



Review

A Survey of Deep Learning for Alzheimer's Disease

Qinghua Zhou¹, Jiaji Wang¹ , Xiang Yu¹, Shuihua Wang¹ and Yudong Zhang^{1,2,*}

¹ School of Computing and Mathematical Sciences, University of Leicester, Leicester LE1 7RH, UK; qz105@le.ac.uk (Q.Z.); jw933@le.ac.uk (J.W.); xy144@le.ac.uk (X.Y.); shuihuawang@ieee.org (S.W.)

² Department of Information Systems, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah 21589, Saudi Arabia

* Correspondence: yudongzhang@ieee.org; Tel.: +44-754-870-0453

Abstract: Alzheimer's and related diseases are significant health issues of this era. The interdisciplinary use of deep learning in this field has shown great promise and gathered considerable interest. This paper surveys deep learning literature related to Alzheimer's disease, mild cognitive impairment, and related diseases from 2010 to early 2023. We identify the major types of unsupervised, supervised, and semi-supervised methods developed for various tasks in this field, including the most recent developments, such as the application of recurrent neural networks, graph-neural networks, and generative models. We also provide a summary of data sources, data processing, training protocols, and evaluation methods as a guide for future deep learning research into Alzheimer's disease. Although deep learning has shown promising performance across various studies and tasks, it is limited by interpretation and generalization challenges. The survey also provides a brief insight into these challenges and the possible pathways for future studies.

Keywords: deep learning; Alzheimer's disease; mild cognitive impairment; neural networks; recent advances



Citation: Zhou, Q.; Wang, J.; Yu, X.; Wang, S.; Zhang, Y. A Survey of Deep Learning for Alzheimer's Disease. *Mach. Learn. Knowl. Extr.* **2023**, *5*, 611–668. <https://doi.org/10.3390/make5020035>

Academic Editor: Abdulhamit Subasi

Received: 18 April 2023

Revised: 24 May 2023

Accepted: 30 May 2023

Published: 9 June 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Deep learning is a field of study that shows great promise for medical image analysis and knowledge discovery that is approaching clinicians' performance in a growing range of tasks [1–3]. The interdisciplinary study of Alzheimer's disease (AD) and deep learning have been a focus of interest for the past 13 years. This paper aims to survey the most current state of deep learning studies related to multiple aspects of research in Alzheimer's disease, ranging from current detection methods to pathways of generalization and interpretation. This section will first provide the current definition of AD and clinical diagnostic methods to provide the basis for deep learning. We then detail this interdisciplinary study's main areas of interest and current challenges.

1.1. Alzheimer's Disease and Mild Cognitive Impairment

Alzheimer's disease is the most common form of dementia and a significant health issue of this era [4]. Brookmeyer et al. [5] predicted that more than 1% of the world population would be affected by AD or related diseases by 2050, with a significant proportion of this cohort requiring a high level of care. AD usually starts from middle-to-old age as a chronic neurodegenerative disorder, but rare cases of early-onset AD can affect individuals of 45–64 years old [6]. AD leads to cognitive decline symptoms: memory impairment [7], language dysfunction [8], and decline in cognition and judgment [9]. An individual with symptoms may require moderate to constant assistance in day-to-day life, depending on the stage of disease progression. These symptoms severely affect patients' quality of life (QOL) and their families. Studies into cost-of-illness for dementia and AD reveal that the higher societal need for elderly care significantly increases overall social-economic pressure [10].

The biological process that leads to AD may begin more than 20 years before symptoms appear [11]. The current understanding of AD pathogenesis is based on amyloid peptide deposition and the accumulation and phosphorylation of tau proteins around neurons [12–14], which leads to neurodegeneration and eventual brain atrophy. Factors associated with AD include age, genetic predisposition [15], Down's syndrome [16], brain injuries [17], and cardiorespiratory fitness [18–20]. AD-related cognitive impairment can be broadly separated into three stages: (1) preclinical AD, where measurable changes in the brain, cerebral spinal fluid (CSF), and blood plasma can be detected; (2) mild cognitive impairment (MCI) due to AD, where biomarker evidence of AD-related brain change can be present; and (3) dementia due to AD, where changes in the brain are evident and noticeable memory, thinking and behavioural changes appear and impair an individual's daily function.

The condition most commonly associated with AD is mild cognitive impairment (MCI), the pre-dementia stage of cognitive impairment. However, not all cases of MCI develop into AD. Since no definite pathological description exists, MCI is currently perceived as the level of cognitive impairment above natural age-related cognitive decline [21,22]. Multiple studies have analyzed the demographics and progression of MCI and have found the following: 15–20% of people age 65 or older have MCI from a range of possible causes [23]; 15% of people age 65 or older with MCI developed dementia at two years follow-up [24]; and 32% developed AD and 38% developed dementia at five years follow-up [25,26]. The early diagnosis of MCI and its subtypes can lead to early intervention, which can profoundly impact patient longevity and QOL [27]. Therefore, better understanding the condition and developing effective and accurate diagnostic methods is of great public interest.

1.2. Diagnostic Methods and Criteria

The current standard diagnosis of AD and MCI is based on a combination of various methods. These methods include cognitive assessments such as the Mini-Mental State Examination [28–30], Clinical Dementia Rating [31,32], and Cambridge Cognitive Examination [33,34]. These exams usually take the form of a series of questions and are often performed with physical and neurological examinations. Medical and family history, including psychiatric history and history of cognitive and behavioral changes, are also considered in the diagnosis. Genetic sequencing for particular biomarkers, such as the APOE-e4 allele [35], is used to determine genetic predisposition.

Neuroimaging is commonly used to inspect various signs of brain changes and exclude other potential causes. Structural magnetic resonance and diffusion tensor imaging are widely applied to check for evidence of symptoms of brain atrophy. Various forms of computed tomography (CT) are also used in AD and MCI diagnosis. Regarding positron emission tomography (PET), FDG-PET [36] inspects brain glucose metabolism, while amyloid-PET is applied to measure beta-amyloid levels. Single-photon emission computed tomography [37] (SPECT) is likely to produce false-positive results and is inadequate in clinical use. However, SPECT variants can be potentially used in diagnosis, e.g., 99 mTc-HMPAO SPECT [38,39]. At the same time, FP-CIT SPECT can visualize discrepancies in the nigrostriatal dopaminergic neurons [40]. In neuroimaging, a combination of multiple modalities is commonly used to utilize the functionality of each modality.

New diagnostic factors of CSF and blood plasma biomarkers have been reported in the literature and have been deployed in clinical practice in recent years. There are three main CSF and blood plasma biomarkers: Amyloid- β 42, t-tau, and p-tau. Other biomarkers include neurofilament light protein (NFL), and neuron-specific enolase (NSE, and HFABP [41,42]. CSF biomarkers are becoming a critical factor in AD diagnostic criteria in some practices. However, the actual 'ground truth' diagnosis of AD can only be made via post-mortem autopsy.

Before this century, the established diagnostic criteria were the NINCDS-ADRDA criteria [43,44]. These criteria were updated by the International Working Group (IWG) in 2007 to include requirements of at least one factor among MRI, PET, and CSF biomarkers [45]. A second revision was introduced in 2010 to include both pre-dementia and dementia phases [46]. This was followed by a third revision to include atypical prodromal Alzheimer's disease that shows cognition deficits other than memory impairment—IWG-2 [47]. Another independent set of criteria, the NIA-AA criteria, was introduced in 2011. These criteria include measures of brain amyloid, neuronal injury, and degeneration [48]. Individual criteria were introduced for each clinical stage, including pre-clinical [49,50], MCI [51,52], dementia [53,54], and post-mortem autopsy [55].

1.3. The Deep Learning Approach

Detailed preprocessing with refined extraction of biomarkers combined with statistical analysis is the accepted practice in current medical research. Risacher et al. [56] applied statistical analysis on biomarkers extracted using voxel-based morphometry and parcellation methods from T1-weighted MRI scans of AD, MCI, and HC. The study reveals statistical significance in multiple measures, including hippocampal volume and entorhinal cortex thickness. Qiu et al. [57] further confirmed this significance by analyzing regional volumetric changes through large deformation diffeomorphic metric mapping (LDDMM). Guevremont et al. [58] focused on robustly detecting microRNAs in plasma and used standardized analysis to identify microRNA biomarkers in different phases of Alzheimer's disease. This study and its statistical analysis yielded useful diagnostic markers reflecting the underlying disease pathology. The different biomarker information extracted was fed into statistical analysis methods with varying numbers of variables to detect changes in biomarkers in disease development [59]. Similar studies also employed other neuroimaging data, genetic data, and CSF biomarkers. These studies supported the use of MRI imaging biomarkers in AD [60] and MCI diagnosis [61], laying the basis for developing automatic diagnostic algorithms.

Machine learning has amassed great popularity among current automated diagnostic algorithms due to its adaptivity to data and the ability to generalize knowledge with lower requirements of expert experience. The study by Klöppel et al. [62] proved the validity of applying machine learning algorithms in diagnosing dementia through a performance comparison between the Support Vector Machine (SVM) classification of local grey matter volumes and human diagnosis by professional radiologists. Janousova et al. [63] proposed penalized regression with resampling to search for discriminative regions to aid Gaussian kernel SVM classification. The regions found by the study coincide with the previous morphological studies. These breakthroughs led to the development of many machine-learning algorithms for AD and MCI detection. Zhang et al. [64] proposed a kernel combination method for the fusion of heterogeneous biomarkers for classification with linear SVM. Liu et al. [65] proposed the Multifold Bayesian Kernelization (MBK) algorithm, where a Bayesian framework derives kernel weights and synthesis analysis provides the diagnostic probabilities of each biomarker. Zhang et al. [66] proposed the extraction of the eigenbrain using Welch's *t*-test (WTT) [67] combined with a polynomial kernel SVM [68] and particle swarm optimization (PSO) [69].

There has also been considerable interest in applying deep learning (DL), a branch of machine learning, to the field of AD and related diseases. Deep learning integrates the two-step feature extraction and classification process into neural networks, universal approximators based on backpropagation parameter training. [70]. Deep learning has made considerable advances in the domain of medical data, e.g., breast cancer [71], tuberculosis [72], and glioma [73]. Instead of hand-crafting features, models, and optimizers, deep learning leverages the layered structure of neural networks for the automated abstraction of various levels of features. For example, Feng et al. [74] used the proposed deep learning model to extract biomarkers for MRI in neuroimaging. The study demonstrates that the deep learning approach outperformed other neuroimaging biomarkers of amyloid and tau

pathology and neurodegeneration in prodromal AD. A visualization of the field of this survey is shown in Figure 1.

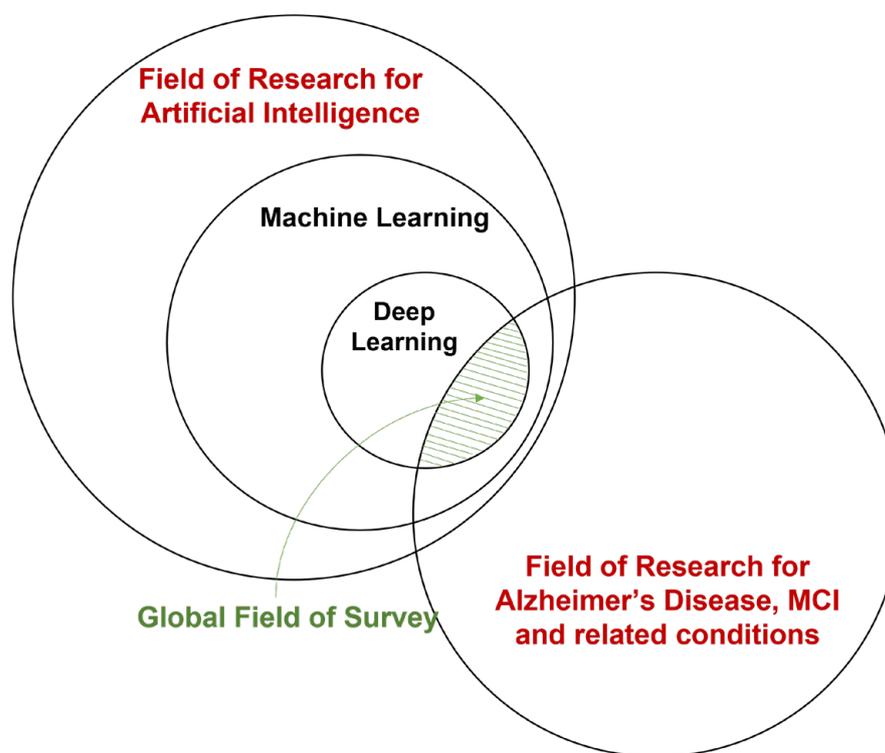


Figure 1. A broad overview of the field of this survey.

1.4. Areas of Interest

The primary aim of the surveyed deep learning studies in Alzheimer's and related diseases is detecting and predicting neurodegeneration to provide early detection and accurate prognosis to support treatment and intervention. The main interests of this interdisciplinary field can be roughly categorized into three areas:

1. Classification of various stages of AD. This area targets diagnosis or efficient progression monitoring. Current studies mostly focus on AD, MCI subtypes, and normal cognitive controls (NC). A few studies contain the subjective cognitive decline (SCD) stage before MCI.
2. Predicting MCI conversion. This area is mainly approached by formulating prediction as a classification problem, which usually involves defining MCI converters and non-converters based on a time threshold from the initial diagnosis. Some studies also aim at the prediction of time-to-conversion for MCI to AD.
3. Prediction of clinical measures. This area aims at producing surrogate biomarkers to reduce cost or invasivity, e.g., neuroimaging to replace lumbar puncture. Prediction of clinical measures, e.g., ADAS-Cog13 [75] and ventricular volume [76], is also used for longitudinal studies and attempts to achieve a more comprehensive evaluation of disease progression and model performance benchmarking.

There are also other areas of interest, including knowledge discovery, where studies attempt to understand AD through data [77]. Another area of interest is phenotyping and sample enrichment for clinical trials of treatments [78], where DL models are used to select patients that will likely respond to treatment and prevent ineffective or unnecessary treatment [79]. Interest also lies in segmentation and preprocessing, where DL models are applied to achieve higher performance or efficiency than conventional pipelines [80].

1.5. Challenges in Research

There is uncertainty in the diagnosis or prognosis of AD or related diseases with still developing diagnostic criteria and scientific understanding. DL-based approaches have already shown potential in the above areas of interest; however, there exists room for improvement and a range of challenges:

1. Numerical representation of the differences between AD stages. Monfared et al. [81] calculated the range of Alzheimer's disease composite scores to assess the severity of the cognitive decline in patients. Sheng et al. [82] made multiple classifications and concluded that the gap between late mild cognitive impairment and early mild cognitive impairment was small, whereas a greater difference exists between early and late MCI patients. Studies comparing clinical and post-mortem diagnoses have shown 10–20% false cases [83]. In addition, autopsy studies in individuals who were cognitively normal for their age found that ~30% had Alzheimer's-related brain changes in the form of plaque and tangles [84,85]. Sometimes the signs that distinguish AD, for example, brain shrinkage [86], can be found in a normal healthy brain of older people.
2. Difficulty in preprocessing. Preprocessing medical data, especially neuroimaging data, often requires complex pipelines. There is no set standard for preprocessing, while a broad range of processing options and relevant parameters exist. Preprocessing quality is also vastly based on the subjective judgment of clinicians.
3. Unavailability of a comprehensive dataset. Though the amount and variety of data available for AD and related diseases are abundant compared with many other conditions, the number of subjects is only moderate compared with large datasets such as Image-Net and is below the optimal requirements for generalization.
4. Differences in diagnostic criteria. The diagnostic criteria, or criteria for ground truth labels, can differ significantly between studies, especially in prior studies before new methods of diagnosis (e.g., CSF biomarkers [87] and genetic sequencing [88]) became accessible.
5. Lack of reproducibility. Most frameworks and models are not publicly available. Without open-source code, implementation details such as specific data cohort selection, preprocessing procedures and parameters, evaluation procedures, and metrics are usually lacking. These are all factors that can significantly impact results. Additionally, few comprehensive frameworks are designed for benchmarking different models based on the same preprocessing/processing and testing standards [89,90].
6. Lack of expert knowledge. Researchers adept at using DL often have no medical background, while medical data are significantly more complicated than natural images or language data. Therefore, these researchers lack expert knowledge, especially in preprocessing and identifying brain regions of interest (ROIs).
7. Generalizability and interpretability. Current DL models are plagued by information leakage and only provide limited measures of generalizability, the model's performance in real-world populations. The inherent 'black box' nature of neural networks impedes the interpretation of model functions and the subsequent feedback of knowledge for clinicians [91].
8. Other practical challenges include the subjectivity of cognitive assessments, the invasiveness of diagnostic techniques such as a lumbar puncture to measure CSF biomarkers and the high cost of neuroimaging such as MRI.

By analyzing the frequency of occurrence, influencing factors, and potential impact on research results for each challenge based on evidence and observations in the literature, we assign weights to each challenge in Table 1.

Table 1. Summary of challenges in applying DL to AD.

Challenge	Description	Weight (1–5)
Numerical representation of AD stages	Variability in Alzheimer’s disease composite scores and difficulty distinguishing between stages of cognitive impairment.	3
Difficulty in preprocessing	Complex pipelines for preprocessing medical data, lack of standardization, subjective judgment of clinicians.	3
Unavailability of a comprehensive dataset	Abundance of data for AD but moderate number of subjects, below optimal requirements for generalization.	2
Difference in diagnostic criteria	Variations in diagnostic criteria and ground truth labels between studies, impacting comparability of results.	3
Lack of reproducibility	Lack of publicly available frameworks, implementation details, and comprehensive benchmarking standards.	4
Lack of expert knowledge	Researchers with DL expertise may lack medical background, particularly in preprocessing and identifying brain regions.	2
Generalizability and interpretability	Limited measures of generalizability, ‘black box’ nature of neural networks, hindering model interpretation and feedback.	5
Practical challenges	Subjectivity of cognitive assessments, invasiveness and cost of diagnostic techniques such as lumbar puncture and MRI.	3

1.6. Survey Protocol

This survey covers DL studies related to AD or related diseases from 2010 to 2023. To identify literature related to our focus, we first queried the online libraries of IEEE, New York, NY, USA, Springer, Berlin/Heidelberg, Germany, and ScienceDirect, Amsterdam, The Netherlands and then concentrated search on:

1. Recognized journals, including Brain, Neuroimage, Medical Image Analysis, Alzheimer’s and Dementia, Nature Communications, and Radiology.
2. Conferences in computer vision and deep learning, including ACM, NeurIPS, CVPR, MICCAI, and ICCV.

The full list of search keywords is as follows: “Alzheimer’s”, “AD”, “Dementia”, “Mild Cognitive Impairment”, “MCI”, “Neural Networks”, “Deep Learning”, “Machine Learning”, “Learning”, “Big Data”, “Autoencoders”, “Generative”, “Multi-Modal”, “Interpretable”, “Explainable”. These keywords were used independently or in combination during the search process, which yielded 360 papers from various sources. A two-stage selection was performed, where the following conditions were first used to select the papers:

1. Related to Alzheimer’s disease, MCI, or other related diseases.
2. Related to deep learning, with the use of neural networks.
3. Contains valid classification/prediction metrics.
4. Utilizes a reasonable form of validation.
5. Written in English or contains a valid translation.
6. Contains a minimum of 180 individual subjects.

An additional constraint of subject number was applied in this survey, where 180 subjects correspond to a 0.01 chance of having an approximately 10% fluctuation in accuracy in generalization according to derivations from Hoeffding inequality [92]. However, this is only a basic requirement since the approximate generalization bound depends on the data available for evaluation and the independence assumptions between the classifier parameters and data. This condition is relaxed for studies using uncommon data types and functional MRI, where the available data are often limited compared with standard data types, e.g., MRI and PET. This selection stage yielded a total of 165 papers.

After the first stage of selection, the papers were evaluated on the year of publication, data source and type, preprocessing technique, feature extraction techniques, model architecture, platform, optimization protocol, and evaluation details. Papers from unknown sources or studies with apparent errors, e.g., information leakage, were excluded from the selection. This selection stage yielded a total of 83 papers. Similar to previous surveys, this survey mainly focused on neuroimaging data [93–96]. This review also expands on the work by Wen, Thibeu-Sutre, Diaz-Melo, Samper-González, Routier, Bottani, Dormont, Durreleman, Burgos and Colliot [89], which focuses on convolutional neural networks, to a broader range of supervised and unsupervised neural networks, including recent advances in graph and geometric neural networks. The survey protocol is visualized in Figure 2.

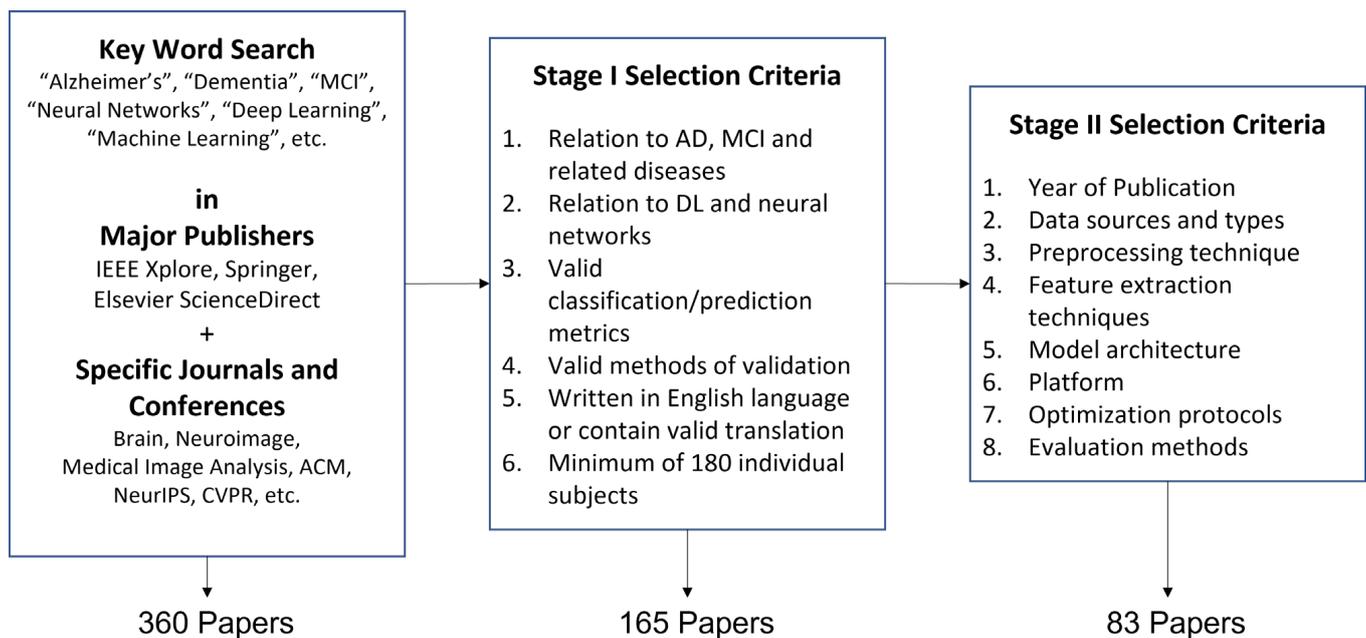


Figure 2. Survey Protocol.

The paper is organized as follows: Section 2 introduces the data types implemented in deep learning research and potential data sources; Section 3 provides detailed summaries of data preprocessing methods for the main data types, followed by four categories of data processing for neural network data input in Section 4. Sections 5–7 constitute the main body of deep learning architectures and methods included in this review, categorized into unsupervised, semi-supervised, and supervised learning methods; typical models and recent advances are included in each category, including recent developments in generative models, recurrent and graph neural networks. Section 8 introduces various techniques, including transfer learning, ensemble learning, and multi-modal fusion. Section 9 details training and evaluation protocols, while Sections 10 and 11 lay possible pathways for future research in interpretability and generalization. A taxonomy of the survey is shown in Figure 3.

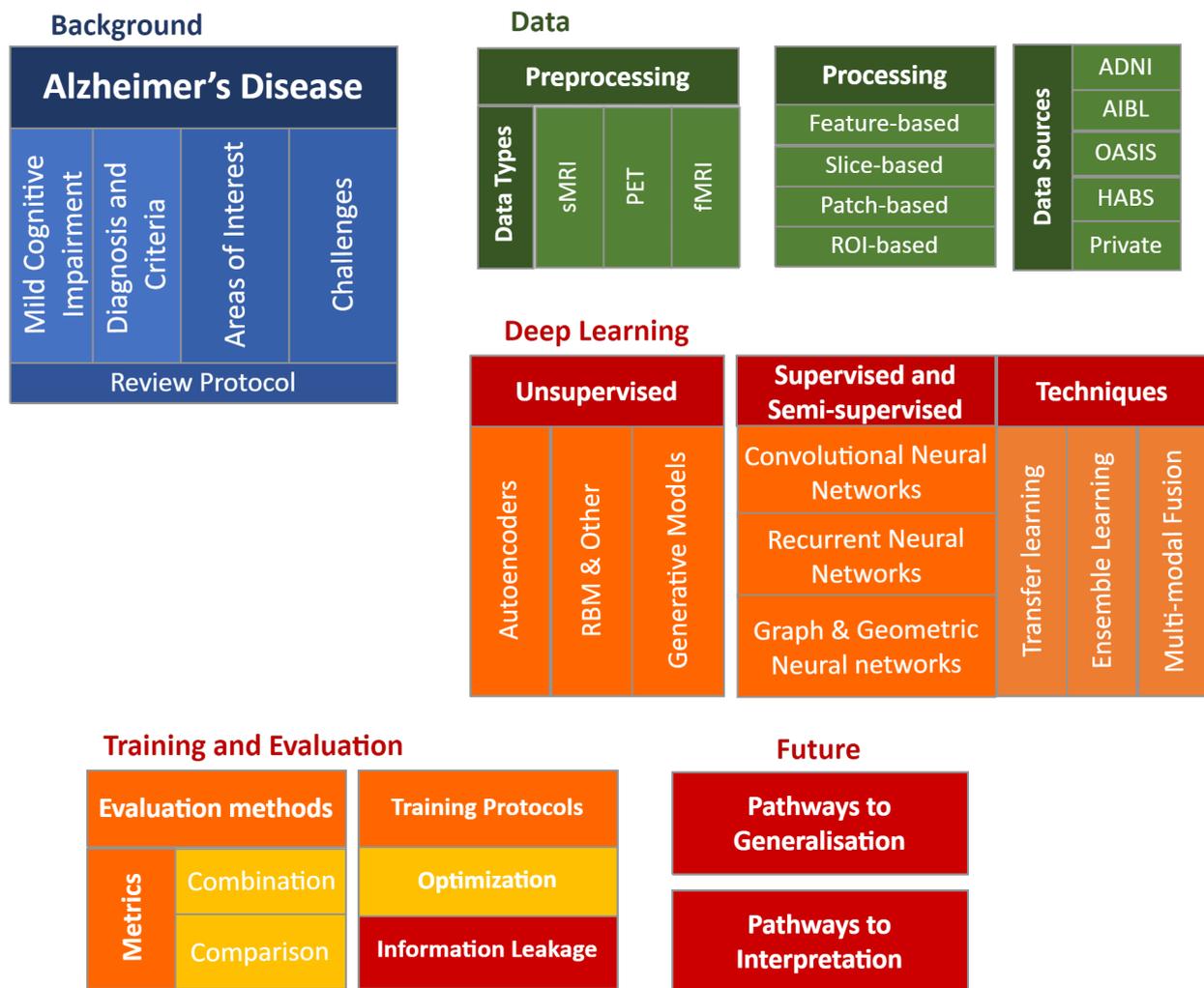


Figure 3. Overview of the survey content.

