018, 42, 85.
92. Tang, Z.; Chuang, K.V.; DeCarli, C.; Jin, L.-W.; Beckett, L.; Keiser, M.J.; Dugger, B.N. Interpretable classification of Alzheimer's disease pathologies with a convolutional neural network pipeline. Nat. Commun. 2019, 10, 2173.
93. Choi, H.-S.; Choe, J.Y.; Kim, H.; Han, J.W.; Chi, Y.K.; Kim, K.; Hong, J.; Kim, T.; Kim, T.H.; Yoon, S.; et al. Deep learning based low-cost high-accuracy diagnostic framework for dementia using comprehensive neuropsychological assessment profiles. BMC Geriatr. 2018, 18, 234.
94. Ieracitano, C.; Mammone, N.; Hussain, A.; Morabito, F.C. A Convolutional Neural Network based self-learning approach for classifying neurodegenerative states from EEG signals in dementia. In Proceedings of the 2020 International Joint Conference on Neural Networks (IJCNN), Glasgow, UK, 19–24 July 2020.
95. Pan, X.; Phan, T.-L.; Adel, M.; Fossati, C.; Gaidon, T.; Wojak, J.; Guedj, E. Multi-View Separable Pyramid Network for AD Prediction at MCI Stage by 18F-FDG Brain PET Imaging. IEEE Trans. Med. Imaging 2020, 40, 81–92.
96. Alavi, A.; Ruffalo, M.; Parvangada, A.; Huang, Z.; Bar-Joseph, Z. A web server for comparative analysis of single-cell RNA-seq data. Nat. Commun. 2018, 9, 4768.
97. Islam, J.; Zhang, Y. Understanding 3D CNN Behavior for Alzheimer's Disease Diagnosis from Brain PET Scan. arXiv 2019, arXiv:1912.04563.
98. Duc, N.T.; Ryu, S.; Qureshi, M.N.I.; Choi, M.; Lee, K.H.; Lee, B. 3D-Deep Learning Based Automatic Diagnosis of Alzheimer's Disease with Joint MMSE Prediction Using Resting-State fMRI. Neuroinformatics 2019, 18, 71–86.
99. Basaia, S.; Agosta, F.; Wagner, L.; Canu, E.; Magnani, G.; Santangelo, R.; Filippi, M. Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks. NeuroImage Clin. 2019, 21, 101645.
100. Qiu, S.; Joshi, P.S.; Miller, M.I.; Xue, C.; Zhou, X.; Karjadi, C.; Chang, G.H.; Joshi, A.S.; Dwyer, B.; Zhu, S.; et al. Development and validation of an interpretable deep learning framework for Alzheimer's disease classification. Brain 2020, 143, 1920–1933.
101. Choi, H.; Initiative, F.T.A.D.N.; Kim, Y.K.; Yoon, E.J.; Lee, J.-Y.; Lee, D.S. Cognitive signature of brain FDG PET based on deep learning: Domain transfer from Alzheimer's disease to Parkinson's disease. Eur. J. Nucl. Med. 2019, 47, 403–412.
102. Basaia, S.; Agosta, F.; Wagner, L.; Canu, E.; Magnani, G.; Santangelo, R.; Filippi, M. Studying the manifold structure of Alzheimer's
103. Disease: A deep learning approach using convolutional autoencoders. IEEE J. Biomed. Health Inform. 2019, 24, 17–26.
104. Payan, A.; Montana, G. Predicting Alzheimer's disease: A neuroimaging study with 3D convolutional neural networks. arXiv2015, arXiv:1502.02506.
105. Ge, C.; Qu, Q.; Gu, I.Y.-H.; Jakola, A.S. Multiscale Deep Convolutional Networks for Characterization and Detection of Alzheimer's Disease Using MR images. In Proceedings of the 2019 IEEE International Conference on Image Processing (ICIP), Taipei, Taiwan,22–25 September 2019.
106. Islam, J.; Zhang, Y. Early Diagnosis of Alzheimer's Disease: A Neuroimaging Study with Deep Learning Architectures. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Salt Lake City, UT, USA, 18–22 June 2018.