2. Data Types and Sources

Data issues are a core aspect of the deep learning approach, and the type and quantity of data directly impact model performance and potential generalizability. A large variety of data from numerous sources have been utilized in the reviewed studies. In this section, we summarize the main types of data available and the sources of these data.

2.1. Types of Data

The available data can be categorized into longitudinal data and cross-sectional data. Longitudinal data correspond to a subject’s disease progression data collected over time, while cross-sectional data are single-instance data that are time-independent. Longitudinal data can also be treated as independent data, time-series data, or comparative data. Demographics (Demo) is often a form of meta-data collected alongside other exams, primarily information regarding age, gender, and education. Neuroimaging data of various modalities are commonly collected. Common modalities include PET, MRI, and CT for diagnosis, and 3D-MRI [97], fMRI [98], and SPECT for research purposes. Various forms of cognitive assessments (CA) are also commonly available, including MMSE [29], CDR [99], ADAS-Cog [100,101], logical memory test [102,103], and postural kinematic analysis [104,105]. CSF, blood plasma biomarkers [106,107], and genetic data are available from several sources. Other less common data types include electroencephalography (EEG) for brain activity monitoring [106,107], mass spectra data collected through surface-enhanced laser desorption and ionization assay of saliva [108], and retinal imaging for

abnormalities [109]. Electronic health records have also been studied to screen dementia and AD [110,111]. Alternative data types such as speech [112,113], activity pattern monitoring [114,115], and eye-tracking [116] have also been studied with a deep learning approach. Few deep-learning-related comparative studies have been performed between different data modalities and types, especially for the less common data types [117].

2.2. Sources of Data

Several open libraries have been created in the past two decades, providing easier access for researchers to available data on subjects with AD or related diseases. One of the main libraries is Alzheimer's Disease Neuroimaging Initiative (ADNI) [118], a large longitudinal study aiming to develop novel biomarkers to detect AD and monitor disease progression. The original ADNI cohort was collected from 2004 to 2010 and contains T1-weighted MRI [119], FDG-PET, blood, and CSF biomarkers from 800 subjects [120]. Additional cohorts, ADNI-Go and ADNI-2, extended the longitudinal study of ADNI-1 while also encompassing a broader range of the stages of AD, adding 200 new subjects with early MCI [121]. A fourth cohort, ADNI3, with additional modalities targeting tau protein tangles, started in 2016 and is due to complete in 2022 [122].

ADNI is the most commonly used open library available for neuroimaging data. Another commonly used open library is the Open Access Series of Imaging Studies (OASIS), which includes a cross-sectional cohort (OASIS-1), a longitudinal cohort (OASIS-2) of demented or non-demented subjects MRI, and an additional longitudinal cohort (OASIS-3) provides MRI and PET in various modalities of 1098 subjects of normal cognition or AD [123]. While ADNI contains genomic data, OASIS only contains neuroimaging and neuropsychology data, i.e., cognitive assessments.

Other open libraries include the Harvard Aging Brain Study (HABS) [124] and Minimal Interval Resonance Imaging in Alzheimer's Disease (MIRIAD) [125]. These open libraries are essential in propagating studies in machine learning or deep learning for AD research. There also exist several local studies modeled similar to ADNI for data compatibility, including Japan ADNI (J-ADNI) [126], the Hong Kong Alzheimer's Disease Study [127], and the Australian Imaging Biomarkers and Lifestyle Study of Ageing (AIBL) [128]. Various institutes have established platforms to provide information and efficient access to available databases and libraries, including NeuGRID [129,130] and the Global Alzheimer's Association Interactive Network [131]: <http://www.gaain.org/> (accessed on 17 April 2023). We provide a shortlist of selected data sources in Table 2.

An alternative source of data for ML practitioners and researchers alike are the challenges hosted either by ADNI or other institutions, such as CADDementia [132], TAD-POLE [133,134], DREAM [135], and the Kaggle international challenge for automated prediction of MCI from MRI data. These challenges may provide pre-selection or preprocessed data, reducing the need for expert knowledge. A few studies have proposed to use brain age as a surrogate measure of cognitive decline and utilized databases of cognitively normal individuals, including UKBioBank [136], NKI, IXI [137], LifespanCN [138], and the Cambridge dataset [139]. Other sources of data that may be available include data from the International Genomics of Alzheimer's Project (IGAP) [140], the Korean Longitudinal Study on Cognitive Aging and Dementia (KLOSCAD) [141,142], the INSIGHT-preAD study [143,144], the Imaging Dementia—Evidence for Amyloid Scanning (IDEA) study [145], and the European version of ADNI—AddNeuroMed [146,147]. Institutes that hold private data collections of AD or related diseases include the National Alzheimer's Coordination Center, the Biobank of Beaumont Reference Laboratory, and IRCCS. As one of the most significant health crises of the era, many studies have collected data for AD and related diseases; therefore, the above-listed sources include only those most commonly used in reviewed literature and examples of alternative sources.

Table 2. Sources of AD and dementia data.

Library	Number of Subjects	Modalities	Link
ADNI	2750	MRI, PET, CSF, Genetic	http://adni.loni.usc.edu/ (accessed on 17 April 2023)
OASIS	1300+	MRI, PET	https://oasis-brains.org/ (accessed on 17 April 2023)
AIBL	1100+	MRI, PET, CSF, Genetic	https://aibl.csiro.au/ (accessed on 17 April 2023)
NACC	47,000+	Neuropathology, Genetic	https://www.alz.washington.edu/ (accessed on 17 April 2023)
EDSD	471	MRI, DTI, Genetic	https://www.neugrid2.eu/ (accessed on 17 April 2023)
ARWIBO	2700+	MRI, PET, Genetic	http://www.arwibo.it/ (accessed on 17 April 2023)
HABS	290	MRI, PET, Genetic	https://habs.mgh.harvard.edu/ (accessed on 17 April 2023)
KLOSCAD	6818	MRI, QOL, Behavioral	http://kloscad.com/ (accessed on 17 April 2023)
VITA	606	MRT, Genetic	https://www.neugrid2.eu/ (accessed on 17 April 2023)

3. Data Preprocessing

The deep learning approach can replace the feature-crafting step of machine learning and reduce the need for preprocessing. Data types such as clock-drawing test images [148], activity monitoring data [115], and speech audio files [112] can be processed in a similar way to natural images and time-series data. However, for the prevalent neuroimaging data, due to the complexity of data and the variety of established pipelines, data preprocessing is a significant component in current DL studies. This survey will focus on imaging data, the most prevalent data category in the intersection between AD and deep learning. Differences in the organization of data adds to the difficulty of preprocessing. Gorgolewski et al. [149] proposed the Brain Imaging Data Structure (BIDS) repository structure. Conversion to a standard data structure such as BIDs is essential when using multiple modalities and data sources.

3.1. Structural MRI Data

MRI is a safe, non-invasive medical imaging technique. High-quality medical images can be generated with good spatial resolution while minimizing patient harm using a powerful magnetic field, radio waves, and a computer. Structural MRI (sMRI) and functional MRI (fMRI) are different MRI techniques used to study the brain. sMRI is a non-invasive brain imaging technique that can investigate changes in brain structure [150]. Changes in brain structure due to worsening cognitive impairment may include atrophy of specific brain regions, loss of brain tissue, and changes in the shape and size of certain brain structures [151,152].

MRI machines are highly complex medical equipment that can vary individually. Inhomogeneity of the B1-field in MRI machines can cause artifact signals known as the bias field. Bias field correction is often the first step in MRI data preprocessing [153,154], usually using B1-scans to correct for the non-uniformity in the MR image. Similarly, gradient non-linearity can be corrected with displacement information and phase mapping, e.g., Gradwarp. These corrections are often in-built into the MRI systems, and its outputs are often the raw data available from data sources. Intensity normalization is essential to mitigate the difference between multiple MRI machines, especially in large-scale multi-center studies or when combining data from multiple sources. The most common method

found in AD-related papers is the N3 nonparametric non-uniform intensity normalization algorithm [155,156], a histogram peak sharpening algorithm that corrects intensity non-uniformity without establishing a tissue model. In some studies, motion correction is used to correct for subject motion artifacts produced during scanning sessions [157].

Brain extraction is a common MRI preprocessing component. It is the removal of non-brain components from the MRI scan. Skull-stripping removes the skull component, e.g., through bootstrapping histogram-based threshold estimations. Other similar procedures include cerebellum removal and neck removal. The extracted brain images are often registered to a brain anatomical template for spatial normalization, usually performed after brain extraction. Registration can be categorized based on the deformation allowed into affine registration and non-linear registration. Affine registration includes linear registration, while non-linear registration allows for local deformations [158–160]. A standard template used is MNI-152 based on 152 subjects [161], while some studies use alternative templates such as Colin27. A potential challenge in this process is that the selected control subjects' age does not match the AD subjects' older age and corresponding brain atrophy. Some studies resolve this issue by constructing study-specific template space based on training data, which can also be aligned with standard templates. Other alignments include AC-PC correction, the alignment of the images with the anterior commissure (AC) and posterior commissure (PC) on the same geometric plane. AC-PC correction can be performed with resampling to $256 \times 256 \times 256$ and intensity normalization using the N3 algorithm with MIPAV. Studies have shown that linear or affine normalization is potentially sufficient for deep learning models [162,163], while other studies have shown that non-rigid registration can improve performance.

Another potential MRI preprocessing procedure is brain region segmentation, the division of the brain MRI into known anatomical regions. This step is usually performed to isolate brain regions related to AD, e.g., grey matter of the medial temporal lobe and the hippocampal region. Segmentation can be performed manually by outlining bounding boxes or precise pixel boundaries. Ideal practices include randomizing the samples and segmenting multiple times or segmentation with multiple expert radiologists [164]. However, manual segmentation is time intensive and not suitable for large datasets. Automated algorithms such as FSL FIRST [165] and the FreeSurfer pipeline can perform segmentation by registering to brain atlases, e.g., AAL. Other methods include using RAVENS maps produced by tissue-preserving image wrapping methods [166] and specific region segmentation, e.g., hippocampus segmentation with MALPEM [167]. With segmentation-based neural networks, multiple studies have applied the deep learning approach to hippocampus segmentation [168,169].

In AD-related studies, downsampling is often performed after preprocessing to reduce the dimensionality of input into the neural network, directly affecting the number of parameters and computational cost and achieving uniformity in input dimensions [170]. Smoothing is also often performed to further improve the signal-to-noise ratio [166] but it results in lower amplitude and increases peak bandwidth. Age correction also considers normal brain atrophy due to increasing age, similar to atrophy due to AD. A potential method to correct this effect is via a voxel-wise linear regression model after registration, which benefits overall model performance [167].

3.2. PET Data

PET utilizes a radioactive tracer to study the activity of cells and tissues in the body [171]. When studying neurological disorders, the tracer binds to specific proteins associated with the disease, such as amyloid beta, a hallmark of AD [172], and tau in the case of AD [173,174]. It can also help identify changes in glucose metabolism, which is altered in the brains of Alzheimer's patients. The preprocessing of PET images is similar to the preprocessing of structural MRI images described in Section 3.1. In AD-related studies, PET data are often used with MRI data due to combined collection in major studies, e.g., ADNI. Preprocessing up to image registration and segmentation is first performed

on the MRI image, while the PET images are registered to the corresponding MRI images through rigid alignment [166]. The post-segmentation steps of downsampling and smoothing are similar to those performed on MRI images. Studies independent of MRI follow either simplified preprocessing methods similar to MRI preprocessing [80,175,176] or only minimal preprocessing [177].

3.3. Functional MRI Data

Functional MRI (fMRI) is a type of magnetic resonance imaging designed to measure brain activity by monitoring blood flow within the brain. Instead of static, single-instance structural MRI, fMRI is temporal, consisting of a series of images. fMRI is used to study changes in brain function related to the disease. These changes can encompass altered connectivity between distinct regions of the brain [178] and variations in how the brain reacts to stimuli [179]. The fMRI can investigate alterations in memory and attention associated with cognitive impairment in MCI and AD [180]. Both sMRI and fMRI can be utilized to monitor the progression of the disease by detecting changes in specific brain regions over time [181,182].

Therefore, preprocessing steps in addition to the preprocessing procedures for structural MRI mentioned in Section 3.1 are required. Slice time correction is required to achieve the time-series exact timing, where fMRI may need to be first corrected for the temporal offset between each scan instance. More extended periods of fMRI scanning and the collection of multiple images in a single session increase the chance of head motion artifacts. Therefore, fMRI scans require additional filtering or correction for motion. Head motion correction of fMRI is usually performed through the spatial alignment to the first scan, or scan of choice, before spatial normalization. High-pass and low-pass filters can also be introduced to the temporal domain to control the fMRI data frequency and period [183]. The preprocessing of fMRI data can be automated using the SPM REST Toolkit, DPABI, or FreeSurfer. Data redundancy reduction methods are often applied to fMRI data; these can be categorized as methods based on common spatial pattern (CSP) or brain functional network (BFN). CSP-based methods produce spatial filters that maximize one group's variance while minimizing another [184]. BFN-based methods use ROI segmentation to construct a brain network where the ROI features are vertices and the functional connections are edges. Brain networks can also be constructed by calculating ROI correlations after segmentation [185]. A recent study has also applied the deep learning approach to construct weighted correlation kernels integrated into neural network architecture to extract dynamic functional connectivity networks [186].

4. Data Processing

Data processing is essential to the deep learning approach, significantly influencing model architecture and performance. Compared with traditional machine learning feature extraction, data processing for deep learning focuses on processing input data to neural networks instead of establishing quantified representations. Data processing aims to preserve and emphasize critical discriminatory information within the preprocessed or raw data while standardizing the input for model readability across samples and modalities. The processing can be categorized into common types of model inputs. A basic summary of the most commonly used input types is illustrated in Figure 4.

4.1. Feature-Based

Feature-based approaches are performed on individual features of the provided data. For neuroimaging, this is also known as the voxel-based approach [96], which is applied to individual image voxels of spatially normalized images. Space co-alignment between images is also essential to ensure comparability between individual voxels across the dataset. To limit the amount of input information, tissue segmentation of the grey matter probability maps is often performed. Machine learning extraction of texture, shape, or other features can also be performed to reduce dimensionality and form an ML-DL hybrid approach [187].

Voxel-based methods for neuroimaging data retain global 2D or 3D information but ignore local information as it treats the entire brain uniformly, regardless of anatomical features. For 3D scans and genetic data with large transcription quantity, higher dimensions of input result in high computational cost; dimensionality reduction through either feature selection or transformations is common. Feature-based approaches are used for most alternative data types such as cognitive assessments, CSF, serum, and genetic biomarkers. Longitudinal data for these types of time-series data such as EEG, activity, and speech requires more stringent processing for sample completeness, e.g., imputation for missing data and time-stamp alignment [188].

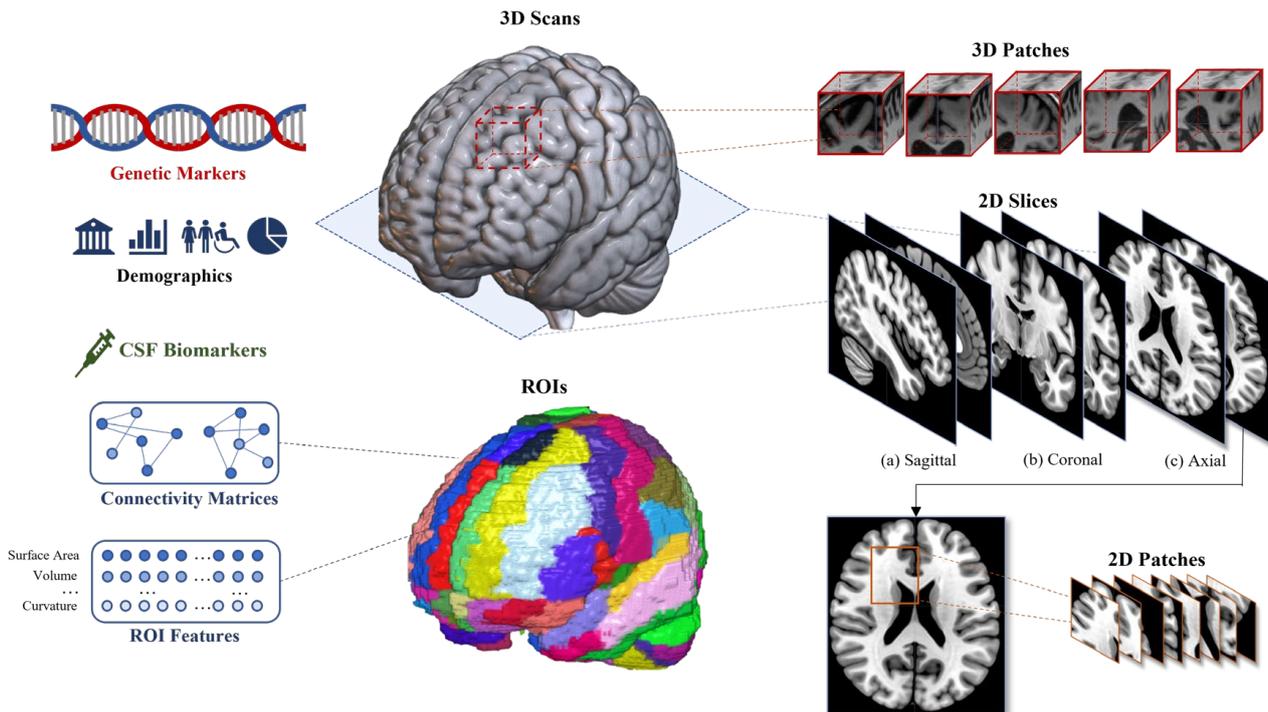


Figure 4. Summary of the most commonly used data types, where an sMRI is provided as an example of neuroimaging data. Feature-based data: demographics, CSF biomarkers, genetic markers, and 3D scans. Slice-based data: 2D slices from three views. Patch-based data: 2D and 3D patches. ROI-based data: ROI features and connectivity matrices.

4.2. Slice-Based

Slice-based approaches use 2D images or data. For 3D information, slice-based approaches assume that 2D information is a sufficient representation of the required information. Practical clinical diagnosis is often based upon a limited number of 2D slices instead of a complete 3D image. Some studies extract single or multiple 2D slices along the sagittal, axial, and coronal planes from the 3D scan. Slices from the axial plane are most commonly extracted, whereas the coronal view might contain the most critical AD-related regions. The selection of slices from 3D scans usually focuses on a particular dissection of the brain and the anatomical components it contains, e.g., the sagittal slices of the hippocampus are a known region of interest. Some studies used sorting procedures to find the most valuable slices, e.g., entropy sorting with greyscale histograms (Choi and Lee 2020). Slice-based approaches can be less computationally expensive than feature-based approaches with limited information quantity. However, the drawback of slice-based approaches is the loss of global and 3D geometric structures. Studies attempt to compensate for this loss by using multiple slices from multiple views, e.g., slices from three projections that show the hippocampal region [189], and multiple modalities, e.g., combining slices from MRI and PET.

4.3. Patch-Based

Instead of using all features or 2D slices, we can use regions of predefined size as input to the model, which is known as the patch-based approach. These regions can be 2D or 3D to suit model requirements [166]. Lin, Tong, Gao, Guo, Du, Yang, Guo, Xiao, Du and Qu [167] combined 2D greyscale patches of the hippocampal region into RGB patches. The patch-based approach can provide a larger sample size, equivalent to the number of patches, in the training procedure. Individual patches have a smaller memory footprint with lower input dimensions, reducing the computational resources required for training. However, additional resources for reconstructing sample-level results will mean costs to efficiency during testing and application. The challenge of patch-based approaches is capturing the most informative regions. Region selection is a vital component in this category; this includes the size of patches, the choice of overlap between the patches, and the choice of essential patches. Studies have attempted to use voxels' statistical significance [56,190] to find patching regions, while landmark-based methods perform patching around anatomically significant landmarks [191]. These patch-based approach is an intermediate form of voxel-based and ROI-based methods. Various kinds of approaches require patch-level data, including the use of a single or a low number of patches from each image for low input dimension models used to localize atrophy [192], patch-level sub-networks for hierarchical models [166], and ensemble learning through networks trained on defined regions.

4.4. ROI-Based

Patch-based methods have predefined regions, and sizes of extraction are often rigid, while ROI-based methods focus on anatomical regions of interest within the brain. These ROIs of anatomical function are often finely selected in the preprocessing stage of registration to brain atlases. The most common atlas used among the reviewed studies is the Automated Anatomical Labeling (AAL) atlas, which contains 93 ROIs. Other atlases include the Kabani reference work [193] and the Harvard-Oxford cortical and subcortical structural atlases [194,195]. Elastic registration, such as HAMMER, has higher registration performance [196,197]. After ROI extraction, the reviewed studies commonly use GM tissue mean intensities, or volumes, of brain ROIs as features from PET, MRI, fMRI or other modalities [198]. Other measures include subcortical volumes [199,200], grey matter densities [201,202], cortical thickness [203,204], brain glucose metabolism [205,206], cerebral amyloid- β accumulation [207,208], and the average regional CMRGlc [209] for PET. The hippocampus is of particular interest in the reviewed papers; ROI-based methods have used 3D data and morphological measurements of its cortical thickness, curvature, surface area, and volume. Aderghal, Benois-Pineau and Afdel [189] proposed using both left and right hippocampal regions through flipping regions along the median plane. The relationships between ROIs are also used as standard input; the correlation between regions provides connectivity matrices that are often divided into cortical and subcortical regions [210].

ROI-based methods are closely linked to anatomical regions and have high interpretability and clinical implementability. However, the close link to a priori knowledge limits its potential in explorative studies. The computational cost is usually between that of the voxel and slice-based approaches, but ROI-based methods can maintain local 3D geometric information. Hierarchical neural network frameworks containing sub-networks at each representation level have also been proposed with effective network pruning to retain complete information [192].

4.5. Voxel-Based

Voxel-based approaches are feature-based approaches that focus on the analysis of individual voxels, which are the three-dimensional pixels that make up a medical image [211]. These voxels represent discrete locations in the brain [212], and their size and number can be adjusted to balance computational efficiency and spatial resolution [213]. Compared with slice-based approaches, voxel-based methods can capture the three-dimensional structure of the brain and its changes, which may not be evident in two-dimensional slices. Due to the complexity of brain structure and differences between subjects, spatial co-alignment (registration) is essential [214]. Registration is the process of spatially aligning image scans to an anatomical reference space [215]. This process involves aligning MRI images of different patients or the same patient at different time points to a standardized template representing a common anatomical space [216]. Many studies segment the aligned images into different tissue types, such as gray matter, white matter, and cerebrospinal fluid, using unique signal features of different tissue types before applying the model [217,218]. Comparing gray and white matter across groups or time points can be a sensitive method for detecting subtle changes in brain structure. However, voxel-based approaches also have limitations. One major limitation is the requirement for high spatial resolution. The paper [219] utilizes functional network topologies to depict neurodegeneration in a low-dimensional form. Furthermore, functional network topologies can be expressed using a low-dimensional manifold, and brain state configurations can be represented in a relatively low-dimensional space.

5. Introduction to Deep Learning

Deep learning (DL) is a branch of machine learning which implements universal approximators of neural networks [70,220], a modern development of the original perceptron [221,222] with chain rule-derived gradient computation [223,224] and backpropagation [225,226]. The fundamental formulation of the neural network can be represented through the formulation of a classifier:

$$y = f'(x) \quad (1)$$

where x is the data. The function f' represents the ideal mapping between the input and the underlying solution y . A neural network defines a mapping $f(x, \theta)$ that provides an approximation of f' by adjusting its parameters θ . This adjustment can be considered a form of learning. For learning, a loss function, $L(f(x; \theta), y)$, can be constructed through the relation between the ideal output and the current output of the neural network. Back-propagation through derivatives of the loss function provides a means of updating the parameters for the learning process with a learning rate of ϵ . DL can abstract latent feature representations with minimal manual interference. Features generated by DL cover the hierarchy of low- to high-level features that extend from lines, dots, or edges to objects or characteristic shapes [227].

$$\theta \leftarrow \theta - \epsilon \frac{\partial L(x, \theta)}{\partial \theta}, \quad (2)$$

Advances in deep learning have achieved performance comparable to healthcare professionals in medical imaging classification [175,228]. Due to its feature as a component-wise universal approximator, it can be formulated in multiple ways, including feature extractors dependent on preprocessing and domain knowledge, classifiers for discrimination between groups, or regressors for the prediction of scenarios. Neural networks can also be used in AD knowledge discovery as the feature representation extracted by neural networks might contain information that is counter-intuitive to human understanding. This review outlines the fundamental techniques of deep learning and the main categories of current approaches to various challenges. As a machine learning sub-branch, deep learning approaches can be categorized into two main categories: unsupervised learning and supervised learning.

6. Unsupervised Learning

Unsupervised learning extracts inferences without ground truth categorization of the provided data samples or labels, while supervised approaches require data sample and label pairs. In deep learning, no architecture is strictly supervised or unsupervised if we decompose them into their base components, e.g., feature extraction and classification components of convolutional neural networks. In this survey, the distinction is made based on the relationship between the optimization target of the main neural network or framework and ground truth labels. Unsupervised learning methods will be summarized in this section, while supervised learning methods will be summarized in Section 7.

6.1. Autoencoder (AE)

Autoencoders are a type of artificial neural network designed to learn efficient data representation. The classical application of autoencoders is an unsupervised learning method with two main components: the encoder f_e , and the decoder f_d . The encoder is a neural network designed to map the input to a latent feature representation, while the decoder is a mirror image of the encoder designed to reconstruct the original input from the compressed representation, i.e.,

$$x' = f_d(f_e(x)), \quad (3)$$

where x' is the reconstructed input, and $f_e(x) = z$, is the latent representation. AE can obtain efficient data representations in an alternative dimension by minimizing a reconstruction loss, e.g., squared errors:

$$L(x, x') = \|x - x'\|^2 \quad (4)$$

The original AE consists of fully connected layers, while a stacked autoencoder consist of multiple layers within the encoder and decoder to allow extraction of high-level representations. This structure can be directly applied to train on extracted features such as the ROI features detailed in Section 4.4. In a previous study, structural features of ROI were combined with texture features extracted from Fractal Brownian Motion co-occurrence matrices [229]. Since the AE is unsupervised, a supervised neural network component is attached after training to enable classification or regression. This component commonly consists of fully connected layers (FCL) and activations. Fine-tuning by re-training the network with the supervised component is often applied to achieve better performance.

Greedy layer-wise training can be applied since the encoder and decoder have similar structures regardless of the number of stacked layers. In this training protocol, layers are continuously added to the encoder and decoder and retrained for hierarchical representation. Liu et al. [230] integrated this protocol with multi-modal fusion to improve multi-class classification with MCI sub-types to 66.47% ACC with 86.98% specificity. The same AE also achieved higher performance for binary classification tasks of AD vs. HC and MCI vs. HC. Another commonly applied method to improve AE performance is using sparsity constraints on the parameters. The constraint can be applied through l_1 -regularization or Kullback–Leibler divergence [231–233] for the model to learn with limited neurons during training instances and thereby reduce overfitting. For the classification between HC and MCI, Ju et al. [234] applied sparsity-constrained AE with functional connection matrices between ROIs in fMRI data. The sparsity constraint AE achieved a classification ACC of 86.47% with an AUC of 0.9164, over 20% higher than the machine learning counterparts of SVM, LDA, and LR. Apart from the training protocol and parameter constraint, some methods moderate the input and output of AE. Denoising AE reformulates the original reconstruction problem of AE to a denoising problem with the introduction of isotropic Gaussian noise.

$$x' = f_d(f_e(x + N(0, 1))) \quad (5)$$

Ithapu et al. [235] utilized this AE variant for feature extraction to construct a quantified marker for sample enrichment. Bhatkoti and Paul [236] applied a k-sparse autoencoder where only the neurons corresponding to the k-largest activations in the output are activated for backpropagation. These studies are representative of innovations in the application and enhancement of the original autoencoder.

The structure of neural networks in the encoder and decoder is not limited to MLP; the convolutional structure is also common among AD-related applications of AE. A study has applied 1D convolutional-AE to derive vector representations of longitudinal EHR data, where the 1D convolution operations act as temporal filters to obtain information on patient history [110]. Similarly, more sophisticated convolutional structures can also be used in the encoder and decoder architecture. Oh et al. [237] applied a convolutional AE with Inception modules, which are groups of layers consisting of multiple parallel filters. The standardized structure of AE makes it adaptable to any input dimension by configuring the encoder and decoder structure. Hosseini-Asl et al. [238], and Oh, Chung, Kim, Kim and Oh [237] applied 3D convolutional autoencoders to compress the representations of 3D MRI, while Er and Goularas [239] applied AE as an unsupervised component of the feature extraction process. AE can also be implemented as a pre-training technique, where after training, fully connected layers are added to the compressed layer of the encoder and used for supervised learning [240].

Apart from structural adjustments to the encoder and decoder layers, a probabilistic variation of AE also exists. These AE are known as variational autoencoders (VAE). For VAE, a single sample of available data x_i can be interpreted as a random sample from the true distribution of data p' , while the encoder can be represented as $q(z|x)$, an approximation to the true marginal distribution of $p(x|z)$. The loss function is, therefore,

$$L = L_1(x, x') + L_{KL}(q(z|x), p(z)), \quad (6)$$

where L_1 is the reconstruction loss and L_2 is the Kullback–Leibler divergence, which regularizes the VAE and enforces the Gaussian prior $p(z) = N(0, 1)$. Through this adjustment, AE learns latent variable distributions instead of representations [241]. A more intuitive formulation is as follows:

$$\mu = f_{h_1}(f_e(x)) \text{ and } \sigma = f_{h_2}(f_e(x)), \quad (7)$$

where f_{h_1} and f_{h_2} represent the mapping to two independent neural network layers representing μ and σ , the set of mean and variance of the latent distributions. The latent representation can be sampled through reparameterization,

$$z = \mu + \sigma \varepsilon \text{ where } \varepsilon \sim N(0, 1), \quad (8)$$

and decoded to reconstruct the input $x' = f_d(z)$. Variational autoencoder has recently been applied to extract latent distributions of eMCI from high-dimensional brain functional networks [242] and provide risk analysis for AD progression [243]. Instead of a single set of latent distributions, a hierarchy of latent distributions can also be learned using ladder VAE. This variant of VAE was applied by Biffi et al. [244] to model HC and AD hippocampal segmentation populations, where latent distribution-generated segmentations for AD showed apparent atrophy compared with HC. By learning latent distributions, new data can be sampled from these distributions to generate new samples. From this perspective, VAE can be considered a generative model and is introduced in the following subsection. The fundamental autoencoder structures are shown in Figure 5.

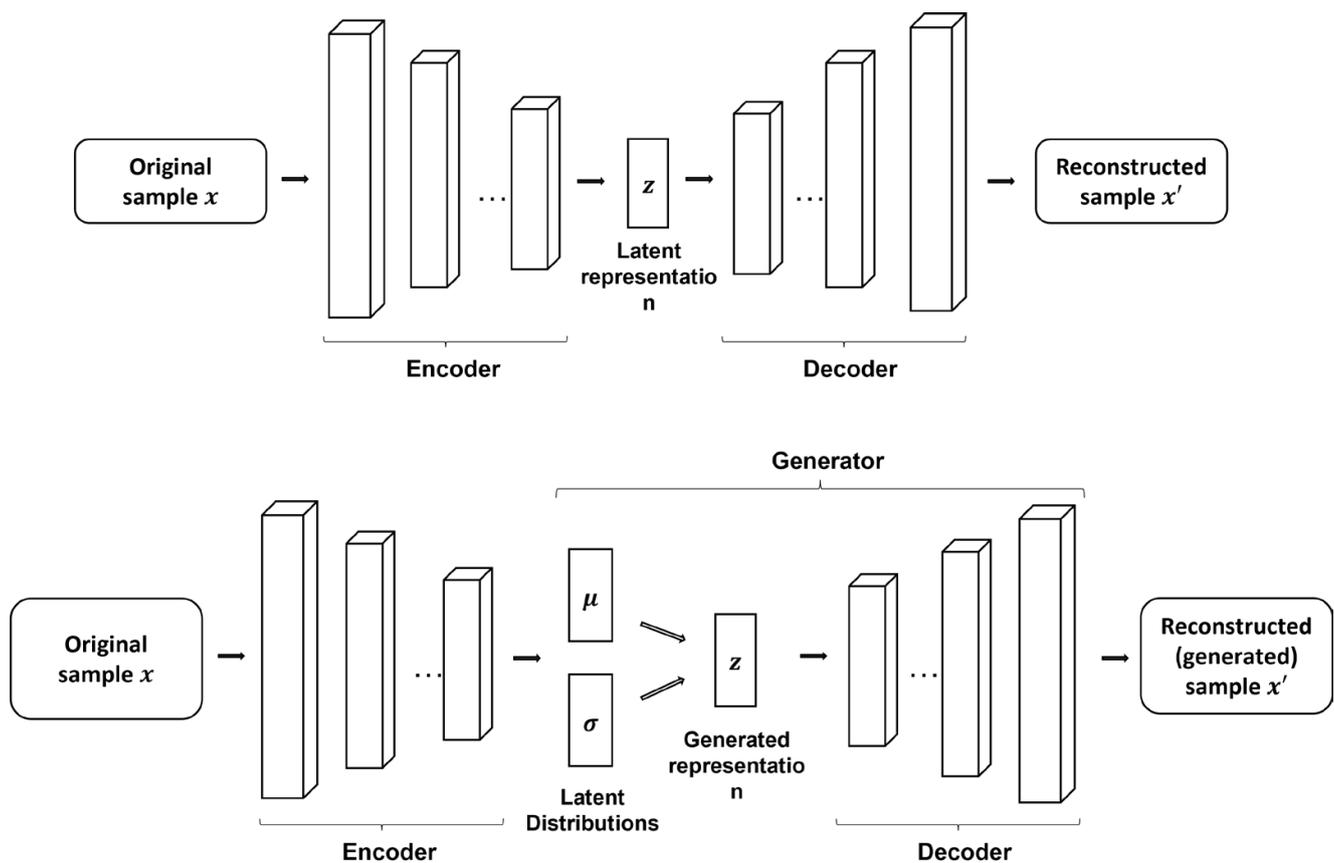


Figure 5. Fundamental autoencoder structures. The top figure represents a stacked 2D autoencoder, where each block represents a convolutional layer composed of a bank of convolutional filters (represented by the rectangular columns). The bottom represents a VAE, where instead of the latent representation, the encoder generates latent distributions represented by mean μ and variance σ , which are then used to generate representations. The convolution operation can be replaced by fully connected layers or complex modules, e.g., the Inception module.

6.2. Generative Models

Generative methods are a form of unsupervised learning that requires the model to recreate new data to supplement an existing data distribution. Variational autoencoders and RBM mentioned in the previous sections are both generative models. Another popular generative method is the construction of generative adversarial networks (GANs), where two or more neural networks compete in a zero-sum game. Classical GAN includes a generative neural network G used to generate dummy data and a discriminator neural network D to determine whether a sample is generated. The generator generates fake images $x' = G(\epsilon)$ from noise ϵ . The generated sample belongs to the generated data distribution $x' \in p_g$. The discriminator attempts to discriminate between generated images x' and real images, $x \in p_r(x)$. The competition between the generator and discriminator can be formulated through their loss function

$$L = E_{G(\epsilon) \sim p_g} \{\log[1 - D(G(\epsilon))]\} + E_{x \sim p_r} \{\log[D(x)]\}, \quad (9)$$

where the objective is to minimize G and maximize D [245]. GAN is widely used for medical image synthesis, reconstruction, segmentation, and classification [246]. Islam and Zhang [247] applied a convolutional GAN to generate synthetic PET images for AD, NC, and MCI. The GAN model generated images with a mean PSNR of 32.83 and a mean SSIM of 77.48. The generated data were then classified using a 2D CNN, which achieved 71.45% ACC. This performance drop illustrates the difficulty in synthesizing quality syn-

thetic images for training. A similar framework was proposed with shared feature maps between the generator and discriminator. With transfer learning, the framework achieved 0.713 AUC for SCD-conversion prediction [248]. Roychowdhury and Roychowdhury [249] implemented a conditional GAN, where the discriminator and generator are conditioned by labels y ,

$$L = E_{G(\epsilon) \sim p_g} \{\log[1 - D(G(\epsilon|y))]\} + E_{x \sim p_r} \{\log[D(x|y)]\}, \quad (10)$$

The conditional GAN was applied to generate longitudinal MRI data by generating and overlaying cortical ribbon images. The generated data provide a potential disease progression model of MCI to AD conversion and brain atrophy. The study showed that the modeled fractal dimension of the cortical image decreases over time. Baumgartner et al. [250] applied an unsupervised Wasserstein GAN, where a K-Lipschitz constraint critic function C replaces the supervised discriminator. The loss of this model can be formulated as:

$$L = E_{x \sim p(y=1)} \{\log[C(x + M(x))]\} + E_{x \sim p(y=0)} \{\log[C(x)]\}, \quad (11)$$

where D is a set of 1-Lipschitz functions, and M is a map generator function that uses existing images x to generate new images $x' = x + M(x)$. An additional regularization component $L_M = \|M(x)\|_1$ is also added to the overall loss function to constrain the map M for minimum change to the original image x . In the study, M is modeled by a 3D U-Net segmentation model. The modified WGAN generated disease effect maps similar to human observations for MRI images of MCI-converted AD. An alternative application of Wasserstein GAN with additional boundary equilibrium constraints was applied by Kim et al. [251]. This study extracted latent representations from autoencoder structure discriminators for classification with FCL and SVM. For AD vs. HC, the model achieved an ACC of 95.14% with an AUC of 0.98. A subsequent study by Rachmadi et al. [252] built upon the Wasserstein GAN structure with an additional critic function C_2 . The loss function corresponding to this additional component is:

$$L_{C_2} = E_{x_1, x_0 \sim p_1, p_0} [C_2(x_1 - x_0)] - E_{x_0 \sim p_0} [C_2(M(x_0))], \quad (12)$$

where x_0 and x_1 are baseline and follow-up images, respectively. Apart from using the original critic C to discriminate between real and fake images, the new critic C_2 is also applied to discriminate between real disease evolution maps $x_1 - x_0$ and generated maps $M(x_0)$. The inclusion of C_2 reformulates the generation of dummy scans to the generation of longitudinal evolution maps. Though this study was applied in monitoring the evolution of white matter hyperintensities in cerebral small vessel disease, the same concept and technique can be migrated to data on Alzheimer's and related diseases [252]. Example GANs are illustrated in Figure 6. Apart from GAN, another type of innovative generative model is invertible neural networks (INN), which create invertible mappings with exact likelihood. Sun et al. [253] used two INN to extract the latent space of MRI and PET data and map them to each other for modality conversion. Conditional INNs, based on the conditional probability of latent space and combined with recurrent neural networks (RNN), were also used to generate longitudinal AD samples [176].

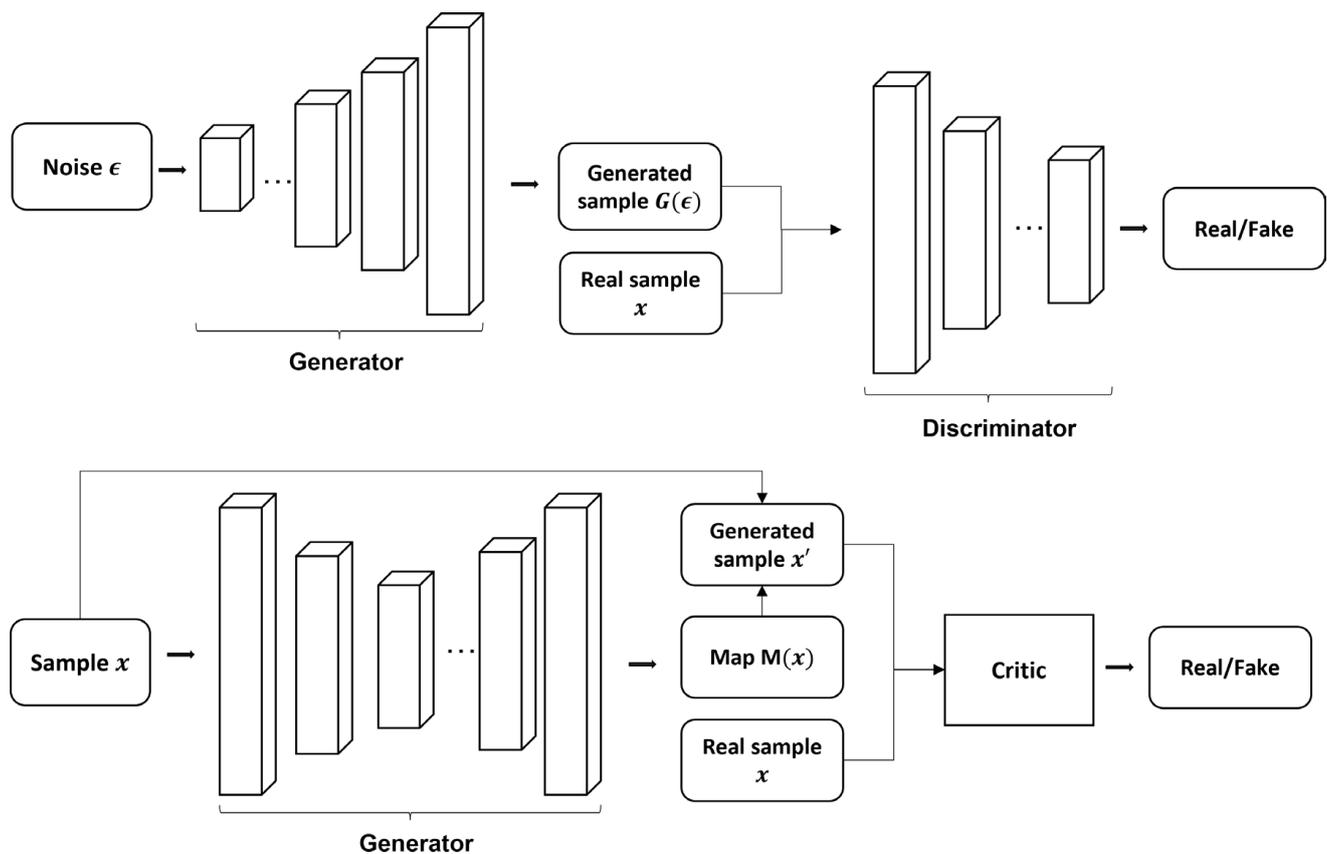


Figure 6. Example generative adversarial networks. The top figure is an example vanilla 3D convolutional generative adversarial network. The bottom figure shows the basic schematics of the modified Wasserstein GAN [250,252]. The structure of each generator and discriminator component can be modified for different neural network architectures.

6.3. Restricted Boltzmann Machine (RBM) and Other Unsupervised Methods

Apart from GAN and AE, numerous unsupervised methods have been applied for AD and related diseases. A well-known category is the restricted Boltzmann machine (RBM). An RBM is a generative network with a bipartite graph used to extract the probability distributions of the input data. RBM consists of two symmetrically-linked layers containing the visible and hidden units, respectively. The units, or neurons, within each layer are not connected. Similar to autoencoders, RBMs encode the input data through the forward pass while reconstructing input data through its backward pass. Two sets of biases for the two different passes aid this process. As an unsupervised method, RBM can also be used for feature extraction. Li et al. [254] applied multiple RBMs to initialize multiple hidden layers one at a time, while Suk et al. [255] combined RBM with the autoencoder learning module by combining layer-wise learning with greedy optimization. Conditional RBM has been applied as a statistical model for unsupervised progression forecasting of MCI, achieving ADAS-Cog13 prediction performance compared with supervised methods [256]. A deep belief network (DBN) is a neural network architecture comprising stacked RBMs. The basic structure of a DBN is shown in Figure 7. A DBN allows a backward pass of generative weights from the extracted feature to the input, making it more robust to noise. However, the layer-by-layer learning procedure for DBN can be computationally expensive. Suk, Lee, Shen and Initiative [166] applied a combination of MLP and DBM for feature extraction from multiple modalities.

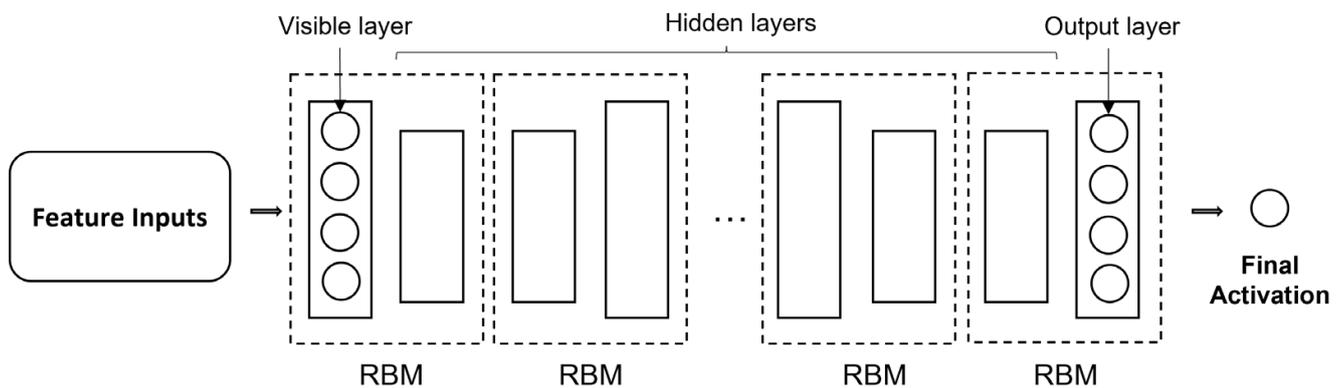


Figure 7. The basic structure of a deep belief network (DBN) consists of multiple restricted Boltzmann machines (RBM).

More recent studies in unsupervised learning applied show great diversity. Razavi et al. [257] applied sparse filtering as an unsupervised pre-training strategy for a 2D CNN. Sparse filtering is an easily applicable pre-training method where a neural network is first trained to output in a specified feature dimension. In this study, the cost function of classification is replaced by minimizing the sparsity of l_2 -normalized features of specified dimensions. Bi et al. [258] combined a CNN with PCA-generated filters and k-means clustering for a fully unsupervised framework for clustering MRI of AD, MCI, and NC. Wang, Xin, Wang, Gu, Zhao and Qian [184] hierarchically applied extreme learning machines for unsupervised feature representation extraction. Extreme learning machines are a variant of feedforward neural networks that applies the Moore–Penrose generalized inverse instead of gradient-based backpropagation. Majumdar and Singhal [259] applied deep dictionary input while using noisy inputs, such as denoising autoencoders, for categorical classification, while Cheng et al. [260] utilized a U-net-based CNN with rigid alignment for cortical surface registration of MRI images.

7. Supervised and Semi-Supervised Learning

Supervised learning involves the use of known labels. In this study, we focus on the use of neural networks to map inputs to definite outputs. This section first introduces architecture classes, such as convolutional and recurrent neural networks, in Sections 7.1 and 7.2. We then present recent advances in transfer learning, ensemble learning, and multimodal fusion in Sections 8.1–8.3. Finally, we introduce the most recent developments in graph and geometric neural networks.

7.1. Convolutional Neural Networks (CNN)

The innovation of convolutional neural networks (CNN), especially the development of the AlexNet [261,262] by Krizhevsky et al. [263], validated neural networks as practical universal approximators with layer-wise feature propagation. In CNN, the dense connections of MLPs are replaced with kernel convolutions:

$$f(x)_{m,n} = \sigma \left(\sum_i^H \sum_j^W \sum_c^C x_{i,j,c} K_{m+i-1, n+j-1,c} \right), \quad (13)$$

where K is the convolutional kernel; σ is a non-linear activation function; and H , W , and C represent the dimensions for height, width, and channel of the input. CNN allows for parameter-efficient hierarchical feature extraction. Besides the reduced computational requirements, CNN has translational invariance and can retain spatial information, making it particularly suitable for neuroimaging data. The effectiveness of CNN is evident in their broad application, both as an independent model and as network components [264].

A typical CNN consists of several convolutional layers followed by non-linear activations. The non-linearity provides the basis for learning through backpropagation. Commonly used activation functions include the rectified linear unit (ReLU), $\sigma = \max(0, x)$, the hyperbolic tangent, $\sigma = \tanh x$, and sigmoid functions, $\sigma = (1 + e^{-x})^{-1}$. Recent new activation functions such as leaky-ReLU and parametric-ReLU are also seen in the reviewed literature [177,265].

Pooling, the downsampling of feature maps through an average or maximum filter approach, is also often applied. Batch normalization, where each mini-batch of data is standardized, is also commonly applied after convolution. A combination of the above procedures forms a convolution block, and a typical CNN comprises multiple convolution blocks. These blocks are often followed by a few fully connected layers and a Softmax activation for classification or a linear activation for regression. The theoretical foundations of CNN can be understood through the decomposition of tensors [266], while in this paper, we will focus on practical applications of CNN for AD-related tasks. The following subsections will provide a summarized introduction to 2D and 3D CNN focusing on recent applications, while more detail can be found in previous reviews [89,96].

7.1.1. 2D-CNN

The original CNN was designed for computer vision pattern recognition of 2D images, allowing an easy application for 2D neuroimaging data. A basic 2D-CNN is shown in Figure 8. Aderghal, Benois-Pineau and Afdel [189] used a two-layer CNN with ReLU and max-pooling of 2D+ ϵ images that project slices from the sagittal, coronal, and axial slices into a three-channel 2D image. Alternatively, when 2D slices are available from multiple planes of a 3D image, an individual 2D-CNN can be used for each image and then ensembled. Neural network depth is associated with an increase in performance. Wang, Phillips, Sui, Liu, Yang and Cheng [265] proposed a deeper eight-layer CNN with leaky rectified learn units to classify single-slice MRI images, while a similar CNN was applied for the classification of Florbetaben-18 PET images [177]. Tang et al. [267] used a CNN model to identify amyloid plaques in AD histology slides. Similar to the aforementioned 2D CNN, the neural network consists of alternating layers of 2D convolution and max-pooling, followed by fully connected layers with ReLU activation and a Softmax activation to produce classification outputs. The CNN model showed excellent performance in the classification of amyloid plaques with an AUC of 0.993. The current state-of-art 2D CNN models are also mostly developed for natural image classification, though these models are easily applicable for 2D AD-related data. The availability of pre-trained state-of-art models provides the basis of transfer learning, as summarized in Section 8.1.

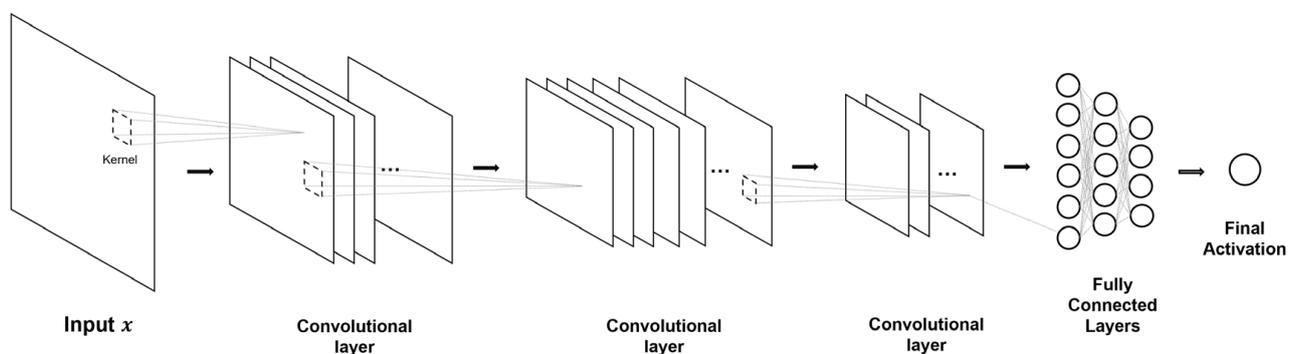


Figure 8. Example 2D-CNN. This architecture provides the foundation for 2D convolutional architectures. Square slices in this figure represent channel-wise feature maps after convolution.