107. Liu, S.; Yadav, C.; Fernandez-Granda, C.; Razavian, N. On the design of convolutional neural networks for automatic detection of Alzheimer's disease. In Proceedings of the Machine Learning for Health Workshop, Vancouver, BC, Canada, 8–14 December 2020.

108. Wang, H.; Shen, Y.; Wang, S.; Xiao, T.; Deng, L.; Wang, X.; Zhao, X. Ensemble of 3D densely connected convolutional network for diagnosis of mild cognitive impairment and Alzheimer's disease. Neurocomputing 2018, 333, 145–156.
109. Spasov, S.; Passamonti, L.; Duggento, A.; Liò, P.; Toschi, N. A parameter-efficient deep learning approach to predict conversion from mild cognitive impairment to Alzheimer's disease. Neuroimage 2019, 189, 276–287.
110. Yang, Z.; Zhuang, X.; Mishra, V.; Sreenivasan, K.; Cordes, D. CAST: A multi-scale convolutional neural network based automated hippocampal subfield segmentation toolbox. Neuroimage 2020, 218, 116947.
111. Pang, S.; Feng, Q.; Lu, Z.; Jiang, J.; Zhao, L.; Lin, L.; Li, X.; Lian, T.; Huang, M.; Yang, W. Hippocampus Segmentation Based on Iterative Local Linear Mapping with Representative and Local Structure-Preserved Feature Embedding. IEEE Trans. Med. Imaging 2019, 38, 2271–2280.
112. Jie, B.; Liu, M.; Lian, C.; Shi, F.; Shen, D. Designing weighted correlation kernels in convolutional neural networks for functional connectivity based brain disease diagnosis. Med. Image Anal. 2020, 63, 101709
113. Liu, M.; Initiative, T.A.D.N.; Cheng, D.; Wang, K.; Wang, Y. Multi-Modality Cascaded Convolutional Neural Networks for Alzheimer's Disease Diagnosis. Neuroinformatics 2018, 16, 295–308.
114. Li, J.; Rong, Y.; Meng, H.; Lu, Z.; Kwok, T.; Cheng, H. Tatc: Predicting Alzheimer's disease with actigraphy data. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, London, UK, 19–23 August 2018
115. Cui, R.; Liu, M.; Li, G. Longitudinal analysis for Alzheimer's disease diagnosis using RNN. In Proceedings of the 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), Washington, DC, USA, 4–7 April 2018. 288.
116. Feng, C.; Elazab, A.; Yang, P.; Wang, T.; Lei, B.; Xiao, X. 3D convolutional neural network and stacked bidirectional recurrent neural network for Alzheimer's disease diagnosis. In Proceedings of the International Workshop on Predictive Intelligence in MEdicine, Granada, Spain, 16 September 2018; Springer: Berlin/Heidelberg, Germany, 2018. 289.
117. Cheng, D.; Liu, M. Combining convolutional and recurrent neural networks for Alzheimer's disease diagnosis using PET images. In Proceedings of the 2017 IEEE International Conference on Imaging Systems and Techniques (IST), Beijing, China, 18–20 October 2017.
118. Xia, Z.; Yue, G.; Xu, Y.; Feng, C.; Yang, M.; Wang, T.; Lei, B. A Novel End-to-End Hybrid Network for Alzheimer's Disease Detection Using 3D CNN and 3D CLSTM. In Proceedings of the 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), Iowa City, IA, USA, 3–7 April 2020
119. Nguyen, M.; He, T.; An, L.; Alexander, D.C.; Feng, J.; Yeo, B.T. Predicting Alzheimer's disease progression using deep recurrent neural networks. NeuroImage 2020, 222, 117203.
120. Tang, F.; Uchendu, I.; Wang, F.; Dodge, H.H.; Zhou, J. Scalable diagnostic screening of mild cognitive impairment using AI dialogue agent. Sci. Rep. 2020, 10, 5732.
121. Pfau, D.; Vinyals, O. Connecting generative adversarial networks and actor-critic methods. arXiv 2016, arXiv:1610.01945.