Due to the two dimension limit, data with multiple slices are either treated as independent or similar. 2D-CNN can also be applied to 1D data, including using the Hilbert space-filling curve to transform 1D cognitive assessment data to 2D [268] or using the time-series data of multi-channel EEG as a 2D matrix [269]. Though limited by dimensionality, 2D-CNN can be more practical in real-world application and deployment, as the data used in clinical practice are often 2D or lacks enough slices to construct the high-dimensional 3D T1-weighted MRI predominately used for medical research. The 3D neuroimaging data in open libraries such as ADNI and OASIS are often processed to obtain 2D slices or patches, as mentioned in Section 4. To retain 3D spatial information, 2D slices or patches from the sagittal, coronal, and axial views are often extracted for multi-view networks [270]. The lower dimensionality of 2D-CNN also makes it suitable for adaptation for 1D data, e.g., Alavi et al. [271] utilized the triplet architecture of face recognition and the Siamese one-shot learning model for automated live comparative analysis of RNA-seq data from GEO.

7.1.2. 3D-CNN

3D-CNN is inherently the same as 2D-CNN apart from an additional dimensionality in all components, including the convolutional kernel. The additional dimension provides 3D-CNN with better spatial information than 2D-CNN as the latter is inherently limited by kernel dimensionality and is, therefore, unable to efficiently capture the spatial information between slices. A basic 3D-CNN is shown in Figure 9. Similar to fundamental 2D-CNN models, Islam and Zhang [272] used a 3D-CNN composed of four 3D convolutional layers with FCL and Softmax with T1-weighted MRI, while Duc et al. [273] applied a similar CNN with rs-fMRI functional networks. A simple two-block 3D-CNN applied by Basaia et al. [274] showed either comparable or better performance than 2D-CNN in binary classification with AD, NC, and various MCI subtypes.

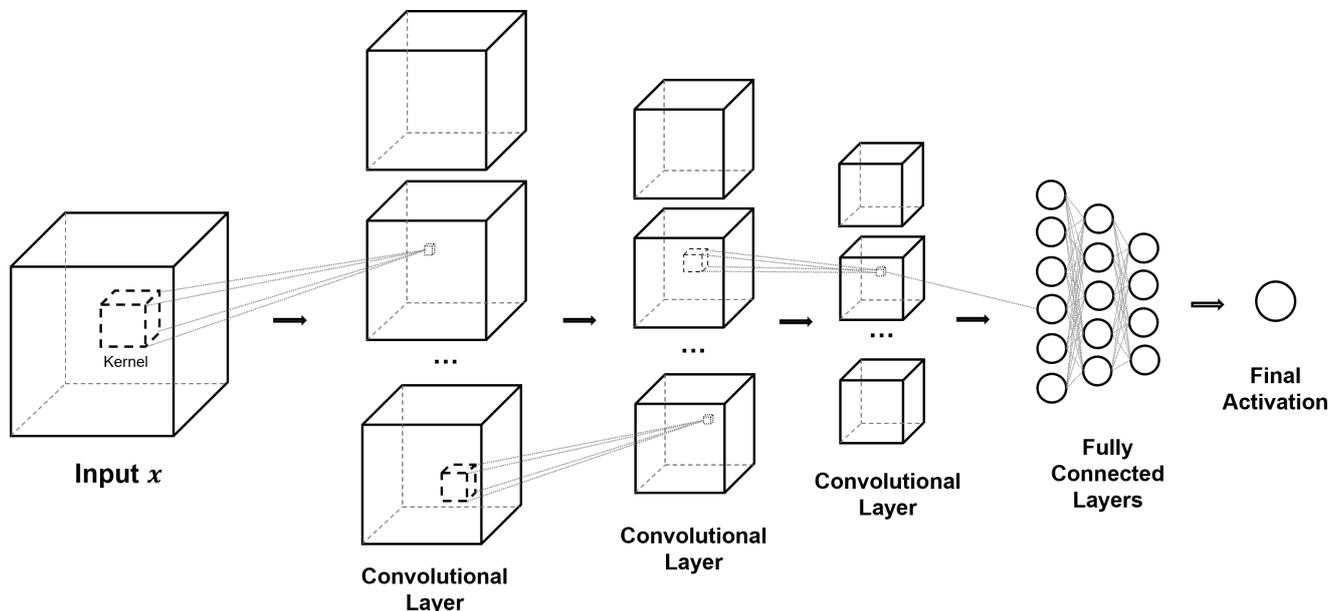


Figure 9. Basic 3D CNN architecture. 3D images and patches from Figure 4 can be used as inputs for this architecture. Individual blocks within a convolutional layer represent channel-wise feature maps after convolution. Modifications such as identity mapping and dense connectivity can be applied with an additional dimension of height. Fully connected layers can be replaced with global average pooling for a fully convolutional neural network, while the final activation can be modified for classification, regression, or additional structure can be applied for alternative tasks such as semantic segmentation.

With the similarity between 2D and 3D CNN, high-performing architectures in two dimensions can easily be adapted to three dimensions; Basaia, Agosta, Wagner, Canu, Magnani, Santangelo, Filippi and Initiative [274] and Qiu et al. [275] both implemented 3D versions of an all convolutional neural network for the classification of AD and MCI, where the FCL + Softmax classification component is replaced with a CNN with a channel number corresponding to the number of categories and global pooling of each channel. A similar application of an all convolutional CNN was applied by Choi et al. [276] for MCI conversion prediction. Unsupervised pre-training has also been tested by Hosseini-Asl, Keynton and El-Baz [238] and Martinez-Murcia et al. [277] with 3D convolutional autoencoders, while features extracted by 3D-CNN have also been used as input for sparse autoencoders [278]. Ge et al. [279] combined a U-Net-structured 3D-CNN for multi-scale feature extraction with XG-Boost feature selection. State-of-art architectures for 2D-CNN have also been adapted to 3D, e.g., a 3D architecture for the Inception-v4 network [280]. Liu et al. [281] used 3D-AlexNet and 3D-ResNet as comparative models. Wang et al. [282] also proposed a probability-based ensemble of densely connected neural networks with 3D kernels to maximize network information flow. This study also revealed ensemble learning as a potential approach to higher performance, which is detailed in Section 8.2.

The additional dimensionality of 3D does not restrict input to the spatial domain. An example is dynamic functional connectivity networks, which are 2D representations of brain ROIs' changes in blood oxygen level-dependent (BOLD) signals over time. For input of FCNs, the 3D-CNN obtains an additional temporal dimension in addition to the 2D spatial representation. With convolution along the temporal dimension, the neural network combines temporal and spatial connectivity to form more dynamic FCNs that can characterize time-dependent interactions considering the different contributions of time points [186].

The additional dimensionality of 3D-CNN corresponds to a significantly higher number of parameters within the model and higher computational cost. In order to reduce the computational cost, Spasov et al. [283] applied parameter-efficient 3D separable convolution, where the original 3D convolution is divided into depth-wise convolution and 1×1 point-wise convolution. Liu, Yadav, Fernandez-Granda and Razavian [281] performed comparative and ablation experiments and found that instance normalization can generalize better than batch normalization. This study also found that early spatial down-sampling negatively impacts model performance, indicating that wider CNN architecture is more beneficial than additional layers and that smaller initial kernel sizes are ideal.

Liu, Cheng, Wang, Wang and Initiative [170] proposed the combined use of an ensemble 3D-CNN and 2D-CNN in a sequential manner, where the 3D-CNN captures spatial correlations with the 3D input. An ensemble of cascading 3D-CNN-generated feature maps is used as input for 2D-CNNs. While most of the studies above focus on categorical classification, disease progression predictions, and prediction of clinical measures, some deep learning studies were applied for different purposes, e.g., segmentation and image processing, which is potentially valuable for future studies in Alzheimer's and related diseases. Yang et al. [284] proposed a 3D-CNN with residual learning architecture for hippocampal segmentation that is significantly more efficient than conventional algorithms. Pang et al. [285] combined a semi-supervised autoencoder with local linear mapping. With the development and availability of more powerful hardware in the past decade, the 3D convolutional neural network has become increasingly popular amongst applications within the reviewed literature.

7.2. Recurrent Neural Networks (RNN)

Longitudinal data of AD provide multiple data instances of a subject, allowing us to find ground truth for MCI conversion and time-to-conversion. However, the temporal nature of a series of instances is often not explored in DNN and CNN architecture. Recurrent neural networks incorporate the temporal domain through adaption to a sequence of input with time-varying activation and sequential synapse-like structure. The fundamental concept was formulated by Goodfellow et al. [227] as:

$$h^t = f(h^{t-1}, x^t), \quad (14)$$

where h^t and x^t represent the state and input at time step t . The state can be unfolded with respect to the past sequence:

$$h^t = g^t(x^1, x^2, \dots, x^t), \quad (15)$$

where g^t is a function. This property of the vanilla RNN allows f to learn on all time steps and sequence lengths. Second-order RNNs consist of more complex neurons with memory components such as long short-term memory. LSTM is composed of a memory cell and three gates. The gates can be formulated as:

$$g_c^t = \sigma(b_c + U_c x^t + W_c h^{t-1}), \quad (16)$$

where for each gate c , σ represents the activation, and W_c and U_c represent the recurrent weight and input weight matrices, respectively. The cell and update protocol can be formulated as:

$$s^t = g_f^t s^{t-1} + g_e^t \sigma(b + U x^t + W h^{t-1}), \quad (17)$$

$$h^t = \tanh(s^t) \cdot g_o^t, \quad (18)$$

where g_f is the forget gate, g_e is the external input gate and g_o is the output gate. The recurrent weight and input weight of the memory cell are represented as U and W , respectively.

LSTM has been applied to brain network graph matrices to extract adjacent positional features from fMRI data; the combination of LSTM and extreme learning machine (ELM) showed a slight improvement over a CNN-ELM model in classification tasks [185]. Gated recurrent unit (GRU) is another gated-RNN structure that shares a similar structure with LSTM but does not contain the forget gate. Therefore, it contains a lower number of parameters and is more suitable for capturing long-term temporal patterns. GRU has been used for classification with temporal clustering of actigraphy time-series obtained through the monitoring of activity for NC, MCI, and AD subjects. In this application, features extracted with CNN and Toeplitz inverse covariance-based clustering were combined and fed into the recurrent neural network [286].

Bi-directional GRU (BGRU) is a GRU variation that can process input both forwards and backwards. It has been applied in a similar manner to MLP and CNN-extracted features in multiple studies [287,288]. Apart from its use as a classification component to replace traditional MLP or machine learning classifiers, RNN can also be utilized for structural data. One study has combined CNN and RNN by inputting a series of 2D slices from 3D scans to capture spatial features; the CNN component captures features within single slices, while the BGRU structures obtain a time-series of CNN-extracted features to extract inter-slice features, which are then used as input for an MLP classifier component [289]. Similarly, instead of features extracted from slice-level data, LSTM architecture variants have been modified to suit 3D structural data, e.g., 3D convolutional LSTM to encode representations extracted by a 3D-CNN [290].

In the range of applications of RNN in deep learning, a key characteristic that stands out is its ability to deal with temporal data. Therefore, one focus of interest is the combination of spatial and temporal information. Wang et al. [291] combined the two types of information for fMRI data through the parallel implementation of multiple LSTM on features corresponding to multiple time series of ROI BOLD signals and the use of convolutional components for time-series segments. While most studies formulate the MCI prediction problem as a classification task [292], one study has used features extracted from an LSTM-based autoencoder for prognosis modeling with a Cox regression model [293]. For time-series or sequential data, sample completeness is a significant challenge in practice. Due to the difficulty in longitudinal data collection, many datasets have missing or delayed collection time points. On top of classical data imputation, Nguyen et al. [294] utilized their proposed minimal RNN model to impute missing data by filling it with model predictions. This study achieved exceptional results in the TADPOLE longitudinal challenge for the 6-year predictions of ADAS-Cog13 and ventricular volume. These studies have shown the effectiveness of RNNs in the temporal modeling of AD and related diseases. With the incrementally increasing amount of longitudinal data collected across various projects, they will significantly impact the direction of the deep learning approach.

7.3. Graph and Geometric Neural Networks (GNNs)

The underlying assumption for conventional neural networks is that the latent distribution of data lies in the Euclidean domain. Graph and geometric neural networks are a branch of deep learning designed for data in the non-Euclidean domain [295], e.g., genetic pathways, brain manifolds, and functional networks. Before the development of GNN, the dominant deep learning approach to graph data was graph kernel methods, where a kernel function is used to map the graph into vector space as input for neural networks. Studies have utilized this method by calculating brain function networks represented as correlation matrices and then using them as input for various neural networks [185,296]. Therefore, most brain function network studies can be considered graph kernel methods. The pipeline to generate these vectors is deterministic, while GNN is learnable and is relatively less penalized by the curse of dimensionality with relational data [297].

There are many categories of GNN, but the most popular GNN in AD research are graph convolutional neural networks (GCNN), which can be sub-categorized into special and spatial GCNN. Song et al. [298] utilized GCNN for multi-class classification and verified a performance advantage over traditional machine learning classifiers with a low sample size. The GCNN was then applied to predict tau protein trajectory with a constraint on loss based on a physical model of tau protein spread in the brain [299]. Song et al. [300] proposed a GCN framework with similarity-aware receptive fields and adaptive adjacency matrices generated through pre-training for better prediction.

A major sub-category of GCNN is spatial GCNN, where we reformulate convolution operations onto the graph nodes to exploit their spatial relationships [301]. A simple formulation of this process is presented by Wu, Pan, Chen, Long, Zhang and Philip [297], where for each layer k ,

$$h^k = \sigma \left(XW^k + \sum_{i=0}^{k-1} Ah^{k-1} \Theta^k \right), \quad (19)$$

where A is an adjacency matrix that contains the connection information between graph nodes; X is the feature matrix of the graph; and W and Θ are matrices of learnable parameters. A key aspect of utilizing GNNs is the generation of graphs or manifolds, e.g., structural connectivity graphs derived from DTI [298] and hypersphere projections of brain-functional networks extracted from fMRI [302]. The graph generation component can also be incorporated into the neural network with embedding and attention-based mechanisms [303], while global attention mechanisms can also be used to build resilience against noise and variance [304].

Spectral GCNN redefines convolution operations to the Fourier domain through the eigendecomposition of graph Laplacian [305]. For a simplified channel-wise example, the spectral GCNN can be formulated as:

$$h' = \sigma(\mathbf{U}\omega\mathbf{U}^T h), \quad (20)$$

where ω denotes the channel component of the filter, which contains trainable parameters, and \mathbf{U} are the eigenvalues of a normalized graph Laplacian. $\mathbf{L} = \mathbf{U}\Lambda\mathbf{U}^T$ and $\mathbf{U}^T x$ are equivalent to the Fourier transform of x [297]. Wee et al. [306] generated graphs based on the cortical thickness of structural MRI and implemented a spectral GCNN for classification between disease stages. The model achieved 92% accuracy in predicting late MCI conversion to AD. Similarly, Zhao et al. [307] utilized a Cheby-GCN-based spectral GCNN with graphs constructed upon MCI functional connectivity networks, hardware, and gender information to predict MCI. Similar to the application of the attention mechanism with spatial-GCNN, Kazi et al. [308] combined spectral-GCNN with an attention module based on LSTM for personalized diagnosis. Huang and Chung [309] implemented Monte-Carlo dropout on a similar network structure for uncertainty estimation in the prediction of MCI conversion. Yu et al. [310] proposed a spectral-GCN framework that simulates random walks with parallel GCN layers and takes a combined input of structural connectivity from DTI and functional connectivity from fMRI. The model showed the difference in the structural connection between different disease stages and achieved 84~93% accuracy for binary classification tasks between NC, early MCI, and late MCI [310].

As an emerging field in deep learning, many AD-related studies focus on various other fields of interest in geometric neural networks, including geometric deep learning manifolds, e.g., Zhen et al. [311] implemented a dilated convolutional architecture designed for sequential manifold-valued data and the application of spectral-temporal neural networks for EEG and fMRI data to capture both spatial and temporal information [312,313]. Geometric and graph neural networks represent a more general structure than the rigid Euclidean domain of conventional neural networks. This property is more suitable for inherently non-Euclidean data and can facilitate better integration of a variety of data types. GNN is becoming a significant area of research for developing future neural networks in AD research.

7.4. Other Methods

Other methods include reinforcement learning, a topic of artificial intelligence research that branches apart from supervised or unsupervised learning. Instead of learning representations, reinforcement learning models focus on agents' actions within an environment. Tang, Uchendu, Wang, Dodge and Zhou [112] applied reinforcement learning with natural language processing techniques for an MCI screening dialogue agent. The reinforcement learning environment was set up with the Actor-Critic method, where a user simulator neural network generates new dialogue data. This set-up is very similar to GAN, but for GAN, the actor cannot affect the reward of the critic function [314]. While the perceptron units of neural networks simulate human brain neurons' fundamental function, it is an oversimplistic representation. Current research in deep learning has attempted to create neural networks based on more representative biological neurons. An example of this research field is spiking neural networks (SNN). Compared with the sequential nature of RNNs, SNNs are neural networks inherently temporal by design. Capecci et al. [315] provided a proof-of-concept with an SNN architecture using EEG data for the prediction of MCI conversion.

We summarize the literature mentioned in this section in Tables 3–5.

Table 3. Binary classification results of selected literature between AD, NC, and MCI.

Study	Data Modalities	Number of Subjects			Classification ACC (%)		Classification AUC	
		AD	NC	MCI	AD vs. NC	MCI vs. NC	AD vs. NC	MCI vs. NC
Suk and Shen [316]	MRI, PET	51	52	99	95.9	85	-	-
Suk, Lee, Shen and Initiative [166]	MRI, PET	93	101	204	95.35	85.67	-	-
Liu, Liu, Cai, Che, Pujol, Kikinis, Feng and Fulham [230]	MRI, PET	85	109	77	82.59	82.10	-	-
Li, Tran, Thung, Ji, Shen and Li [254]	MRI, PET, CSF	51	99	52	91.4	77.4	-	-
Aderghal, Benois-Pineau and Afdel [189]	MRI	188	228	399	69.53	91.41	-	-
Suk et al. [317]	MRI	186	393	226	91.02	-	0.927	-
Majumdar and Singhal [259]	MRI, PET, CSF	51	99	52	95.4	85.7	-	-
Cui, Liu and Li [287]	MRI	198	229	-	89.69	-	0.9214	-
Shi et al. [318]	MRI, PET	51	52	99	97.13	87.24	0.972	0.901
Liu, Wang, Tang, Hu, Wu and Pan [210]	MRI	-	303	83	-	90.9	-	-
Lu et al. [319]	PET	226	304	521	93.58	-	-	-
Ning et al. [320]	MRI, Genetic	138	225	358	-	-	0.992	-
Liu, Cheng, Wang, Wang and Initiative [170]	MRI, PET	93	100	204	93.26	74.34	0.957	0.802
Ge, Qu, Gu and Jakola [279]	MRI	193	139	-	93.53	-	-	-
Ju, Hu and Li [234]	fMRI	-	79	91	-	86.47	-	0.916
Liu, Zhang, Adeli and Shen [191]	MRI	227	249	390	93.7	-	-	-
Islam and Zhang [272]	PET	169	400	661	88.76	-	-	-
Wen, Thibeau-Sutre, Diaz-Melo, Samper-González, Routier, Bottani, Dormont, Durrleman, Burgos and Colliot [89]	MRI	336	330	787	87 ^b	-	-	-
Liu, Li, Yan, Wang, Ma, Shen, Xu and Initiative [169]	MRI	97	119	233	88.9	76.2	0.925	0.775
Lee et al. [321]	MRI	198	229	374	92.75	89.22	0.980	0.957
Lian, Liu, Zhang and Shen [192]	MRI	358	205	2964	89.5	-	0.959	-
Cui and Liu [187]	MRI	192	223	396	92.29	74.64	0.75	0.797
Martinez-Murcia, Ortiz, Gorriz, Ramirez and Castillo-Barnes [277]	MRI	99	168	212	84.9	-	-	-
Duc, Ryu, Qureshi, Choi, Lee and Lee [273]	fMRI	133	198	-	85.3 ^b	-	-	-

Table 3. Cont.

Study	Data Modalities	Number of Subjects			Classification ACC (%)		Classification AUC	
		AD	NC	MCI	AD vs. NC	MCI vs. NC	AD vs. NC	MCI vs. NC
Kim, Lee, Lee, Oh, Yun and Yoo [251]	PET	212	415	-	94.82	-	0.98	-
Choi, Kim, Yoon, Lee, Lee and Initiative [276]	PET	243	393	666			0.94	-
Xia, Yue, Xu, Feng, Yang, Wang and Lei [290]	MRI	198	299	408	94.19	79.01	0.96	0.88
Ieracitano, Mammone, Hussain and Morabito [269]	EEG	63	63	63	85.78	85.34	-	-
Islam and Zhang [247]	PET	98	105	208	71.45	-	-	-
Qiu, Joshi, Miller, Xue, Zhou, Karjadi, Chang, Joshi, Dwyer and Zhu [275]	MRI, Demo, CA	488	978	-	96.8	-	0.996	-
Bashyam et al. [322]	MRI	353	833	513	86	70.2	0.91	0.743
Pan, Phan, Adel, Fossati, Gaidon, Wojak and Guedj [270]	PET	237	242	526	93.13	-	0.9747	-

^b Some studies applied balanced accuracy, where accuracy is weighted by categorical distribution.

Table 4. Results from selected studies of binary classification between cMCI and ncMCI.

Study	Data Modalities	Time to Conversion	Number of Subjects		ACC (%)	AUC
			cMCI	ncMCI		
Suk, Lee, Shen and Initiative [166]	MRI, PET		78	128	75.92	
Suk, Lee, Shen and Initiative [317]	MRI	18 M	167	226	74.82	0.754
Ning, Chen, Sun, Hobel, Zhao, Matloff, Toga and Initiative [320]	MRI, Genetic	24 M	166	192		0.835
Lu, Popuri, Ding, Balachandar, Beg and Initiative [319]	PET	36 M	112	409	82.51	
Cui and Liu [187]	MRI		165	231	74.64	0.777
Spasov, Passamonti, Duggento, Liò, Toschi and Initiative [283]	MRI, Demo, CA, Genetic	36 M	181	228	86	0.925
Lee, Choi, Kim, Suk and Initiative [321]	MRI	18 M	160	214	88.52	
Choi, Kim, Yoon, Lee, Lee and Initiative [276]	PET	36 M	167	274		0.82
Lian, Liu, Zhang and Shen [192]	MRI	36 M	205	465	80.9	0.781
Wen, Thibeau-Sutre, Diaz-Melo, Samper-González, Routier, Bottani, Dormont, Durrleman, Burgos and Colliot [89]	MRI	36 M	295	298	76	
Er and Goularas [239]	MRI		125	169	87.2	
Pan, Phan, Adel, Fossati, Gaidon, Wojak and Guedj [270]	PET	36 M	166	360	83.05	0.868

Abbreviations: M—months, cMCI—MCI converters, ncMCI—non-converters, w.r.t.—time of conversion.

Table 5. Multi-class classification results of selected studies.

Study	Data Modalities	Classes	Accuracy
Liu, Liu, Cai, Che, Pujol, Kikinis, Feng and Fulham [230]	MRI, PET	AD, cMCI, ncMCI, NC	64.07
Dolph, Alam, Shboul, Samad and Iftekharuddin [229]	MRI	AD, MCI, NC	58
Shi, Zheng, Li, Zhang and Ying [318]	MRI, PET	AD, cMCI, ncMCI, NC	57.00
Liu, Zhang, Adeli and Shen [191]	MRI	AD, pMCI, sMCI, NC	51.8
Lee, Choi, Kim, Suk and Initiative [321]	MRI	AD, MCI, NC	71.17
Liu, Yadav, Fernandez-Granda and Razavian [281]	MRI	AD, MCI, NC	70

8. Deep Learning Techniques

8.1. Transfer Learning

With the popularity of deep neural networks in medical diagnostic systems, common challenges exist in practical applications [323]. These challenges include the availability of medical data and relevant labels. Current computer vision success is based on the ImageNet [324] hierarchical database, which contains millions of annotated images [325]. However, medical images are much smaller in quantity and require expert knowledge for labeling [326–329]. A potential solution to this problem is transfer learning—the transfer of knowledge across domains [330]. In image classification applications, transfer learning is commonly implemented by transferring model structure, weights, or parameters for classification in different feature spaces and distributions. Neural networks with transferred parameters have been shown to outperform the same neural networks with randomized parameters in convergence and have lower requirements for complicated and time-consuming hyperparameter searches [331].

There are three types of transfer learning: (1) transfer from Image-Net pre-trained models, e.g., Ding, Sohn, Kawczynski, Trivedi, Harnish, Jenkins, Lituiev, Copeland, Aboian and Mari Aparici [175] used the pre-trained Inception-V3 for the classification of AD vs. MCI, Bae et al. [332] applied a modified Inception-V4 with custom preprocessing for classification of AD vs. CN, Lin et al. [333] used the pre-trained AlexNet with RVR for regression, and Chen, Stromer, Alabdalahim, Schwab, Weih and Maier [148] selected pre-trained ResNet-152, VGG-16, and DenseNet-121 for screening and scoring of dementia using clock-drawing test images; (2) transfer from pre-trained networks for similar classification or prediction tasks, e.g., using a pre-trained network trained on one dataset for another dataset [90,281]; and (3) transfer from pre-trained networks used for different classification or prediction tasks, e.g., using an AD vs. NC pre-trained model for classification between pMCI and sMCI [192,237], or for MCI vs. NC [334].

Chen, Hsu, Yang, Tung, Luo, Liu, Hwang, Hwu and Tseng [163] transferred domain knowledge between different datasets for brain age prediction, while in a large-scale study, Bashyam, Erus, Doshi, Habes, Nasrallah, Truelove-Hill, Srinivasan, Mamourian, Pomponio and Fan [322] transferred a model used for brain age prediction to AD vs. NC and MCI vs. NC. Similar domain transfer has also been applied for transfer from Alzheimer’s disease to Parkinson’s disease [276]. Transfer learning is also applied for other data types, including eye-tracking, where datasets such as MIT GazeCapture, which are unrelated to Alzheimer’s or related diseases, can be utilized for gaze location estimation [116].

8.2. Ensemble Learning

Ensemble Learning in deep learning is the combination of multiple representations to achieve higher overall performance. Ensemble learning allows multiple representations and mitigates errors within individual neural networks [335]. These errors are not limited to misclassification but can also include underfitting or overfitting on training data. Underfitting occurs when the gradient descent is trapped in local minima, and the neural network fails to capture the underlying manifold of the training data. Conversely, overfitting occurs when irrelevant fluctuations in the training data are also captured by the neural network, resulting in lower generalization [336]. Ensemble learning has been widely applied in medical image classification [282,337] and the classification of AD and related diseases [338,339].

Ensemble learning can be performed at three levels: input, feature, and output. The input-level ensemble combines data prior to input into the neural network, e.g., the combination of adjacent slices of hippocampal data to construct mimic RGB channels [189] and the use of zero-masking for the fusion of concatenated MRI and PET inputs [340]. The feature-level ensemble combines features from patch-level, region-level, and subject-level sub-networks as input features for a classification module; Lian, Liu, Zhang and Shen [192] is an ideal example of a feature-level ensemble with hierarchical sub-networks at each level, where the outputs of each level are concatenated and used as input for the next level. The feature-level ensemble was also applied at individual feature levels, i.e., an ensemble of multi-scale patch-level sub-networks [319]. The output-level ensemble combines the predictions of component neural networks, e.g., through majority voting of prediction results [157]. Suk, Lee, Shen and Initiative [317] combined the outputs of multiple sparse regression models with varying regularization parameters for classification. Wang, Shen, Wang, Xiao, Deng, Wang and Zhao [282] utilized a probability-based fusion of softmax outputs from an ensemble of 3D-DenseNets. Apart from combining the outputs of neural networks, output-level ensembles also allow for the ensemble between neural networks and traditional machine learning classifiers [341]. A sub-category of ensemble learning is multi-view learning. Multi-view learning for AD neuroimaging is commonly linked to the 3D nature of available neuroimaging data and slice-based preprocessing, as described in Section 3.

Pan, Phan, Adel, Fossati, Gaidon, Wojak and Guedj [270] created a pyramid network of multiple CNN subnetworks with separable convolutions for each of the three views. The features were added for each view and concatenated for classification [270]. It is worth noting that although ensembling at all three levels is common amongst reviewed papers, there are only a few applications of the boosting method. Boosting is standard in machine learning applications [279]. In this method, individual components are trained sequentially in an adaptive manner. The ensemble with multiple modalities, also known as multi-modal fusion, is introduced in the subsequent section.

8.3. Multi-Modal Fusion

Individual modalities are fundamentally limited in their information content, e.g., genetic data cannot provide information on texture information of neuroimaging data, and MRI has good soft-tissue resolution but is not directly associated with Amyloid- β protein depositions. Fusing information from different modalities can provide a more comprehensive perspective of AD and related diseases. Multi-modal fusion is a common practice in the reviewed literature due to the availability of multi-modal data for AD and related diseases. The standard fusion method is the feature-level ensemble mentioned in Section 8.2, where at a particular stage of the model architecture, features produced by modality-dependent components are fused through concatenation or merging [166,318,342,343].

Liu, Liu, Cai, Che, Pujol, Kikinis, Feng and Fulham [230] performed the zero-masking of a single modality for a stacked autoencoder, which took both MRI and PET as input and achieved the fusion of the two modalities through data reconstruction of one zero-masked modality with only the other modality. Demographics and genetic biomarkers

are often fused with neuroimaging data through the concatenation of features extracted with fully connected layers [191]. For studies with 1D data or engineered features, the direct fusion of multi-modal data through concatenation or combined processing is achievable, e.g., the fusion of cognitive scores, volumetric features, gene expression, and CSF biomarkers [254,344,345].

Multimodal fusion is often combined with multi-scale or multi-view learning, e.g., studies have trained individual neural networks for patches of different sizes by processing MRI and PET images, where the inputs of the neural networks are concatenated for classification [170,320,346]. Through intricately designed connections between 1D and 3D network structures, Senanayake et al. [347] fused MRI and neuropsychological data. Likewise, Spasov, Passamonti, Duggento, Liò, Toschi and Initiative [283] constructed a more extensive architecture with the fusion of additional demographic and APOE-e4 genetic markers to input and Jacobian of sMRI images. Using multi-modal data is expected to improve the performance of neural networks. However, multi-modal fusion can also be limited by the availability of multi-modal data, especially for longitudinal studies. A basic overview of the common multi-modal fusion methods is shown in Figure 10.

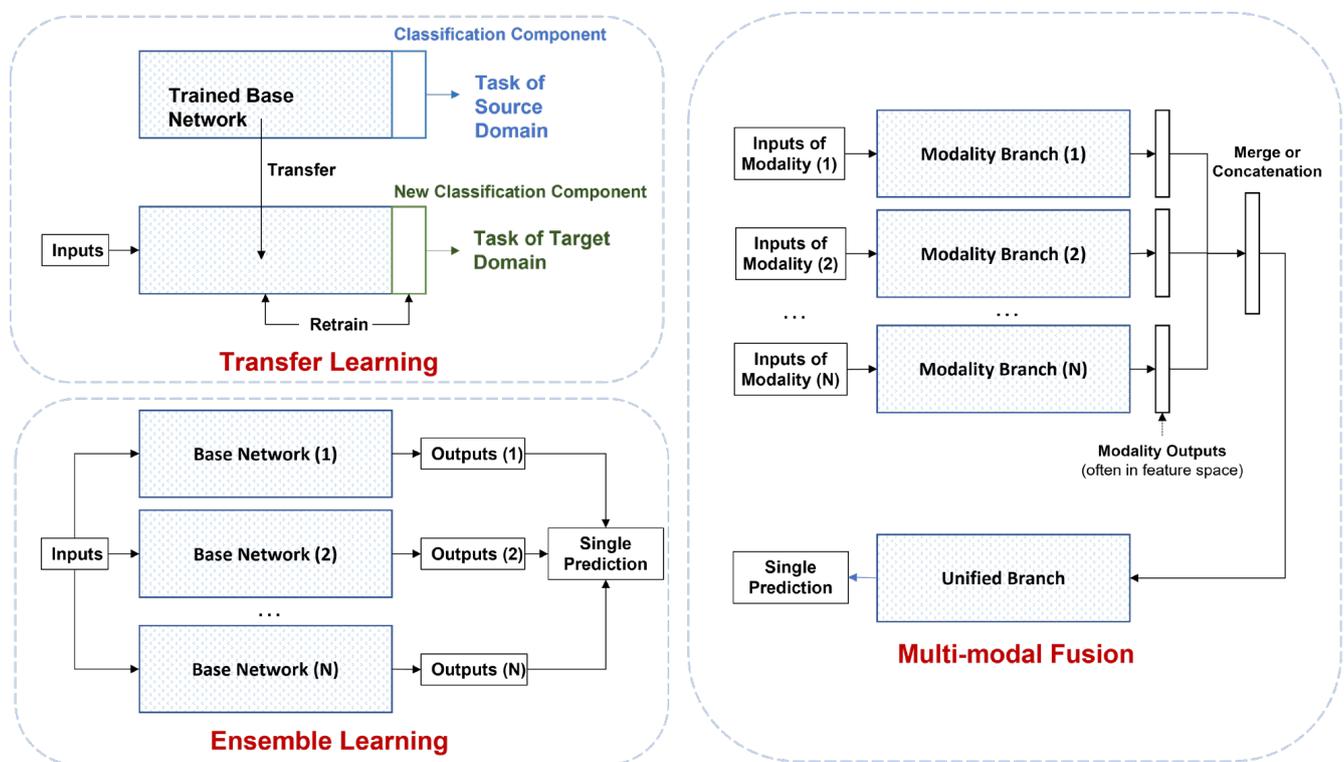


Figure 10. Overview of common multi-modal fusion methods applied in the field of AD, MCI and related diseases.

9. Training and Evaluation

The previous sections provide an overview of the myriad approaches to deep learning and relevant techniques applicable to its use for AD and related diseases. In general, two additional dependencies exist for any study/application of deep learning: how the method is trained and evaluated. Varying these two dependencies can generate wildly different results using the same approach and techniques. Different training and evaluation methods can also affect the interpretation and understanding of the results. In the following sections, we will first explore methods of evaluation in Section 9.1. This is the basis for an introduction to commonly applied training protocols in Section 9.2.

9.1. Evaluation Methods

9.1.1. Hold-Out and Cross-Validation

Data-driven models commonly suffer from the effects of overfitting, where the model learns from the noise and variance within the training data. The fundamental evaluation method for any deep learning algorithm is an independent test set sampled from the same distribution as the training set. This method, also called hold-out, provides unbiased measures of evaluation. The test set is generally 20–50% of the entire cohort, split according to the number of individual subjects. The number of test subjects directly affects the approximate measures of generalization, including the Hoeffding inequality bounds.

Another commonly implemented method is cross-validation (CV) [348,349], a measure of model robustness. There are various types of cross-validation, including k -fold cross-validation [350], balanced cross-validation, randomized cross-validation, and leave-one-out cross-validation (LOOCV) [351,352].

The fundamental cross-validation method, also known as k -fold cross-validation, is performed by splitting the data into k equal folds of similar categorical distributions to the original cohort. For each of k rounds, a single fold is used as the validation set, while the remainder is used for training an independent model. Categorical imbalance can cause biased performance, whereas balanced cross-validation undersamples, or oversamples components of the cross-validation split to provide a balanced training or testing set. Randomized cross-validation does not adhere to the rigid k -folds; a random split is provided for each of the unlimited cross-validation rounds. LOOCV is a variation of k -fold cross-validation where $k = 1$; this is commonly used for data with a limited number of subjects.

9.1.2. Metrics for Classification

The most common form of metric for classification is accuracy in predictions of defined labels. Apart from the basic classification accuracy, there are also sample-wise, subject-wise, and balanced accuracy, which weigh classification performance by categorical distribution. The measure of accuracy, or the correct classification rate, can be separated into a range of prediction measures, including true positive (TP), true negative (TN), false positive (FP), and false negative (FN). Similar metrics that measure various aspects of classification include the positive predictive value (precision), true positive rate (sensitivity), true negative rate (specificity), and the weighted combination of precision and sensitivity, the F_1 score. Simple variations of these metrics for binary classification:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}}, \quad (21)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \quad (22)$$

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad (23)$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}}, \quad (24)$$

$$\text{F}_1 \text{ score} = \frac{2\text{TP}}{2\text{TP} + (\text{FP} + \text{FN})}, \quad (25)$$

Apart from these fundamental metrics, other metrics are also used for specific purposes. A common measure for imbalanced datasets is balanced accuracy (BAC), accuracy weighted by class distribution. Kim et al. [353] used Cohen's Kappa to provide a comparison between observed and random accuracy, while Son, Oh, Oh, Kim, Lee, Roh and Kim [177] and Mårtensson et al. [354] applied it to assess inter-method agreement. The receiver operating characteristic (ROC) curve visualizes the trade-off between specificity and sensitivity. The area under the curve (AUC) for the ROC curve indicates separability between binary class

probabilities. AUC is one of the most common metrics of classification in AD-related publications. Other metrics include the Gini-coefficient, derived from the AUC-ROC, and the Kolmogorov–Smirnov statistic which compares categorical probability distributions. These terms are commonly defined in binary terms but can be easily generalized to multi-class scenarios, providing a group of metrics for each category.

9.1.3. Metrics for Prediction

Since prediction can be formulated into classification problems, most classification metrics in Section 9.1.2 can be applied as prediction metrics. If the prediction problem is formulated as a regression, regression metrics can be used. These include measures of error such as the mean absolute error (MAE) and mean squared error (MSE):

$$\begin{cases} \text{MAE} = \frac{1}{n} \sum_i |y_i - y'_i| \\ \text{MSE} = \frac{1}{n} \sum_i (y_i - y'_i)^2 \end{cases} \quad (26)$$

where n is the number of samples, y_i are labels and y'_i are predictions. Similar metrics include errors compared with simple predictors including the relative absolute error (RAE) and relative squared error (RSE),

$$\begin{cases} \text{RAE} = \frac{1}{n} \frac{\sum_i |y_i - y'_i|}{\sum_i |\bar{y}_i - y_i|} \\ \text{RSE} = \frac{1}{n} \frac{\sum_i (y_i - y'_i)^2}{\sum_i (\bar{y}_i - y_i)^2} \end{cases} \quad (27)$$

or the proportion of predictable variance, the coefficient of determination (R^2). Standard residual plots and residual analysis metrics can also be applied in this case. Since AD is a chronic medical condition, prediction can also be formulated as prognosis problems. Similar to the challenges in survival models, due to limitations in data collection from subjects suffering from AD or related diseases, there are cases of missing values or uncertain post-study outcomes. Metrics such as Harrell's C-index, or concordance index, take these 'censored' data into account by measuring the relationship between concordant and discordant pairs as follows:

$$\text{C index} = \frac{\sum_{i \neq j} \mathbf{1}_{t_i > t_j} \cdot \mathbf{1}_{\eta_i < \eta_j} \cdot \delta_j}{\sum_{i \neq j} \mathbf{1}_{t_i > t_j} \cdot \delta_j} \quad (28)$$

where t is time, η represents risk scores, and $\delta \in (0, 1)$ are auxiliary variables indicating 'censorship'. Li et al. [355] measured concordance with other survival analysis measures, such as the Kaplan–Meier estimate. However, there is a lack of deep learning studies extending this approach to provide individualized risk models, such as the Cox proportional hazard model, with metrics such as the cumulative hazard. Moreover, there is a lack of deep learning-related research into the treatment effect of AD treatment methods where we measure individualized treatment effect (ITE) and C-for-benefit.

9.1.4. Other Metrics

A range of other metrics are also used for various purposes including data reconstruction and generation. Such metrics include the peak signal-to-noise ratio (PSNR):

$$\text{PSNR} = \frac{n \times \max(y)^2}{\text{MSE}} \quad (29)$$

Another such metric is the structural similarity index (SSIM):

$$\text{SSIM}_{x,y} = \frac{(2\sigma_{x,y} + (K_2L)^2)(2\mu_x\mu_y + (K_1L)^2)}{(\mu_x^2 + \mu_y^2 + (K_1L)^2)(\sigma_x^2 + \sigma_y^2 + (K_2L)^2)}, \quad (30)$$

where x and y are two patches or images, K_1 and K_2 are constant values, μ_x is luminance, and σ_x is contrast.

$$\mu_x = \frac{1}{n} \sum_i x_i \text{ and } \sigma_x = \sqrt{\frac{1}{N-1} \sum_i (x_i - \mu_x)^2}, \quad (31)$$

Both PSNR and SSIM are metrics of generative models. Normalized cross-correlation is another metric that is used to measure the quality of feature selection in the form of visual attributions:

$$\text{NCC} = \frac{1}{n} \sum_{x,y} y_{x,y} \cdot y'_{x,y'} \quad (32)$$

where y' is the ground truth map for AD-affected regions. Segmentation metrics include the dice similarity coefficient, formulated in the same way as the F1-score in Section 9.1.2, where pixel-wise localization success is used instead of classification prediction.

9.1.5. Level of Evaluation

Data for AD and related diseases are inhomogeneous, with diverse data types and sources. Additional variance in preprocessing and processing data can provide significantly different inputs to the deep learning models. These differences give rise to problems in evaluation.

We can categorize two primary levels of evaluation: sample level and subject level. Sample-level evaluation is based on the model's performance in classifying or predicting data samples, while subject-level evaluation is based on individual subjects, e.g., AD or MCI patients. Sample-level evaluation occurs when multiple samples from the same subject are used in evaluation or when the data source does not indicate independence between samples. Subject-level evaluation can be based on either a single sample of data or multiple sample-level results; this provides a better representation in a real-world application.

With multiple data sources available for AD and related diseases, another level of evaluation has become more common in recent studies: validation with alternative datasets. This validation process involves using a trained model on a single dataset to provide outputs for data originating from an alternative source, e.g., a separate cohort or study. As the largest open library, data from ADNI is often used to train deep learning models, which are subsequently tested on data from other open libraries such as AIBL, OASIS [89], and private datasets [332].

9.1.6. Combination of Evaluation Methods

In current AD-related deep learning studies, a combination of evaluation techniques and metrics is often applied. A typical combination of evaluation techniques is cross-validation on the training set for hyper-parameter optimization and hold-out testing on the independent test set. The study-specific combinations are dependent on the overall objective of the deep learning model and training protocols applied to achieve this objective. Training protocols are discussed in detail in Section 9.2. In regard to classification, a combination of accuracy, sensitivity, and precision metrics is commonly measured for evaluation. Most MCI-conversion prediction problems are formulated as a classification based on a conversion time limit and share similar evaluation metrics.

9.1.7. Comparison and Ablation

Comparative studies provide insight into overall model performance, improvement, or limitation compared with the state-of-art studies of similar methods. In current AD-related deep learning literature, most studies apply comparative methods. These applications are often baseline machine learning or deep learning methods, such as SVM, Decision Tree, basic 2D-CNN, or variants of the proposed method. Most studies also contain a comparison of metrics from the literature. A comparison between models is commonly achieved by comparing the same or similar metrics under the assumption that the evaluation method and training protocols are similar. However, these comparisons can only serve as a rough performance evaluation due to differences in metrics definitions, data sampling, and processing methods. A major study to counter this problem of in-comparability and lack of reproducibility is the development of a standardized framework for machine learning algorithms, Clinica [90], and the extension of the framework for neural network evaluation [89].

With recent studies of increasingly advanced deep learning approaches, the performance increase between subsequent studies is small. For some studies, the performance gain in comparative models is within the approximated generalization error bounds. Therefore, instead of basic comparisons, some studies conducted statistical tests to validate performance gain. The DeLong test [268,356], which produces a confidence interval and standard error of difference, can compare the AUC of comparative models [355]. Apart from comparing model performance, the comparison of model architecture, data processing pipelines, and evaluation methods are also vital to propagating innovation. To establish a valid comparison within a single study, some scholars have used ablation studies to evaluate the individual components' importance in the overall modeling process [294]. In ablation studies, individual model components, feature inputs, or processing steps are removed to assess their importance. These studies, along with in-model comparisons, should be encouraged for all future studies to assess the effectiveness of the wide variety of model structures, techniques, and pipelines.

9.2. Training Protocols

As a data-driven approach, the practical application of deep learning to AD classification or prediction typically consists of a model or framework that acts as the basis of training to achieve the objective of individual algorithmic implementations. For most prediction and classification studies, this implies that the training and evaluation protocol has limited dependence on core architecture and utility. The protocol followed by training and evaluation impacts the models' performance, quality, and potential generalizability. This section will first introduce typical training and evaluation protocols, in addition to those methods mentioned in Section 7.1, and then highlight the hazard of information leakage. Then we will discuss appropriate optimization methods and the use of comparative studies.

9.2.1. Training and Evaluation Protocols

Standard training protocols involve using a single type of hold-out or cross-validation. However, using a validation set pre-selected from the training set or performing cross-validation on the training set for hyperparameter optimization is common. The metrics of performance on the validation set or CV of the training set provide a basis for the optimization. The use of a validation set is more commonly applied due to lower computational costs. When cross-validation is performed on the training set, testing can be performed on either a model from the cross-validation process or by retraining a new model with the entire training set.

More complex training and evaluation protocols can also be applied with sufficient data samples and computational resources. These protocols include component-wise parameter optimization, where each component of a neural network framework is trained or optimized separately. Random seeding is an evaluation procedure commonly used in machine learning, where multiple tests are run with different initial seeds for the ran-

dom number generators. As an interdisciplinary field between computer science and medicine, another procedure for evaluating model performance is human evaluation by medical practitioners.

9.2.2. Information Leakage

A significant concern under-addressed in current machine learning research in the field of AD or other diseases is the problem of information leakage [89]. This refers to the leakage of information from the training or validation set to the test set, introducing bias that can skew or invalidate the testing results. Information leakage can be categorized into three main types: (1) lack of test set, (2) invalid split, and (3) leakage in design.

The lack of an independent test set means the study cannot evaluate overfitting and generalization. The test set can also be invalidated if this subset is involved in hyperparameter optimization instead of a separate validation set. In these scenarios, the metrics do not provide a valid approximation of the model's actual performance and generalizability. Reported performance in these studies is typically significantly overstated quantitatively and should be considered as training performance only.

A test set can also be biased due to an invalid splitting process. A cross-sectional study might use the same subject's data samples at different time points as independent samples with longitudinal datasets in multiple data sources. However, if the split is performed according to images instead of subjects, the anatomical features of individual subjects could introduce bias that overfits the model at the subject level. A relative performance difference of 8% was found by Bäckström et al. [357]. Similar scenarios related to the invalid splitting of training and testing sets can also occur in other stages of the training process, e.g., data augmentation of an entire dataset before sample-wise splitting.

Apart from an invalid test set, information leakage can also occur through other factors in the studies. These factors include flawed data sourcing, where the same individual in multiple data sources is treated as independent. This leakage is possible for data sources such as ADNI, where some individuals are enrolled in multiple rounds of data collection and separated into different cohorts. Similar problems can occur with transfer learning where the source and target domains contain an amount of overlap. Information leakage is not limited to between the test and training sets but can also occur with intermediate subsets such as the validation set. With the validation set as an intermediate measure of model performance between the training and independent test set, we expect to utilize the validation set and, therefore, leak information. However, extensive hyperparameter optimization on the validation set can cause more information to be passed from this set to the model, causing overfitting on the validation set. Validation overfitting can negatively impact testing set performance and the overall generalizability of the model.

9.2.3. Optimization Protocols

Optimization is an essential component of overall training protocols and is divided into two main parts: optimization of parameters in the training process, and optimization of hyperparameters in the training protocol. Optimizers have become an essential component of current neural networks as they dictate the trajectory and means of gradient descent. Today's main optimizers are standardized, such as the stochastic gradient descent with momentum (SGDM) and RMSProp. Recent developments in neural network optimizers such as Adam and Adadelta have reduced dependence on learning rates and are more adaptive. Some studies include custom modifications, while others rely on model-dependent machine learning optimizers such as Limited memory BFGS [198,234]. The other main optimization component is the choice of hyperparameters, which define the overall structure and training specifications. Current methods are mostly based on grid search and random search, where a definite or random selection of hyperparameters is combined to train a model and evaluate performance. Statistical methods such as Bayesian optimization have only limited success due to the large search space and high computational cost.

9.3. Development Platforms

The modern development platforms are mainly based on MATLAB, R, and Python. MATLAB is a proprietary computational platform for science and engineering; its Deep Learning Toolbox provides an optimized framework for the efficient development and deployment of neural networks. The MathWorks File Exchange provides a platform for open-source code sharing, but the core platform is inherently closed-source. The most popular programming language for deep learning implementation and research is Python. The large open-source community has provided researchers with deep-learning libraries such as Tensorflow [358], Caffe [359], Theano [360], and PyTorch [361].

Higher-level APIs for these packages, such as Keras for Tensorflow and Fast.ai for Pytorch, have also been developed to lower scripting requirements and difficulty for researchers outside the field of bioinformatics and computer science. The majority of AD-related open-source deep learning packages or scripts are in Python. Keras is also available as a package for R, a popular programming language and platform for statistical computing and graphics. The access, availability, and reduced application difficulty of these platforms promote research into Alzheimer's disease and related diseases from an interdisciplinary perspective.

10. Path to Interpretation of Deep Learning Models

A significant challenge in applying DL to AD research is the lack of interpretability inherent in these often over-parameterized and highly complex data-driven models. A large number of studies have attempted to improve interpretability from different perspectives. Basic interpretation can be achieved through simple methods, e.g., correlation analysis and clustering of neural network features or predictions. Lin, Wu, Wu and Wu [333] analyzed the correlation between prediction error and individual features to validate the relationship between APOE-e4 and brain aging. Ding, Sohn, Kawczynski, Trivedi, Harnish, Jenkins, Lituiev, Copeland, Aboian and Mari Aparici [175] performed t-distributed stochastic neighbor embedding (t-SNE) on neural network-generated features to validate the model's understanding of AD disease stages, and a similar analysis with additional principal component analysis was performed by Son, Oh, Oh, Kim, Lee, Roh and Kim [177].

These simple methods offer a primer to the various methods utilized in the surveyed studies to explain model predictions and improve interpretability. In machine learning, the approaches to interpretability can be categorized into post hoc and intrinsic. Post hoc interpretation methods refer to probing and manipulation after the model is trained, while the intrinsic approach attempts to build a level of interpretability directly into the model architecture. However, since neural networks are inherent "black boxes," most deep learning methods surveyed focus on the post hoc approach. In this section, we detail three branches of the post hoc approach: (1) data-based methods, (2) architecture-based methods, and (3) model-agnostic visualization methods.

How data are processed and inputted into deep neural networks can fundamentally impact interpretability. ROI-based methods can provide a level of basic interpretability, which can be further translated to ROI sensitivity and feature stability. These measures can be projected to functional regions [198]. Feature maps can also be directly projected to ROIs [234]. Similarly, patch-based methods have some basic interpretability, e.g., Liu, Cheng, Wang, Wang and Initiative [170] visualized network attention areas by finding critical local patches that significantly affect class prediction probability, i.e., a drop in performance if they are removed. Graphical data also provide benefits for interpretability. Li, Rong, Meng, Lu, Kwok and Cheng [286] used analytic graph measures such as PageRank to determine the importance of each vertex in the input graph data, while Ju, Hu and Li [234] utilized brain networks of fMRI imaging to isolate functional regions of importance. At the voxel level, Duc, Ryu, Qureshi, Choi, Lee and Lee [273] visualized independent components of individual component analysis results as saliency maps on MRI and utilized these maps as inputs for the classification of AD and regression of MMSE. Methods utilizing

the inherent types and properties of data are often hybrids of traditional machine learning and deep learning that attempt to gain advantages from both approaches.

The choice of neural network architectures can also impact interpretability. As a classic example, the decoder component of convolutional autoencoders often contains transposed convolutional layers, or deconvolution layers, which are often used to reconstruct the input from the encoded feature space. The deconvolution process generates reconstructed images and feature maps, which can be compared globally with the input image, or locally with the anatomical structures of the input image [238].