122. Capecci, E.; Doborjeh, Z.G.; Mammone, N.; La Foresta, F.; Morabito, F.C.; Kasabov, N. Longitudinal study of alzheimer's disease degeneration through EEG data analysis with a NeuCube spiking neural network model. In Proceedings of the 2016 International Joint Conference on Neural Networks (IJCNN), Vancouver, BC, Canada, 24–29 July 2016
123. Lundervold, A.S.; Lundervold, A. An overview of deep learning in medical imaging focusing on MRI. Z. Med. Phys. 2019, 29, 102–127
124. Morid, M.A.; Borjali, A.; Del Fiol, G. A scoping review of transfer learning research on medical image analysis using ImageNet. Comput. Biol. Med. 2020, 128, 104115
125. Kam, H.J.; Kim, H.Y. Learning representations for the early detection of sepsis with deep neural networks. Comput. Biol. Med. 2017, 89, 248–255.
126. Wang, S.-H.; Zhan, T.-M.; Chen, Y.; Zhang, Y.; Yang, M.; Lu, H.-M.; Wang, H.-N.; Liu, B.; Phillips, P. Multiple Sclerosis Detection Based on Biorthogonal Wavelet Transform, RBF Kernel Principal Component Analysis, and Logistic Regression. IEEE Access 2016, 4, 7567–7576.
127. Zhang, X.; Yan, L.-F.; Hu, Y.-C.; Li, G.; Yang, Y.; Han, Y.; Sun, Y.-Z.; Liu, Z.-C.; Tian, Q.; Han, Z.-Y.; et al. Optimizing a machine learning based glioma grading system using multi-parametric MRI histogram and texture features. Oncotarget 2017, 8, 47816–47830.
128. Zhang, Y.; Wang, S.; Sui, Y.; Yang, M.; Liu, B.; Cheng, H.; Sun, J.; Jia, W.; Phillips, P.; Gorriz, J.M. Multivariate Approach for Alzheimer's Disease Detection Using Stationary Wavelet Entropy and Predator-Prey Particle Swarm Optimization. J. Alzheimer's Dis. 2018, 65, 855–869.
129. Zhuang, F.; Qi, Z.; Duan, K.; Xi, D.; Zhu, Y.; Zhu, H.; Xiong, H.; He, Q. A Comprehensive Survey on Transfer Learning. Proc. IEEE 2021, 109, 43–76.
130. Yosinski, J.; Clune, J.; Bengio, Y.; Lipson, H. How transferable are features in deep neural networks? In Proceedings of the Advances in Neural Information Processing Systems, NIPS 2014, Montréal, QC, Canada, 8–13 December 2014.
131. Mårtensson, G.; Ferreira, D.; Granberg, T.; Cavallin, L.; Oppedal, K.; Padovani, A.; Rektorova, I.; Bonanni, L.; Pardini, M.; Kramberger, M.G.; et al. The reliability of a deep learning model in clinical out-of-distribution MRI data: A multicohort study. Med. Image Anal. 2020, 66, 101714.
132. Li, Q.; Xing, X.; Sun, Y.; Xiao, B.; Wei, H.; Huo, Q.; Zhang, M.; Zhou, X.S.; Zhan, Y.; Xue, Z.; et al. Novel iterative attention focusing strategy for joint pathology localization and prediction of MCI progression. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Shenzhen, China, 13–17 October 2019; Springer: Berlin/Heidelberg, Germany, 2019.
133. Hoeffding, W. Probability inequalities for sums of bounded random variables. In The Collected Works of Wassily Hoeffding; Springer: Berlin/Heidelberg, Germany, 1994; pp. 409–426.
134. Kawaguchi, K.; Kaelbling, L.; Bengio, Y. Generalization in deep learning. arXiv 2017, arXiv:1710.05468.
135. Jin, P.; Lu, L.; Tang, Y.; Karniadakis, G.E. Quantifying the generalization error in deep learning in terms of data distribution and neural network smoothness. Neural Netw. 2020, 130, 85–99. [CrossRef]
136. Geirhos, R.; Temme, C.R.; Rauber, J.; Schütt, H.H.; Bethge, M.; Wichmann, F.A. Generalisation in humans and deep neural networks. In Proceedings of the Advances in Neural Information Processing Systems, NIPS 2018, Montréal, QC, Canada, 2–8 December 2018.