A prime example is the use of generative models, such as variational autoencoders or GAN, to generate representational reconstructions through averaging iterative generations and structural transformations [244]. Alternatively, neural network architectures, such as the fully convolutional networks, which replace fully connected layers with global average pooling and SoftMax, can be designed to generate probabilistic maps [275,276]. The hierarchical framework implemented by Lee, Choi, Kim, Suk and Initiative [321] also allows for abnormality detection at various levels of voxels, patches, and regions, which can be combined to form a unified regional abnormality map. Visualization and interpretation methods that are dependent on architecture are also fundamentally constrained by the rigidity of the architecture and may not be able to adapt to new data or modalities. The data-based and architecture-based methods can be considered partially ad hoc. However, most data-model frameworks do not intrinsically provide functionality for tracing the decision process from inputs to classification probabilities and, therefore, cannot be considered intrinsically interpretable.

Transformer technology is a relatively new and powerful technique. The main area of application for transformers is language-based tasks. In future Alzheimer's disease research, transformers can extract meaningful information from medical records, patient interviews, and research articles by applying natural language processing techniques. Their ability to capture long-range dependencies in sequential data makes them highly suitable for analyzing textual data related to Alzheimer's disease.

Furthermore, transformers offer the potential for multimodal fusion in Alzheimer's disease research. Integrating data from multiple modalities, including imaging data, genetic information, and clinical assessments, can provide a comprehensive understanding of the disease. Transformers can facilitate the fusion of diverse data sources, capturing complex interactions and uncovering hidden relationships between different data types. One notable advantage of transformers is their attention mechanism, which enhances explainability. By highlighting relevant regions in images or identifying important words in the text, attention weights provide insights into the model's predictions. This interpretability feature can be valuable for medical professionals in understanding and validating the results of transformer models.

Alternatively, model-agnostic techniques exist to visualize feature saliency. The probabilistic maps generated through FCN by Qiu, Joshi, Miller, Xue, Zhou, Karjadi, Chang, Joshi, Dwyer and Zhu [275] are examples of CAM's dependence on the model structure. Recent developments of Grad-CAM utilize gradient information, allowing visualization of feature maps of various layers throughout the neural network. Tang, Chuang, DeCarli, Jin, Beckett, Keiser and Dugger [267] utilized a guided Grad-CAM with feature occlusion to identify amyloid- β plaques on immunohistochemically-stained slides. Similar methods were also applied to whole-brain MRI and identified GM regions around the hippocampus and ventricles that were consistent with anatomical pathology [281,290,322]. Monitoring model output with perturbations in the input is another method to interpret neural network function. An example of this approach is the swap test proposed by Nigri et al. [362], where patches of the image of interest are replaced by patches from reference images of an alternative class. The hippocampal region showed the highest impact on model predictions for the swap test. Mean relevance maps can also be generated for each category to interpret disease stages and progression from the perspective of groups instead of individuals [355]. Regional saliency maps were also combined with hippocampal segmentation by Liu, Li,

Yan, Wang, Ma, Shen, Xu and Initiative [169], while attention maps can also be included in the neural network framework to enhance performance and localization results [363]. However, as with the methods mentioned above, providing quantitative assessments of these visualizations is difficult.

Data-based methods, architectural-based methods, and model-agnostic visualization techniques are all constrained by their fundamental limits, e.g., the information content of patch-based ensembles is limited by the patch dimensions. The generative models summarized in Section 6.2 emphasize new approaches that dedicate modeling to interpretation by changing the core aim to visual attribution [250] and designing neural networks that are inherently interpretable, e.g., invertible neural networks [176]. These approaches present the most novel aspects on the path to interpretation. A basic summary is presented in Figure 11.

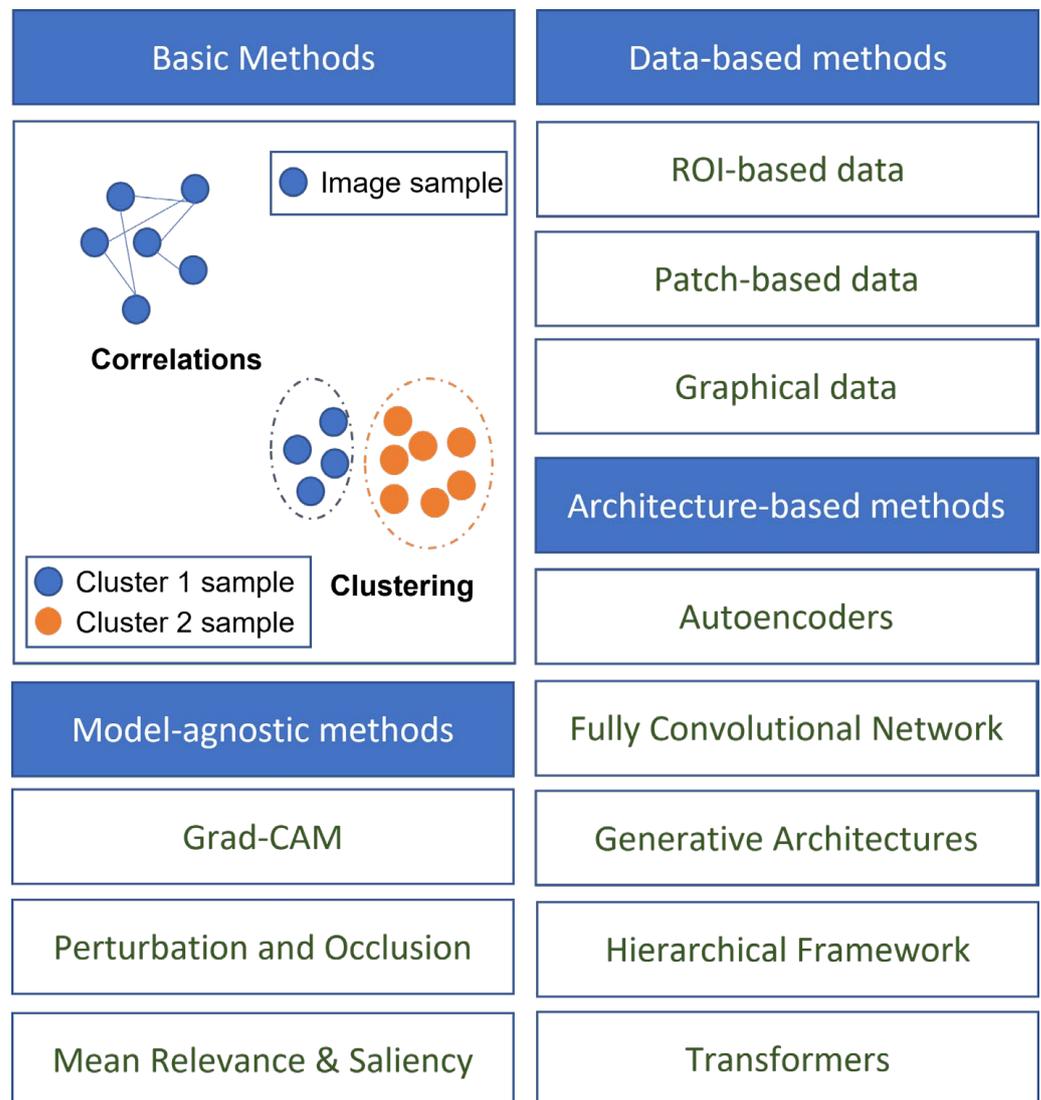


Figure 11. Overview of common interpretation methods found in surveyed literature.

11. Path to Generalization in the Real World

A major challenge in DL is the real-world generalization of models. Generalization is heavily affected by the data used to train the models. Most of the literature surveyed utilized data collected under strict acquisition protocols with specified modalities, types, and hardware and are often not representative of clinical settings. Preprocessing is commonly applied to eliminate some of these variabilities, but its variations and subjectivity can introduce uncertainty and different levels of quality and is, therefore, a double-edged sword. Mårtensson, Ferreira, Granberg, Cavallin, Oppedal, Padovani, Rektorova, Bonanni, Pardini and Kramberger [354] extensively assessed training and testing in different data domains.

While the proposed recurrent CNN showed consistency across different datasets, the recent study confirmed the expected degradation of performance in evaluating data collected through protocols that differ from those used for training data. As a solution, including a broader range of protocols in training, increased generalization performance in unseen data. Though this study is limited to a single CNN-based model and does not provide a definite conclusion, it provides valuable insight into the possible generalization challenges and the importance of data heterogeneity in countering them. It is established that generalization is heavily affected by the amount of data used in training and evaluation. Apart from collecting new data, methods to increase data quantity and heterogeneity during training include implementing lower-dimensional data (i.e., use of 2D slices of 3D scans), data augmentation, and the careful use of generative models. An alternative approach focuses on reducing the model's training data requirement, utilizing a train-test split of 50% or lower. These approaches are often semi-supervised and provide a larger testing set and a more accurate approximation of generalizability.

The theoretical forefront of this problem lies in estimating the 'generalization gap,' the difference between metrics derived from the independent test set and real-world scenarios. The approximate generalization bounds derived from the Hoeffding inequality [364] are based on a range of assumptions but provide a core insight into the relationship between the testing set and the approximate generalization gap—the bound is proportional to the inverted root of the sample size. Even though the amount of data required by these worst-case bounds is likely impossible to achieve in practical data collection, generalizability benefits from a larger sample size. A complete estimation is based on model complexity, usually measured through the Vapnik–Chervonenkis dimension. Alternative methods to derive generalization bounds have been explored, which include using the validation set [365], measurement of network smoothness [366], and comparison of generalization error between deep neural networks and humans [367]. An alternative approach to estimating generalization is to tackle the concept of label inhomogeneity due to misdiagnosis. Wu et al. [368] proposed using unsure data models to account for discordant MCI samples for which conversion is uncertain.

Apart from the technical and theoretical pathways to generalization, another key consideration is the practical generalization to clinical use, especially in mass screening. False positives produced by deep learning models in small-scale studies have been found to increase radiologist workload. In large-scale screening, the overdiagnosis caused by the number of false positives may significantly affect cost and efficiency [369,370]. Close monitoring of false-positive rates alongside generalization gap approximations should be a key aspect of evaluation in these scenarios.

12. Conclusions

In the past 13 years, many deep-learning studies have been conducted for AD and related diseases, producing various techniques, models, and protocols. We have provided a comprehensive summary of these major components that contribute to a deep learning study and a summary of the most recent advances, including recurrent neural networks, graph and geometric neural networks, as well as generative modeling.

These studies have shown promising results for a broad range of tasks, including image processing, disease categorical classification, and disease progression prediction. However, the wide variety of approaches shows a lack of consistency, and few studies provide standardized benchmarks for comparison. Most of these studies are research-oriented; few studies have conducted or simulated evaluations in clinical settings. These issues contribute to the challenges of interpretation and generalization of deep learning.

This review provides a glimpse into the possible solutions for interpretation, e.g., visualization techniques and inherently interpretable architectures. It also provides insights into potential pathways for generalization, e.g., data heterogeneity, data quantity, and generalization gap approximation. Apart from the key aspects of interpretation and generalization of neural networks, there are many other aspects of potential research, e.g., deep learning for polygenic studies [371] and the application of transformer-based foundational models. Combined with the continuously developing model architectures, these pathways will guide us toward more robust and clinically feasible deep-learning models for AD and related diseases.

Author Contributions: Q.Z.: Conceptualization, Software, Formal analysis, Writing—Original Draft, Visualization. J.W.: Software, Validation, Investigation, Writing—Original Draft, Visualization. X.Y.: Methodology, Formal analysis, Resources. S.W.: Methodology, Validation, Resources, Writing—Review and Editing, Supervision, Funding acquisition. Y.Z.: Conceptualization, Investigation, Writing—Review and Editing, Supervision, Project administration, Funding acquisition. All authors have read and agreed to the published version of the manuscript.

Funding: This paper is partially supported by MRC, UK (MC_PC_17171); Royal Society, UK (RP202G0230); BHF, UK (AA/18/3/34220); Hope Foundation for Cancer Research, UK (RM60G0680); GCRF, UK (P202PF11); Sino-UK Industrial Fund, UK (RP202G0289); LIAS, UK (P202ED10, P202RE969); Data Science Enhancement Fund, UK (P202RE237); Fight for Sight, UK (24NN201); Sino-UK Education Fund, UK (OP202006); BBSRC, UK (RM32G0178B8).

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

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. [[CrossRef](#)]
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. [[CrossRef](#)] [[PubMed](#)]
3. Nam, D.; Chapiro, J.; Paradis, V.; Seraphin, T.P.; Kather, J.N. Artificial intelligence in liver diseases: Improving diagnostics, prognostics and response prediction. *JHEP Rep.* **2022**, *4*, 100443. [[CrossRef](#)] [[PubMed](#)]
4. Tatulian, S.A. Challenges and hopes for Alzheimer's disease. *Drug Discov. Today* **2022**, *27*, 1027–1043. [[CrossRef](#)] [[PubMed](#)]
5. Brookmeyer, R.; Johnson, E.; Ziegler-Graham, K.; Arrighi, H.M. Forecasting the global burden of Alzheimer's disease. *Alzheimer's Dement.* **2007**, *3*, 186–191. [[CrossRef](#)]
6. Loi, S.M.; Pijnenberg, Y.; Velakoulis, D. Recent research advances in young-onset dementia. *Curr. Opin. Psychiatry* **2023**, *36*, 126–133. [[CrossRef](#)] [[PubMed](#)]
7. Zhang, W.; Xu, C.; Sun, J.; Shen, H.-M.; Wang, J.; Yang, C. Impairment of the autophagy–lysosomal pathway in Alzheimer's diseases: Pathogenic mechanisms and therapeutic potential. *Acta Pharm. Sin. B* **2022**, *12*, 1019–1040. [[CrossRef](#)] [[PubMed](#)]
8. Boeve, B.F.; Boxer, A.L.; Kumfor, F.; Pijnenburg, Y.; Rohrer, J.D. Advances and controversies in frontotemporal dementia: Diagnosis, biomarkers, and therapeutic considerations. *Lancet Neurol.* **2022**, *21*, 258–272. [[CrossRef](#)]
9. Sögis, E.; Dauvillier, J.; Leontjeva, A.; Adler, P.; Hindie, V.; Moncion, T.; Collura, V.; Daudin, R.; Loe-Mie, Y.; Hérault, Y.; et al. HENA, heterogeneous network-based data set for Alzheimer's disease. *Sci. Data* **2019**, *6*, 151. [[CrossRef](#)]
10. Wimo, A.; Jönsson, L.; Bond, J.; Prince, M.; Winblad, B.; Alzheimer Disease International. The worldwide economic impact of dementia 2010. *Alzheimer's Dement.* **2013**, *9*, 1–11.e3. [[CrossRef](#)] [[PubMed](#)]
11. López-Cuenca, I.; Nebreda, A.; García-Colomo, A.; Salobar-García, E.; de Frutos-Lucas, J.; Bruña, R.; Ramírez, A.I.; Ramírez-Torano, F.; Salazar, J.J.; Barabash, A.; et al. Early visual alterations in individuals at-risk of Alzheimer's disease: A multidisciplinary approach. *Alzheimer's Res. Ther.* **2023**, *15*, 19. [[CrossRef](#)] [[PubMed](#)]

12. Toschi, N.; Baldacci, F.; Zetterberg, H.; Blennow, K.; Kilimann, I.; Teipel, S.J.; Cavado, E.; dos Santos, A.M.; Epelbaum, S.; Lamari, F. Alzheimer's disease biomarker-guided diagnostic workflow using the added value of six combined cerebrospinal fluid candidates: Ab1–42, total-tau, phosphorylated-tau, NFL, neurogranin, and YKL-40. *Alzheimer's Dement.* **2017**, *1*, 10.
13. Scheltens, P.; Blennow, K.; Breteler, M.M.B.; de Strooper, B.; Frisoni, G.B.; Salloway, S.; Van der Flier, W.M. Alzheimer's disease. *Lancet* **2016**, *388*, 505–517. [[CrossRef](#)]
14. Vogt, A.-C.S.; Jennings, G.T.; Mohsen, M.O.; Vogel, M.; Bachmann, M.F. Alzheimer's Disease: A Brief History of Immunotherapies Targeting Amyloid β . *Int. J. Mol. Sci.* **2023**, *24*, 3895. [[CrossRef](#)] [[PubMed](#)]
15. Van der Lee, S.J.; Wolters, F.J.; Ikram, M.K.; Hofman, A.; Ikram, M.A.; Amin, N.; van Duijn, C.M. The effect of APOE and other common genetic variants on the onset of Alzheimer's disease and dementia: A community-based cohort study. *Lancet Neurol.* **2018**, *17*, 434–444. [[CrossRef](#)] [[PubMed](#)]
16. Fortea, J.; Vilaplana, E.; Carmona-Iragui, M.; Benejam, B.; Videla, L.; Barroeta, I.; Fernández, S.; Altuna, M.; Pegueroles, J.; Montal, V. Clinical and biomarker changes of Alzheimer's disease in adults with Down syndrome: A cross-sectional study. *Lancet* **2020**, *395*, 1988–1997. [[CrossRef](#)]
17. Brett, B.L.; Gardner, R.C.; Godbout, J.; Dams-O'Connor, K.; Keene, C.D. Traumatic Brain Injury and Risk of Neurodegenerative Disorder. *Biol. Psychiatry* **2021**, *91*, 498–507. [[CrossRef](#)]
18. Letnes, J.M.; Nes, B.M.; Wisløff, U. Age-related decline in peak oxygen uptake: Cross-sectional vs. longitudinal findings. A review. *Int. J. Cardiol. Cardiovasc. Risk Prev.* **2023**, *16*, 200171. [[CrossRef](#)]
19. Tari, A.R.; Nauman, J.; Zisko, N.; Skjellegrind, H.K.; Bosnes, I.; Bergh, S.; Stensvold, D.; Selbæk, G.; Wisløff, U. Temporal changes in cardiorespiratory fitness and risk of dementia incidence and mortality: A population-based prospective cohort study. *Lancet Public Health* **2019**, *4*, e565–e574. [[CrossRef](#)] [[PubMed](#)]
20. Birkenhäger, W.H.; Forette, F.; Seux, M.-L.; Wang, J.-G.; Staessen, J.A. Blood Pressure, Cognitive Functions, and Prevention of Dementias in Older Patients with Hypertension. *Arch. Intern. Med.* **2001**, *161*, 152–156. [[CrossRef](#)]
21. Donaghy, P.C.; Ciafone, J.; Durcan, R.; Hamilton, C.A.; Barker, S.; Lloyd, J.; Firbank, M.; Allan, L.M.; O'Brien, J.T.; Taylor, J.-P.; et al. Mild cognitive impairment with Lewy bodies: Neuropsychiatric supportive symptoms and cognitive profile. *Psychol. Med.* **2020**, *52*, 1147–1155. [[CrossRef](#)] [[PubMed](#)]
22. Burns, A.; Iliffe, S. Alzheimer's disease. *BMJ Br. Med. J. (Int. Ed.)* **2009**, *338*, 467–471. [[CrossRef](#)] [[PubMed](#)]
23. Roberts, R.; Knopman, D.S. Classification and Epidemiology of MCI. *Clin. Geriatr. Med.* **2013**, *29*, 753–772. [[CrossRef](#)] [[PubMed](#)]
24. Petersen, R.C.; Lopez, O.; Armstrong, M.J.; Getchius, T.S.; Ganguli, M.; Gloss, D.; Gronseth, G.S.; Marson, D.; Pringsheim, T.; Day, G.S.; et al. Author response: Practice guideline update summary: Mild cognitive impairment: Report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology. *Neurology* **2018**, *91*, 373–374. [[CrossRef](#)]
25. Ward, A.; Tardiff, S.; Dye, C.; Arrighi, H.M. Rate of Conversion from Prodromal Alzheimer's Disease to Alzheimer's Dementia: A Systematic Review of the Literature. *Dement. Geriatr. Cogn. Disord. Extra* **2013**, *3*, 320–332. [[CrossRef](#)]
26. Mitchell, A.J.; Shiri-Feshki, M. Rate of progression of mild cognitive impairment to dementia—Meta-analysis of 41 robust inception cohort studies. *Acta Psychiatr. Scand.* **2009**, *119*, 252–265. [[CrossRef](#)]
27. Sherman, D.S.; Mauser, J.; Nuno, M.; Sherzai, D. The Efficacy of Cognitive Intervention in Mild Cognitive Impairment (MCI): A Meta-Analysis of Outcomes on Neuropsychological Measures. *Neuropsychol. Rev.* **2017**, *27*, 440–484. [[CrossRef](#)] [[PubMed](#)]
28. Ala, T.; Bakir, D.; Goel, S.; Feller, N.; Botchway, A.; Womack, C. A Mini-Mental State Examination Formula May Help to Distinguish Alzheimer's Disease from Dementia with Lewy Bodies. *J. Alzheimer's Dis.* **2022**, *89*, 1119–1129. [[CrossRef](#)] [[PubMed](#)]
29. McGurn, M.; Dworkin, J.D.; Chapman, S.; Huey, E.D.; Cosentino, S.; Louis, E.D. Can the Montreal Cognitive Assessment and Mini-Mental State Examination detect cognitive decline in elderly patients with essential tremor? *Clin. Neuropsychol.* **2022**, 1–18. [[CrossRef](#)] [[PubMed](#)]
30. Folstein, M.F.; Folstein, S.E.; McHugh, P.R. "Mini-Mental State". A Practical Method for Grading the Cognitive State of Patients for the Clinician. *J. Psychiatr. Res.* **1975**, *12*, 189–198. [[CrossRef](#)] [[PubMed](#)]
31. Tzeng, R.-C.; Yang, Y.-W.; Hsu, K.-C.; Chang, H.-T.; Chiu, P.-Y. Sum of boxes of the clinical dementia rating scale highly predicts conversion or reversion in predementia stages. *Front. Aging Neurosci.* **2022**, *14*, 1021792. [[CrossRef](#)] [[PubMed](#)]
32. Hughes, C.P.; Berg, L.; Danziger, W.L.; Coben, L.A.; Martin, R.L. A New Clinical Scale for the Staging of Dementia. *Br. J. Psychiatry* **1982**, *140*, 566–572. [[CrossRef](#)]
33. Titheradge, D.; Isaac, M.; Bremner, S.; Tabet, N. Cambridge Cognitive Examination and Hachinski Ischemic Score as predictors of MRI confirmed pathology in dementia: A cross-sectional study. *Int. J. Clin. Pract.* **2019**, *74*, e13446. [[CrossRef](#)] [[PubMed](#)]
34. Schmand, B.; Walstra, G.; Lindeboom, J.; Teunisse, S.; Jonker, C. Early detection of Alzheimer's disease using the Cambridge Cognitive Examination (CAMCOG). *Psychol. Med.* **2000**, *30*, 619–627. [[CrossRef](#)] [[PubMed](#)]
35. López-Cuenca, I.; Marcos-Dolado, A.; Yus-Fuertes, M.; Salobar-García, E.; Elvira-Hurtado, L.; Fernández-Albarral, J.A.; Salazar, J.J.; Ramírez, A.I.; Sánchez-Puebla, L.; Fuentes-Ferrer, M.E.; et al. The relationship between retinal layers and brain areas in asymptomatic first-degree relatives of sporadic forms of Alzheimer's disease: An exploratory analysis. *Alzheimer's Res. Ther.* **2022**, *14*, 79. [[CrossRef](#)] [[PubMed](#)]

36. Rocha, A.; Bellaver, B.; Souza, D.G.; Schu, G.; Fontana, I.C.; Venturin, G.T.; Greggio, S.; Fontella, F.U.; Schiavenin, M.L.; Machado, L.S.; et al. Clozapine induces astrocyte-dependent FDG-PET hypometabolism. *Eur. J. Nucl. Med.* **2022**, *49*, 2251–2264. [[CrossRef](#)]
37. Oe, K.; Zeng, F.; Niikura, T.; Fukui, T.; Sawauchi, K.; Matsumoto, T.; Nogami, M.; Murakami, T.; Kuroda, R. Influence of Metal Implants on Quantitative Evaluation of Bone Single-Photon Emission Computed Tomography/Computed Tomography. *J. Clin. Med.* **2022**, *11*, 6732. [[CrossRef](#)]
38. Madetko-Alster, N.; Alster, P.; Migda, B.; Nieciecki, M.; Koziowski, D.; Królicki, L. The Use of Cerebellar Hypoperfusion Assessment in the Differential Diagnosis of Multiple System Atrophy with Parkinsonism and Progressive Supranuclear Palsy-Parkinsonism Predominant. *Diagnostics* **2022**, *12*, 3022. [[CrossRef](#)]
39. Charpentier, P.; Lavenu, I.; Defebvre, L.; Duhamel, A.; Lecouffe, P.; Pasquier, F.; Steinling, M. Alzheimer's disease and frontotemporal dementia are differentiated by discriminant analysis applied to 99mTc HmPAO SPECT data. *J. Neurol. Neurosurg. Psychiatry* **2000**, *69*, 661–663. [[CrossRef](#)]
40. Garriga, M.; Emila, M.; Mir, M.; Eal-Baradie, R.; Ehuertas, S.; Ecastejon, C.; Ecasas, L.; Badenes, D.; Gimenez, N.; Font, M.A.; et al. 123I-FP-CIT SPECT imaging in early diagnosis of dementia in patients with and without a vascular component. *Front. Syst. Neurosci.* **2015**, *9*, 99. [[CrossRef](#)]
41. Fortea, J.; Carmona-Iragui, M.; Benejam, B.; Fernández, S.; Videla, L.; Barroeta, I.; Alcolea, D.; Pegueroles, J.; Muñoz, L.; Belbin, O.; et al. Plasma and CSF biomarkers for the diagnosis of Alzheimer's disease in adults with Down syndrome: A cross-sectional study. *Lancet Neurol.* **2018**, *17*, 860–869. [[CrossRef](#)] [[PubMed](#)]
42. Olsson, B.; Lautner, R.; Andreasson, U.; Öhrfelt, A.; Portelius, E.; Bjerke, M.; Hölttä, M.; Rosén, C.; Olsson, C.; Strobel, G.; et al. CSF and blood biomarkers for the diagnosis of Alzheimer's disease: A systematic review and meta-analysis. *Lancet Neurol.* **2016**, *15*, 673–684. [[CrossRef](#)]
43. Chen, C.L.; Lu, Q.; Moorakonda, R.B.; Kandiah, N.; Tan, B.Y.; Villaraza, S.G.; Cano, J.; Venketasubramanian, N. Alzheimer's Disease THERapy with NEuroaid (ATHENE): A Randomized Double-Blind Delayed-Start Trial. *J. Am. Med. Dir. Assoc.* **2021**, *23*, 379–386.e3. [[CrossRef](#)] [[PubMed](#)]
44. McKhann, G.; Drachman, D.; Folstein, M.; Katzman, R.; Price, D.; Stadlan, E.M. Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology* **1984**, *34*, 939–944. [[CrossRef](#)] [[PubMed](#)]
45. Dubois, B.; Feldman, H.H.; Jacova, C.; DeKosky, S.T.; Barberger-Gateau, P.; Cummings, J.L.; Delacourte, A.; Galasko, D.; Gauthier, S.; Jicha, G.A.; et al. Research Criteria for the Diagnosis of Alzheimer's Disease: Revising the NINCDS-ADRDA Criteria. *Lancet Neurol.* **2007**, *6*, 734–746. [[CrossRef](#)] [[PubMed](#)]
46. Dubois, B.; Feldman, H.H.; Jacova, C.; Cummings, J.L.; DeKosky, S.T.; Barberger-Gateau, P.; Delacourte, A.; Frisoni, G.; Fox, N.C.; Galasko, D.; et al. Revising the definition of Alzheimer's disease: A new lexicon. *Lancet Neurol.* **2010**, *9*, 1118–1127. [[CrossRef](#)] [[PubMed](#)]
47. Dubois, B.; Feldman, H.H.; Jacova, C.; Hampel, H.; Molinuevo, J.L.; Blennow, K.; DeKosky, S.T.; Gauthier, S.; Selkoe, D.; Bateman, R.; et al. Advancing research diagnostic criteria for Alzheimer's disease: The IWG-2 criteria. *Lancet Neurol.* **2014**, *13*, 614–629. [[CrossRef](#)] [[PubMed](#)]
48. Jack, C.R., Jr.; Albert, M.; Knopman, D.S.; McKhann, G.M.; Sperling, R.A.; Carillo, M.; Thies, W.; Phelps, C.H. Introduction to revised criteria for the diagnosis of Alzheimer's disease: National Institute on Aging and the Alzheimer Association Workgroups. *Alzheimer's Dement. J. Alzheimer's Assoc.* **2011**, *7*, 257. [[CrossRef](#)]
49. Zhou, J.; Benoit, M.; Sharoar, G. Recent advances in pre-clinical diagnosis of Alzheimer's disease. *Metab. Brain Dis.* **2021**, *37*, 1703–1725. [[CrossRef](#)]
50. Sperling, R.A.; Aisen, P.S.; Beckett, L.A.; Bennett, D.A.; Craft, S.; Fagan, A.M.; Iwatsubo, T.; Jack, C.R., Jr.; Kaye, J.; Montine, T.J.; et al. Toward Defining the Preclinical Stages of Alzheimer's Disease: Recommendations from the National Institute on Aging-Alzheimer's Association Workgroups on Diagnostic Guidelines for Alzheimer's Disease. *Alzheimer's Dement.* **2011**, *7*, 280–292. [[CrossRef](#)]
51. Chao, L.; Mueller, S.; Buckley, S.; Peek, K.; Raptentsetseng, S.; Elman, J.; Yaffe, K.; Miller, B.; Kramer, J.; Madison, C.; et al. Evidence of neurodegeneration in brains of older adults who do not yet fulfill MCI criteria. *Neurobiol. Aging* **2010**, *31*, 368–377. [[CrossRef](#)] [[PubMed](#)]
52. Albert, M.S.; DeKosky, S.T.; Dickson, D.; Dubois, B.; Feldman, H.H.; Fox, N.C.; Gamst, A.; Holtzman, D.M.; Jagust, W.J.; Petersen, R.C.; et al. The Diagnosis of Mild Cognitive Impairment due to Alzheimer's Disease: Recommendations from the National Institute on Aging-Alzheimer's Association Workgroups on Diagnostic Guidelines for Alzheimer's Disease. *Alzheimer's Dement.* **2011**, *7*, 270–279. [[CrossRef](#)] [[PubMed](#)]
53. McGrattan, A.M.; Pakpahan, E.; Siervo, M.; Mohan, D.; Reidpath, D.D.; Prina, M.; Allotey, P.; Zhu, Y.; Shulin, C.; Yates, J. Risk of conversion from mild cognitive impairment to dementia in low-and middle-income countries: A systematic review and meta-analysis. *Alzheimer's Dement. Transl. Res. Clin. Interv.* **2022**, *8*, e12267. [[CrossRef](#)] [[PubMed](#)]

54. McKhann, G.M.; Knopman, D.S.; Chertkow, H.; Hyman, B.T.; Jack, C.R., Jr.; Kawas, C.H.; Klunk, W.E.; Koroshetz, W.J.; Manly, J.J.; Mayeux, R. The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's Dement.* **2011**, *7*, 263–269. [[CrossRef](#)] [[PubMed](#)]
55. Hyman, B.T.; Phelps, C.H.; Beach, T.G.; Bigio, E.H.; Cairns, N.J.; Carrillo, M.C.; Dickson, D.W.; Duyckaerts, C.; Frosch, M.P.; Masliah, E. National Institute on Aging-Alzheimer's Association guidelines for the neuropathologic assessment of Alzheimer's disease. *Alzheimer's Dement.* **2012**, *8*, 1–13. [[CrossRef](#)]
56. Risacher, S.; Saykin, A.; Wes, J.; Shen, L.; Firpi, H.; McDonald, B. Baseline MRI Predictors of Conversion from MCI to Probable AD in the ADNI Cohort. *Curr. Alzheimer Res.* **2009**, *6*, 347–361. [[CrossRef](#)]
57. Qiu, A.; Fennema-Notestine, C.; Dale, A.M.; Miller, M.I. Regional shape abnormalities in mild cognitive impairment and Alzheimer's disease. *Neuroimage* **2009**, *45*, 656–661. [[CrossRef](#)]
58. Guévremont, D.; Tsui, H.; Knight, R.; Fowler, C.J.; Masters, C.L.; Martins, R.N.; Abraham, W.C.; Tate, W.P.; Cutfield, N.; Williams, J.M. Plasma microRNA vary in association with the progression of Alzheimer's disease. *Alzheimer's Dement.* **2022**, *14*, e12251. [[CrossRef](#)]
59. Mesa-Herrera, F.; Marin, R.; Torrealba, E.; Santos, G.; Diaz, M. Neuronal ER-Signalosome Proteins as Early Biomarkers in Prodromal Alzheimer's Disease Independent of Amyloid-beta Production and Tau Phosphorylation. *Front. Mol. Neurosci.* **2022**, *15*, 1–20. [[CrossRef](#)]
60. Shahid, S.S.; Wen, Q.; Risacher, S.L.; Farlow, M.R.; Unverzagt, F.W.; Apostolova, L.G.; Foroud, T.M.; Zetterberg, H.; Blennow, K.; Saykina, A.J.; et al. Hippocampal-subfield microstructures and their relation to plasma biomarkers in Alzheimer's disease. *Brain* **2022**, *145*, 2149–2160. [[CrossRef](#)]
61. Vaghari, D.; Kabir, E.; Henson, R.N. Late combination shows that MEG adds to MRI in classifying MCI versus controls. *Neuroimage* **2022**, *252*, 119054. [[CrossRef](#)] [[PubMed](#)]
62. 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. [[CrossRef](#)]
63. 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.
64. 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. [[CrossRef](#)]
65. 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.
66. 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. [[CrossRef](#)]
67. Hong, S.; Coelho, C.A.; Park, J. An Exact and Near-Exact Distribution Approach to the Behrens–Fisher Problem. *Mathematics* **2022**, *10*, 2953. [[CrossRef](#)]
68. Esteki, S.; Naghsh-Nilchi, A.R. Frequency component Kernel for SVM. *Neural Comput. Appl.* **2022**, *34*, 22449–22464. [[CrossRef](#)]
69. Nayak, J.; Swapnarekha, H.; Naik, B.; Dhiman, G.; Vimal, S. 25 Years of Particle Swarm Optimization: Flourishing Voyage of Two Decades. *Arch. Comput. Methods Eng.* **2022**, *30*, 1663–1725. [[CrossRef](#)]
70. Sonoda, S.; Murata, N. Neural network with unbounded activation functions is universal approximator. *Appl. Comput. Harmon. Anal.* **2017**, *43*, 233–268. [[CrossRef](#)]
71. McKinney, S.M.; Sieniek, M.; Godbole, V.; Godwin, J.; Antropova, N.; Ashrafian, H.; Back, T.; Chesus, M.; Corrado, G.S.; Darzi, A.; et al. International evaluation of an AI system for breast cancer screening. *Nature* **2020**, *577*, 89–94. [[CrossRef](#)]
72. Zaidi, S.M.A.; Habib, S.S.; Van Ginneken, B.; Ferrand, R.A.; Creswell, J.; Khowaja, S.; Khan, A. Evaluation of the diagnostic accuracy of Computer-Aided Detection of tuberculosis on Chest radiography among private sector patients in Pakistan. *Sci. Rep.* **2018**, *8*, 12339. [[CrossRef](#)] [[PubMed](#)]
73. Kang, D.; Park, J.E.; Kim, Y.-H.; Kim, J.H.; Oh, J.Y.; Kim, J.; Kim, Y.; Kim, S.T.; Kim, H.S. Diffusion radiomics as a diagnostic model for atypical manifestation of primary central nervous system lymphoma: Development and multicenter external validation. *Neuro-Oncology* **2018**, *20*, 1251–1261. [[CrossRef](#)]
74. 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. [[CrossRef](#)] [[PubMed](#)]
75. Zheng, Q.; Zhang, Y.Y.; Li, H.L.; Tong, X.R.; Ouyang, M.H. How segmentation methods affect hippocampal radiomic feature accuracy in Alzheimer's disease analysis? *Eur. Radiol.* **2022**, *32*, 6965–6976. [[CrossRef](#)]
76. Aly, M.F.A.; Kleijn, S.A.; van Lenthe, J.H.; Menken-Negroiu, R.F.; Robbers, L.F.; Beek, A.M.; Kamp, O. Prediction of prognosis in patients with left ventricular dysfunction using three-dimensional strain echocardiography and cardiac magnetic resonance imaging. *Neth. Heart J.* **2022**, *30*, 572–579. [[CrossRef](#)] [[PubMed](#)]

77. 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. [[CrossRef](#)]
78. 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. [[CrossRef](#)]
79. 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. [[CrossRef](#)]
80. 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. [[CrossRef](#)]
81. 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. [[CrossRef](#)] [[PubMed](#)]
82. 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. [[CrossRef](#)]
83. 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. [[CrossRef](#)] [[PubMed](#)]
84. 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. [[CrossRef](#)]
85. 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. [[CrossRef](#)] [[PubMed](#)]
86. Gopinadhan, A.; Prasanna, G.; Anbarasu, S. AD-EHS: Alzheimer's disease severity detection using efficient hybrid image segmentation. *Adv. Eng. Softw.* **2022**, *173*, 103234. [[CrossRef](#)]
87. Krell-Roesch, J.; Rakusa, M.; Syrjanen, J.A.; van Harten, A.C.; Lowe, V.J.; Jack, C.R.; Kremers, W.K.; Knopman, D.S.; Stokin, G.B.; Petersen, R.C.; et al. Association between CSF biomarkers of Alzheimer's disease and neuropsychiatric symptoms: Mayo Clinic Study of Aging. *Alzheimer's Dement.* **2022**, 1–9. [[CrossRef](#)]
88. Mol, M.O.; van der Lee, S.J.; Hulsman, M.; Pijnenburg, Y.A.L.; Scheltens, P.; Seelaar, H.; van Swieten, J.C.; Kaat, L.D.; Holstege, H.; van Rooij, J.G.J.; et al. Mapping the genetic landscape of early-onset Alzheimer's disease in a cohort of 36 families. *Alzheimer's Res. Ther.* **2022**, *14*, 1–14. [[CrossRef](#)]
89. Wen, J.; Thibeau-Sutre, E.; Diaz-Melo, M.; Samper-González, J.; Routier, A.; Bottani, S.; Dormont, D.; Durrleman, S.; Burgos, N.; Colliot, O. Convolutional Neural Networks for Classification of Alzheimer's Disease: Overview and Reproducible Evaluation. *Med. Image Anal.* **2020**, *63*, 101694. [[CrossRef](#)]
90. Samper-Gonzalez, J.; Burgos, N.; Bottani, S.; Fontanella, S.; Lu, P.; Marcoux, A.; Routier, A.; Guillon, J.; Bacci, M.; Wen, J. Reproducible evaluation of classification methods in Alzheimer's disease: Framework and application to MRI and PET data. *NeuroImage* **2018**, *183*, 504–521. [[CrossRef](#)]
91. Fraternali, P.; Milani, F.; Torres, R.N.; Zangrando, N. Black-box error diagnosis in Deep Neural Networks for computer vision: A survey of tools. *Neural Comput. Appl.* **2022**, *35*, 3041–3062. [[CrossRef](#)]
92. Garnier, R.; Langhendries, R. Concentration inequalities for non-causal random fields. *Electron. J. Stat.* **2022**, *16*, 1681–1725. [[CrossRef](#)]
93. Adali, T.; Calhoun, V.D. Reproducibility and replicability in neuroimaging data analysis. *Curr. Opin. Neurol.* **2022**, *35*, 475–481. [[CrossRef](#)] [[PubMed](#)]
94. Medeiros, G.C.; Twose, C.; Weller, A.; Dougherty, J.W.; Goes, F.S.; Sair, H.I.; Smith, G.S.; Roy, D. Neuroimaging Correlates of Depression after Traumatic Brain Injury: A Systematic Review. *J. Neurotrauma* **2022**, *39*, 755–772. [[CrossRef](#)] [[PubMed](#)]
95. Rathore, S.; Habes, M.; Iftikhar, M.A.; Shacklett, A.; Davatzikos, C. A review on neuroimaging-based classification studies and associated feature extraction methods for Alzheimer's disease and its prodromal stages. *NeuroImage* **2017**, *155*, 530–548. [[CrossRef](#)] [[PubMed](#)]
96. Ebrahimiaghnavieh, M.A.; Luo, S.; Chiong, R. Deep learning to detect Alzheimer's disease from neuroimaging: A systematic literature review. *Comput. Methods Programs Biomed.* **2020**, *187*, 105242. [[CrossRef](#)]
97. Fernando, K.R.M.; Tsokos, C.P. Deep and statistical learning in biomedical imaging: State of the art in 3D MRI brain tumor segmentation. *Inf. Fusion* **2023**, *92*, 450–465. [[CrossRef](#)]
98. Du, B.; Cheng, X.; Duan, Y.; Ning, H. fMRI Brain Decoding and Its Applications in Brain-Computer Interface: A Survey. *Brain Sci.* **2022**, *12*, 228. [[CrossRef](#)]
99. Patel, B.; Irwin, D.J.; Kaufer, D.; Boeve, B.F.; Taylor, A.; Armstrong, M.J. Outcome Measures for Dementia with Lewy Body Clinical Trials A Review. *Alzheimer Dis. Assoc. Disord.* **2022**, *36*, 64–72. [[CrossRef](#)]
100. Zhang, T.J.; Sui, Y.X.; Lu, Q.; Xu, X.J.; Zhu, Y.; Dai, W.J.; Shen, Y.; Wang, T. Effects of rTMS treatment on global cognitive function in Alzheimer's disease: A systematic review and meta-analysis. *Front. Aging Neurosci.* **2022**, *14*, 984708. [[CrossRef](#)]

101. Skinner, J.; Initiative, F.T.A.D.N.; Carvalho, J.O.; Potter, G.G.; Thames, A.D.; Zelinski, E.M.; Crane, P.; Gibbons, L.E. The Alzheimer's Disease Assessment Scale-Cognitive-Plus (ADAS-Cog-Plus): An expansion of the ADAS-Cog to improve responsiveness in MCI. *Brain Imaging Behav.* **2012**, *6*, 489–501. [[CrossRef](#)]
102. Vyhnalek, M.; Jester, D.J.; Andel, R.; Horakova, H.; Nikolai, T.; Laczó, J.; Matuskova, V.; Cechova, K.; Sheardova, K.; Hort, J. Contribution of Memory Tests to Early Identification of Conversion from Amnesic Mild Cognitive Impairment to Dementia. *J. Alzheimer's Dis.* **2022**, *88*, 1397–1409. [[CrossRef](#)] [[PubMed](#)]
103. Abikoff, H.; Alvir, J.; Hong, G.; Sukoff, R.; Orazio, J.; Solomon, S.; Saravay, S. Logical memory subtest of the wechsler memory scale: Age and education norms and alternate-form reliability of two scoring systems. *J. Clin. Exp. Neuropsychol.* **1987**, *9*, 435–448. [[CrossRef](#)]
104. Mills, S.J.; Mackintosh, S.; McDonnell, M.N.; Thewlis, D. Improvement in postural alignment is associated with recovery of mobility after complex acquired brain injury: An observational study. *Physiother. Theory Pract.* **2022**, *39*, 1274–1286. [[CrossRef](#)]
105. Costa, L.; Gago, M.F.; Yelshyna, D.; Ferreira, J.; Silva, H.D.; Rocha, L.; Sousa, N.; Bicho, E. Application of Machine Learning in Postural Control Kinematics for the Diagnosis of Alzheimer's Disease. *Comput. Intell. Neurosci.* **2016**, *2016*, 1–15. [[CrossRef](#)]
106. Gannouni, S.; Aledaily, A.; Belwafi, K.; Aboalsamh, H. Electroencephalography based emotion detection using ensemble classification and asymmetric brain activity. *J. Affect. Disord.* **2022**, *319*, 416–427. [[CrossRef](#)]
107. Morabito, F.C.; Campolo, M.; Ieracitano, C.; Ebadi, J.M.; Bonanno, L.; Bramanti, A.; Desalvo, S.; Mammone, N.; Bramanti, P. Deep convolutional neural networks for classification of mild cognitive impaired and Alzheimer's disease patients from scalp EEG recordings. In Proceedings of the 2016 IEEE 2nd International Forum on Research and Technologies for Society and Industry Leveraging a Better Tomorrow (RTSI), Bologna, Italy, 7–9 September 2016.
108. Anyaiwe, D.E.; Wilson, G.D.; Geddes, T.J.; Singh, G.B. Harnessing mass spectra data using KNN principle: Diagnosing Alzheimer's disease. *ACM SIGBioinformatics Rec.* **2018**, *7*, 1–7. [[CrossRef](#)]
109. Wisely, C.E.; Wang, D.; Henao, R.; Grewal, D.S.; Yoon, S.P.; Polascik, B.; Thompson, A.C.; Burke, J.R.; Carin, L.; Fekrat, S. Deep learning algorithm for diagnosis of Alzheimer's disease using multimodal retinal imaging. *Investig. Ophthalmol. Vis. Sci.* **2019**, *60*, 1461.
110. Landi, I.; Glicksberg, B.S.; Lee, H.-C.; Cherng, S.; Landi, G.; Danieleto, M.; Dudley, J.T.; Furlanello, C.; Miotto, R. Deep representation learning of electronic health records to unlock patient stratification at scale. *NPJ Digit. Med.* **2020**, *3*, 96. [[CrossRef](#)]
111. Park, J.H.; Cho, H.E.; Kim, J.H.; Wall, M.M.; Stern, Y.; Lim, H.; Yoo, S.; Kim, H.S.; Cha, J. Machine learning prediction of incidence of Alzheimer's disease using large-scale administrative health data. *NPJ Digit. Med.* **2020**, *3*, 46. [[CrossRef](#)]
112. 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. [[CrossRef](#)]
113. Chien, Y.-W.; Hong, S.-Y.; Cheah, W.-T.; Yao, L.-H.; Chang, Y.-L.; Fu, L.-C. An Automatic Assessment System for Alzheimer's Disease Based on Speech Using Feature Sequence Generator and Recurrent Neural Network. *Sci. Rep.* **2019**, *9*, 19597. [[CrossRef](#)]
114. Lam, K.-Y.; Tsang, N.W.-H.; Han, S.; Zhang, W.; Ng, J.K.-Y.; Nath, A. Activity tracking and monitoring of patients with Alzheimer's disease. *Multimed. Tools Appl.* **2015**, *76*, 489–521. [[CrossRef](#)]
115. Toosizadeh, N.; Ehsani, H.; Wendel, C.; Zamrini, E.; Connor, K.O.; Mohler, J. Screening older adults for amnesic mild cognitive impairment and early-stage Alzheimer's disease using upper-extremity dual-tasking. *Sci. Rep.* **2019**, *9*, 10911. [[CrossRef](#)] [[PubMed](#)]
116. Haque, R.U.; Pongos, A.L.; Manzanares, C.M.; Lah, J.J.; Levey, A.I.; Clifford, G.D. Deep Convolutional Neural Networks and Transfer Learning for Measuring Cognitive Impairment Using Eye-Tracking in a Distributed Tablet-Based Environment. *IEEE Trans. Biomed. Eng.* **2020**, *68*, 11–18. [[CrossRef](#)]
117. Farina, F.; Emek-Savas, D.; Rueda-Delgado, L.; Boyle, R.; Kiiski, H.; Yener, G.; Whelan, R. A comparison of resting state EEG and structural MRI for classifying Alzheimer's disease and mild cognitive impairment. *Neuroimage* **2020**, *215*, 116795. [[CrossRef](#)] [[PubMed](#)]
118. Ashford, M.T.; Raman, R.; Miller, G.; Donohue, M.C.; Okonkwo, O.C.; Mindt, M.R.; Nosheny, R.L.; Coker, G.A.; Petersen, R.C.; Aisen, P.S.; et al. Screening and enrollment of underrepresented ethnocultural and educational populations in the Alzheimer's Disease Neuroimaging Initiative (ADNI). *Alzheimer's Dement.* **2022**, *18*, 2603–2613. [[CrossRef](#)]
119. Nanayakkara, N.D.; Arnott, S.R.; Scott, C.J.; Solovey, I.; Liang, S.; Fonov, V.S.; Gee, T.; Broberg, D.N.; Haddad, S.M.; Ramirez, J.; et al. Increased brain volumetric measurement precision from multi-site 3D T1-weighted 3 T magnetic resonance imaging by correcting geometric distortions. *Magn. Reson. Imaging* **2022**, *92*, 150–160. [[CrossRef](#)]
120. Weiner, M.W.; Aisen, P.S.; Jack, C.R., Jr.; Jagust, W.J.; Trojanowski, J.Q.; Shaw, L.; Saykin, A.J.; Morris, J.C.; Cairns, N.; Beckett, L.A. The Alzheimer's disease neuroimaging initiative: Progress report and future plans. *Alzheimer's Dement.* **2010**, *6*, 202–211.e7. [[CrossRef](#)] [[PubMed](#)]
121. Weiner, M.W.; Veitch, D.P.; Aisen, P.S.; Beckett, L.A.; Cairns, N.J.; Cedarbaum, J.; Donohue, M.C.; Green, R.C.; Harvey, D.; Jack, C.R., Jr. Impact of the Alzheimer's disease neuroimaging initiative, 2004 to 2014. *Alzheimer's Dement.* **2015**, *11*, 865–884. [[CrossRef](#)]
122. Weiner, M.W.; Veitch, D.P.; Aisen, P.S.; Beckett, L.A.; Cairns, N.J.; Green, R.C.; Harvey, D.; Jack, C.R., Jr.; Jagust, W.; Morris, J.C. The Alzheimer's Disease Neuroimaging Initiative 3: Continued innovation for clinical trial improvement. *Alzheimer's Dement.* **2017**, *13*, 561–571. [[CrossRef](#)]

123. LaMontagne, P.J.; Benzinger, T.L.; Morris, J.C.; Keefe, S.; Hornbeck, R.; Xiong, C.; Grant, E.; Hassenstab, J.; Moulder, K.; Vlassenko, A.G.; et al. OASIS-3: Longitudinal neuroimaging, clinical, and cognitive dataset for normal aging and Alzheimer disease. *medRxiv* **2019**, *12*, 19014902.
124. Dagley, A.; LaPoint, M.; Huijbers, W.; Hedden, T.; McLaren, D.G.; Chatwal, J.P.; Papp, K.V.; Amariglio, R.E.; Blacker, D.; Rentz, D.M.; et al. Harvard Aging Brain Study: Dataset and accessibility. *Neuroimage* **2017**, *144*, 255–258. [[CrossRef](#)]
125. Malone, I.; Cash, D.; Ridgway, G.; MacManus, D.; Ourselin, S.; Fox, N.; Schott, J. MIRIAD (Minimal Interval Resonance Imaging in Alzheimer’s Disease). *NeuroImage* **2013**, *70*, 33–36. [[CrossRef](#)] [[PubMed](#)]
126. Iwatsubo, T. Japanese Alzheimer’s Disease Neuroimaging Initiative: Present status and future. *Alzheimer’s Dement.* **2010**, *6*, 297–299. [[CrossRef](#)]
127. Sun, W.; Wu, Q.; Chen, H.; Yu, L.; Yin, J.; Liu, F.; Tian, R.; Song, B.; Qu, B.; Xing, M.; et al. A Validation Study of the Hong Kong Brief Cognitive Test for Screening Patients with Mild Cognitive Impairment and Alzheimer’s Disease. *J. Alzheimer’s Dis.* **2022**, *88*, 1523–1532. [[CrossRef](#)]
128. Ellis, K.A.; Bush, A.I.; Darby, D.; De Fazio, D.; Foster, J.; Hudson, P.; Lautenschlager, N.T.; Lenzo, N.; Martins, R.N.; Maruff, P. The Australian Imaging, Biomarkers and Lifestyle (AIBL) study of aging: Methodology and baseline characteristics of 1112 individuals recruited for a longitudinal study of Alzheimer’s disease. *Int. Psychogeriatr.* **2009**, *21*, 672–687. [[CrossRef](#)] [[PubMed](#)]
129. Nigri, A.; Ferraro, S.; Wheeler-Kingshott, C.A.M.G.; Tosetti, M.; Redolfi, A.; Forloni, G.; D’Angelo, E.; Aquino, D.; Biagi, L.; Bosco, P.; et al. Quantitative MRI Harmonization to Maximize Clinical Impact: The RIN-Neuroimaging Network. *Front. Neurol.* **2022**, *13*, 855125. [[CrossRef](#)]
130. Redolfi, A.; McClatchey, R.; Anjum, A.; Zijdenbos, A.; Manset, D.; Barkhof, F.; Spenger, C.; Legré, Y.; Wahlund, L.-O.; Pietro, C.B.d.S.; et al. Grid infrastructures for computational neuroscience: The neuGRID example. *Futur. Neurol.* **2009**, *4*, 703–722. [[CrossRef](#)]
131. Toga, A.W.; Neu, S.C.; Bhatt, P.; Crawford, K.L.; Ashish, N. The global Alzheimer’s association interactive network. *Alzheimer’s Dement.* **2016**, *12*, 49–54. [[CrossRef](#)]
132. Bron, E.E.; Smits, M.; van der Flier, W.M.; Vrenken, H.; Barkhof, F.; Scheltens, P.; Papma, J.M.; Steketee, R.M.; Orellana, C.M.; Meijboom, R.; et al. Standardized evaluation of algorithms for computer-aided diagnosis of dementia based on structural MRI: The CADDementia challenge. *Neuroimage* **2015**, *111*, 562–579. [[CrossRef](#)]
133. Hernandez, M.; Ramon-Julvez, U.; Ferraz, F. With the ADNI Consortium Explainable AI toward understanding the performance of the top three TADPOLE Challenge methods in the forecast of Alzheimer’s disease diagnosis. *PLoS ONE* **2022**, *17*, e0264695. [[CrossRef](#)]
134. Marinescu, R.V.; Oxtoby, N.P.; Young, A.L.; Bron, E.E.; Toga, A.W.; Weiner, M.W.; Barkhof, F.; Fox, N.C.; Klein, S.; Alexander, D.C. Tadpole challenge: Prediction of longitudinal evolution in Alzheimer’s disease. *arXiv* **2018**, arXiv:1805.03909.
135. Allen, G.I.; Amoroso, N.; Anghel, C.; Balagurusamy, V.; Bare, C.J.; Beaton, D.; Bellotti, R.; Bennett, D.A.; Boehme, K.L.; Boutros, P.C. Crowdsourced estimation of cognitive decline and resilience in Alzheimer’s disease. *Alzheimer’s Dement.* **2016**, *12*, 645–653. [[CrossRef](#)]
136. El-Gazzar, A.; Thomas, R.M.; van Wingen, G. Dynamic Adaptive Spatio-Temporal Graph Convolution for fMRI Modelling. In *Machine Learning in Clinical Neuroimaging. Proceedings of the 4th International Workshop, MLCN 2021, Held in Conjunction with MICCAI 2021, Strasbourg, France, 27 September 2021*; Proceedings 4; Springer International Publishing: Berlin/Heidelberg, Germany, 2021; pp. 125–134. [[CrossRef](#)]
137. Varzandian, A.; Razo, M.A.S.; Sanders, M.R.; Atmakuru, A.; Di Fatta, G. Classification-Biased Apparent Brain Age for the Prediction of Alzheimer’s Disease. *Front. Neurosci.* **2021**, *15*, 673120. [[CrossRef](#)]
138. Fu, Y.; Huang, Y.; Wang, Y.; Dong, S.; Xue, L.; Yin, X.; Yang, Q.; Shi, Y.; Zhuo, C. OTFPF: Optimal Transport-Based Feature Pyramid Fusion Network for Brain Age Estimation with 3D Overlapped ConvNeXt. *arXiv* **2022**, arXiv:2205.04684.
139. Bycroft, C.; Freeman, C.; Petkova, D.; Band, G.; Elliott, L.T.; Sharp, K.; Motyer, A.; Vukcevic, D.; Delaneau, O.; O’connell, J.; et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature* **2018**, *562*, 203–209. [[CrossRef](#)] [[PubMed](#)]
140. Huang, K.-l.; Marcora, E.; Pimenova, A.A.; Di Narzo, A.F.; Kapoor, M.; Jin, S.C.; Harari, O.; Bertelsen, S.; Fairfax, B.P.; Czajkowski, J. A common haplotype lowers PU. 1 expression in myeloid cells and delays onset of Alzheimer’s disease. *Nat. Neurosci.* **2017**, *20*, 1052–1061. [[CrossRef](#)] [[PubMed](#)]
141. Zhu, X.; Luchetti, M.; Aschwanden, D.; Sesker, A.A.; Stephan, Y.; Sutin, A.R.; Terracciano, A. Satisfaction With Life and Risk of Dementia: Findings From the Korean Longitudinal Study of Aging. *J. Gerontol. Ser. B* **2022**, *77*, 1831–1840. [[CrossRef](#)]
142. Suh, S.; Han, J.; Oh, S.; Kim, K. Impact of sleep on future cognition in non-demented elderly: Results from the korean longitudinal study on cognitive aging and dementia (kloscad). *J. Neurol. Sci.* **2017**, *381*, 182. [[CrossRef](#)]
143. Sakr, F.A.; Grothe, M.J.; Cavedo, E.; Jelistratova, I.; Habert, M.-O.; Dyrba, M.; Gonzalez-Escamilla, G.; Bertin, H.; Locatelli, M.; Lehericy, S.; et al. Applicability of in vivo staging of regional amyloid burden in a cognitively normal cohort with subjective memory complaints: The INSIGHT-preAD study. *Alzheimer’s Res. Ther.* **2019**, *11*, 15. [[CrossRef](#)]
144. Dubois, B.; Epelbaum, S.; Nyasse, F.; Bakardjian, H.; Gagliardi, G.; Uspenskaya, O.; Houot, M.; Lista, S.; Cacciamani, F.; Potier, M.-C. Cognitive and neuroimaging features and brain β -amyloidosis in individuals at risk of Alzheimer’s disease (INSIGHT-preAD): A longitudinal observational study. *Lancet Neurol.* **2018**, *17*, 335–346. [[CrossRef](#)]

145. Wilkins, C.H.; Windon, C.C.; Dilworth-Anderson, P.; Romanoff, J.; Gatsonis, C.; Hanna, L.; Apgar, C.; Gareen, I.F.; Hill, C.V.; Hillner, B.E. Racial and Ethnic Differences in Amyloid PET Positivity in Individuals with Mild Cognitive Impairment or Dementia: A Secondary Analysis of the Imaging Dementia—Evidence for Amyloid Scanning (IDEAS) Cohort Study. *JAMA Neurol.* **2022**, *79*, 1139–1147. [[CrossRef](#)]
146. Silva, T.C.; Zhang, W.; Young, J.I.; Gomez, L.; Schmidt, M.A.; Varma, A.; Chen, X.S.; Martin, E.R.; Wang, L. Distinct sex-specific DNA methylation differences in Alzheimer’s disease. *Alzheimer’s Res. Ther.* **2022**, *14*, 1–21. [[CrossRef](#)]
147. Lovestone, S.; Francis, P.; Kloszewska, I.; Mecocci, P.; Simmons, A.; Soininen, H.; Spenger, C.; Tsolaki, M.; Vellas, B.; Wahlund, L.O.; et al. AddNeuroMed—The European collaboration for the discovery of novel biomarkers for Alzheimer’s disease. *Ann. N. Y. Acad. Sci.* **2009**, *1180*, 36–46. [[CrossRef](#)]
148. Chen, S.; Stromer, D.; Alabdalrahim, H.A.; Schwab, S.; Weih, M.; Maier, A. Automatic dementia screening and scoring by applying deep learning on clock-drawing tests. *Sci. Rep.* **2020**, *10*, 1–11. [[CrossRef](#)]
149. Gorgolewski, K.J.; Auer, T.; Calhoun, V.D.; Craddock, R.C.; Das, S.; Duff, E.P.; Flandin, G.; Ghosh, S.S.; Glatard, T.; Halchenko, Y.O.; et al. The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Sci. Data* **2016**, *3*, 1–9. [[CrossRef](#)]
150. Hu, Z.; Wang, Z.; Jin, Y.; Hou, W. VGG-TSwinformer: Transformer-based deep learning model for early Alzheimer’s disease prediction. *Comput. Methods Programs Biomed.* **2023**, *229*, 107291. [[CrossRef](#)]
151. Houria, L.; Belkhamas, N.; Cherfa, A.; Cherfa, Y. Multi-modality MRI for Alzheimer’s disease detection using deep learning. *Phys. Eng. Sci. Med.* **2022**, *45*, 1043–1053. [[CrossRef](#)]
152. Pan, D.; Zeng, A.; Yang, B.Y.; Lai, G.Y.; Hu, B.; Song, X.W.; Jiang, T.Z. Deep Learning for Brain MRI Confirms Patterned Pathological Progression in Alzheimer’s Disease. *Adv. Sci.* **2022**, *10*, 2204717. [[CrossRef](#)]
153. Jindal, S.K.; Banerjee, S.; Patra, R.; Paul, A. Deep learning-based brain malignant neoplasm classification using MRI image segmentation assisted by bias field correction and histogram equalization. In *Brain Tumor MRI Image Segmentation Using Deep Learning Techniques*; Elsevier: Amsterdam, The Netherlands, 2022; pp. 135–161.
154. Gispert, J.D.; Reig, S.; Pascau, J.; Vaquero, J.J.; García-Barreno, P.; Desco, M. Method for bias field correction of brain T1-weighted magnetic resonance images minimizing segmentation error. *Hum. Brain Mapp.* **2004**, *22*, 133–144. [[CrossRef](#)]
155. Wu, L.; He, T.; Yu, J.; Liu, H.; Zhang, S.; Zhang, T. Volume and surface coil simultaneous reception (VSSR) method for intensity inhomogeneity correction in MRI. *Technol. Health Care* **2022**, *30*, 827–838. [[CrossRef](#)]
156. Sled, J.G.; Zijdenbos, A.P.; Evans, A.C. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans. Med. Imaging* **1998**, *17*, 87–97. [[CrossRef](#)] [[PubMed](#)]
157. Ji, H.; Liu, Z.; Yan, W.Q.; Klette, R. Early diagnosis of Alzheimer’s disease using deep learning. In Proceedings of the 2nd International Conference on Control and Computer Vision, Jeju, Republic of Korea, 15–18 June 2019.
158. Bhattacharjee, R.; Heitz, F.; Noblet, V.; Sharma, S.; Sharma, N. Evaluation of a Learning-based Deformable Registration Method on Abdominal CT Images. *IRBM* **2020**, *42*, 94–105. [[CrossRef](#)]
159. Andersson, J.L.; Jenkinson, M.; Smith, S. *Non-Linear Registration Aka Spatial Normalisation FMRIB Technial Report TR07JA2*; FMRIB Analysis Group of the University of Oxford: Oxford, UK, 2007; pp. 1–22.
160. Jenkinson, M.; Smith, S. A global optimisation method for robust affine registration of brain images. *Med. Image Anal.* **2001**, *5*, 143–156. [[CrossRef](#)] [[PubMed](#)]
161. Mazziotta, J.; Toga, A.; Evans, A.; Fox, P.; Lancaster, J.; Zilles, K.; Woods, R.; Paus, T.; Simpson, G.; Pike, B. A four-dimensional probabilistic atlas of the human brain. *J. Am. Med. Inform. Assoc.* **2001**, *8*, 401–430. [[CrossRef](#)]
162. Ramon-Julvez, U.; Hernandez, M.; Mayordomo, E. Adni Analysis of the Influence of Diffeomorphic Normalization in the Prediction of Stable VS Progressive MCI Conversion with Convolutional Neural Networks. In Proceedings of the 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), Iowa City, IA, USA, 3–7 April 2020. [[CrossRef](#)]
163. Chen, C.-L.; Hsu, Y.-C.; Yang, L.-Y.; Tung, Y.-H.; Luo, W.-B.; Liu, C.-M.; Hwang, T.-J.; Hwu, H.-G.; Tseng, W.-Y.I. Generalization of diffusion magnetic resonance imaging-based brain age prediction model through transfer learning. *Neuroimage* **2020**, *217*, 116831. [[CrossRef](#)]
164. Ahmed, S.; Choi, K.Y.; Lee, J.J.; Kim, B.C.; Kwon, G.-R.; Lee, K.H.; Jung, H.Y. Ensembles of Patch-Based Classifiers for Diagnosis of Alzheimer Diseases. *IEEE Access* **2019**, *7*, 73373–73383. [[CrossRef](#)]
165. Patenaude, B.; Smith, S.M.; Kennedy, D.N.; Jenkinson, M. A Bayesian model of shape and appearance for subcortical brain segmentation. *NeuroImage* **2011**, *56*, 907–922. [[CrossRef](#)]
166. 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. [[CrossRef](#)]
167. Lin, W.; Tong, T.; Gao, Q.; Guo, D.; Du, X.; Yang, Y.; Guo, G.; Xiao, M.; Du, M.; Qu, X.; et al. Convolutional Neural Networks-Based MRI Image Analysis for the Alzheimer’s Disease Prediction from Mild Cognitive Impairment. *Front. Neurosci.* **2018**, *12*, 777. [[CrossRef](#)]
168. Basher, A.; Choi, K.Y.; Lee, J.J.; Lee, B.; Kim, B.C.; Lee, K.H.; Jung, H.Y. Hippocampus Localization Using a Two-Stage Ensemble Hough Convolutional Neural Network. *IEEE Access* **2019**, *7*, 73436–73447. [[CrossRef](#)]
169. Liu, M.; Li, F.; Yan, H.; Wang, K.; Ma, Y.; Shen, L.; Xu, M. A multi-model deep convolutional neural network for automatic hippocampus segmentation and classification in Alzheimer’s disease. *Neuroimage* **2020**, *208*, 116459. [[CrossRef](#)]

170. 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. [[CrossRef](#)]
171. Prem Kumar, A.; Singh, N.; Nair, D.; Justin, A. Neuronal PET tracers for Alzheimer's disease. *Biochem. Biophys. Res. Commun.* **2022**, *587*, 58–62. [[CrossRef](#)]
172. Zhou, D.A.; Xu, K.; Zhao, X.B.; Chen, Q.; Sang, F.; Fan, D.; Su, L.; Zhang, Z.J.; Ai, L.; Chen, Y.J. Spatial Distribution and Hierarchical Clustering of beta-Amyloid and Glucose Metabolism in Alzheimer's Disease. *Front. Aging Neurosci.* **2022**, *14*, 788567. [[CrossRef](#)]
173. Tanner, J.A.; Iaccarino, L.; Edwards, L.; Asken, B.M.; Gorno-Tempini, M.L.; Kramer, J.H.; Pham, J.; Perry, D.C.; Possin, K.; Malpetti, M.; et al. Amyloid, tau and metabolic PET correlates of cognition in early and late-onset Alzheimer's disease. *Brain* **2022**, *145*, 4489–4505. [[CrossRef](#)]
174. Lagarde, J.; Olivieri, P.; Tonietto, M.; Tissot, C.; Rivals, I.; Gervais, P.; Caillé, F.; Moussion, M.; Bottlaender, M.; Sarazin, M. Tau-PET imaging predicts cognitive decline and brain atrophy progression in early Alzheimer's disease. *J. Neurol. Neurosurg. Psychiatry* **2022**, *93*, 459–467. [[CrossRef](#)]
175. Ding, Y.; Sohn, J.H.; Kawczynski, M.G.; Trivedi, H.; Harnish, R.; Jenkins, N.W.; Lituiev, D.; Copeland, T.P.; Aboian, M.S.; Aparici, C.M.; et al. A Deep Learning Model to Predict a Diagnosis of Alzheimer Disease by Using ¹⁸F-FDG PET of the Brain. *Radiology* **2019**, *290*, 456–464. [[CrossRef](#)]
176. Hwang, S.J.; Tao, Z.; Singh, V.; Kim, W.H. Onditional recurrent flow: Conditional generation of longitudinal samples with applications to neuroimaging. In Proceedings of the IEEE International Conference on Computer Vision, Seoul, Republic of Korea, 27–28 October 2019. [[CrossRef](#)]
177. 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. [[CrossRef](#)]
178. Palmer, W.C.; Park, S.M.; Levendovszky, S.R. Brain state transition analysis using ultra-fast fMRI differentiates MCI from cognitively normal controls. *Front. Neurosci.* **2022**, *16*, 1531. [[CrossRef](#)]
179. Tondelli, M.; Benuzzi, F.; Ballotta, D.; Molinari, M.A.; Chiari, A.; Zamboni, G. Eliciting Implicit Awareness in Alzheimer's Disease and Mild Cognitive Impairment: A Task-Based Functional MRI Study. *Front. Aging Neurosci.* **2022**, *14*, 816648. [[CrossRef](#)] [[PubMed](#)]
180. Han, X.-M.; Gu, X.-Q.; Liu, Y.; Gu, J.-B.; Li, L.-F.; Fu, L.-L. Correlations between hippocampal functional connectivity, structural changes, and clinical data in patients with relapsing-remitting multiple sclerosis: A case-control study using multimodal magnetic resonance imaging. *Neural Regen. Res.* **2022**, *17*, 1115. [[CrossRef](#)] [[PubMed](#)]
181. Miao, D.W.; Zhou, X.G.; Wu, X.Y.; Chen, C.D.; Tian, L. Distinct profiles of functional connectivity density aberrance in Alzheimer's disease and mild cognitive impairment. *Front. Psychiatry* **2022**, *13*, 1079149. [[CrossRef](#)] [[PubMed](#)]
182. Luo, J.; Agboola, F.; Grant, E.; Morris, J.C.; Masters, C.L.; Albert, M.S.; Johnson, S.C.; McDade, E.M.; Fagan, A.M.; Benzinger, T.L.S.; et al. Accelerated longitudinal changes and ordering of Alzheimer disease biomarkers across the adult lifespan. *Brain* **2022**, *145*, 4459–4473. [[CrossRef](#)] [[PubMed](#)]
183. Sarraf, S.; Desouza, D.D.; Anderson, J.A.E.; Saverino, C. MCADNet: Recognizing Stages of Cognitive Impairment Through Efficient Convolutional fMRI and MRI Neural Network Topology Models. *IEEE Access* **2019**, *7*, 155584–155600. [[CrossRef](#)]
184. 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. [[CrossRef](#)]
185. Bi, X.; Zhao, X.; Huang, H.; Chen, D.; Ma, Y. Functional Brain Network Classification for Alzheimer's Disease Detection with Deep Features and Extreme Learning Machine. *Cogn. Comput.* **2019**, *12*, 513–527. [[CrossRef](#)]
186. 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. [[CrossRef](#)]
187. Cui, R.; Liu, M. Hippocampus Analysis by Combination of 3-D DenseNet and Shapes for Alzheimer's Disease Diagnosis. *IEEE J. Biomed. Health Inform.* **2018**, *23*, 2099–2107. [[CrossRef](#)]
188. Jung, W.; Mulyadi, A.; Suk, H.-I. Unified modeling of imputation, forecasting, and prediction for ad 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.
189. 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.
190. Huang, H.; Zheng, S.; Yang, Z.; Wu, Y.; Li, Y.; Qiu, J.; Cheng, Y.; Lin, P.; Guan, J.; Mikulis, D.J.; et al. Voxel-based morphometry and a deep learning model for the diagnosis of early Alzheimer's disease based on cerebral gray matter changes. *Cereb. Cortex* **2022**, *33*, 754–763. [[CrossRef](#)]
191. Liu, M.; Zhang, J.; Adeli, E.; Shen, D. Joint classification and regression via deep multi-task multi-channel learning for Alzheimer's disease diagnosis. *IEEE Trans. Biomed. Eng.* **2018**, *66*, 1195–1206. [[CrossRef](#)]
192. Lian, C.; Liu, M.; Zhang, J.; Shen, D. Hierarchical fully convolutional network for joint atrophy localization and Alzheimer's Disease diagnosis using structural MRI. *IEEE Trans. Pattern Anal. Mach. Intell.* **2018**, *42*, 880–893. [[CrossRef](#)]
193. Kabani, N.J.; MacDonald, D.J.; Holmes, C.J.; Evans, A.C. 3D Anatomical Atlas of the Human Brain. *Neuroimage* **1998**, *7*, S717. [[CrossRef](#)]

194. Sydnor, V.J.; Cieslak, M.; Duprat, R.; Deluisi, J.; Flounders, M.W.; Long, H.; Scully, M.; Balderston, N.L.; Sheline, Y.I.; Bassett, D.S.; et al. Cortical-subcortical structural connections support transcranial magnetic stimulation engagement of the amygdala. *Sci. Adv.* **2022**, *8*, eabn5803. [[CrossRef](#)]
195. Du, Y.; Yang, W.; Zhang, J.; Liu, J. Changes in ALFF and ReHo values in methamphetamine abstinent individuals based on the Harvard-Oxford atlas: A longitudinal resting-state fMRI study. *Addict. Biol.* **2021**, *27*, e13080. [[CrossRef](#)]
196. Sengupta, D.; Gupta, P.; Biswas, A. A survey on mutual information based medical image registration algorithms. *Neurocomputing* **2021**, *486*, 174–188. [[CrossRef](#)]
197. Shen, D.; Davatzikos, C. HAMMER: Hierarchical attribute matching mechanism for elastic registration. *IEEE Trans. Med. Imaging* **2002**, *21*, 1421–1439. [[CrossRef](#)]
198. Liu, S.; Liu, S.; Cai, W.; Pujol, S.; Kikinis, R.; Feng, D. Early diagnosis of Alzheimer’s disease with deep learning. In Proceedings of the 2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI), Beijing, China, 29 April–2 May 2014.
199. Li, T.; Hoogman, M.; Mota, N.R.; Buitelaar, J.K.; Vasquez, A.A.; Franke, B.; Rooij, D.; The ENIGMA-ASD Working Group. Dissecting the heterogeneous subcortical brain volume of autism spectrum disorder using community detection. *Autism Res.* **2021**, *15*, 42–55. [[CrossRef](#)]
200. Song, H.; Bharadwaj, P.K.; Raichlen, D.A.; Habeck, C.G.; Huentelman, M.J.; Hishaw, G.A.; Trouard, T.P.; Alexander, G.E. Association of homocysteine-related subcortical brain atrophy with white matter lesion volume and cognition in healthy aging. *Neurobiol. Aging* **2023**, *121*, 129–138. [[CrossRef](#)]
201. Chen, W.; Li, H.; Hou, X.; Jia, X. Gray matter alteration in medication overuse headache: A coordinates-based activation likelihood estimation meta-analysis. *Brain Imaging Behav.* **2022**, *16*, 2307–2319. [[CrossRef](#)]
202. Vercellino, M.; Marasciulo, S.; Grifoni, S.; Vallino-Costassa, E.; Bosa, C.; Pasanisi, M.B.; Crociara, P.; Casalone, C.; Chio, A.; Giordana, M.T.; et al. Acute and chronic synaptic pathology in multiple sclerosis gray matter. *Mult. Scler. J.* **2022**, *28*, 369–382. [[CrossRef](#)] [[PubMed](#)]
203. White, M.F.; Tanabe, S.; Casey, C.; Parker, M.; Bo, A.; Kunkel, D.; Nair, V.; Pearce, R.A.; Lennertz, R.; Prabhakaran, V.; et al. Relationships between preoperative cortical thickness, postoperative electroencephalogram slowing, and postoperative delirium. *Br. J. Anaesth.* **2021**, *127*, 236–244. [[CrossRef](#)] [[PubMed](#)]
204. Demirci, N.; Holland, M.A. Cortical thickness systematically varies with curvature and depth in healthy human brains. *Hum. Brain Mapp.* **2022**, *43*, 2064–2084. [[CrossRef](#)] [[PubMed](#)]
205. Jiang, J.; Sheng, C.; Chen, G.; Liu, C.; Jin, S.; Li, L.; Jiang, X.; Han, Y.; Weiner, M.W.; Aisen, P.; et al. Glucose metabolism patterns: A potential index to characterize brain ageing and predict high conversion risk into cognitive impairment. *Geroscience* **2022**, *44*, 2319–2336. [[CrossRef](#)]
206. Choi, J.H.; Kim, M.-S. Homeostatic Regulation of Glucose Metabolism by the Central Nervous System. *Endocrinol. Metab.* **2022**, *37*, 9–25. [[CrossRef](#)]
207. Rabin, J.S.; Nichols, E.; La Joie, R.; Casaletto, K.B.; Palta, P.; Dams-O’connor, K.; Kumar, R.G.; George, K.M.; Satizabal, C.L.; Schneider, J.A.; et al. Cerebral amyloid angiopathy interacts with neuritic amyloid plaques to promote tau and cognitive decline. *Brain* **2022**, *145*, 2823–2833. [[CrossRef](#)]
208. Saito, S.; Yamashiro, T.; Yamauchi, M.; Yamamoto, Y.; Noguchi, M.; Tomita, T.; Kawakami, D.; Shikata, M.; Tanaka, T.; Ihara, M. Complement 3 Is a Potential Biomarker for Cerebral Amyloid Angiopathy. *J. Alzheimer’s Dis.* **2022**, *89*, 381–387. [[CrossRef](#)]
209. Wang, M.; Cui, B.; Shan, Y.; Yang, H.; Yan, Z.; Sundar, L.K.S.; Alberts, I.; Rominger, A.; Wendler, T.; Shi, K.; et al. Non-Invasive Glucose Metabolism Quantification Method Based on Unilateral ICA Image Derived Input Function by Hybrid PET/MR in Ischemic Cerebrovascular Disease. *IEEE J. Biomed. Health Inform.* **2022**, *26*, 5122–5129. [[CrossRef](#)]
210. Liu, J.; Wang, J.; Tang, Z.; Hu, B.; Wu, F.-X.; Pan, Y. Improving Alzheimer’s disease classification by combining multiple measures. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2017**, *15*, 1649–1659. [[CrossRef](#)]
211. Messina, D.; Borrelli, P.; Russo, P.; Salvatore, M.; Aiello, M. Voxel-Wise Feature Selection Method for CNN Binary Classification of Neuroimaging Data. *Front. Neurosci.* **2021**, *15*, 630747. [[CrossRef](#)]
212. Gerber, S.; Niethammer, M.; Ebrahim, E.; Piven, J.; Dager, S.R.; Styner, M.; Aylward, S.; Enquobahrie, A. Optimal transport features for morphometric population analysis. *Med. Image Anal.* **2023**, *84*, 102696. [[CrossRef](#)]
213. Wu, S.; Zhao, W.; Ji, S. Real-time dynamic simulation for highly accurate spatiotemporal brain deformation from impact. *Comput. Methods Appl. Mech. Eng.* **2022**, *394*, 114913. [[CrossRef](#)]
214. Bao, Z.; Zhang, T.; Pan, T.; Zhang, W.; Zhao, S.; Liu, H.; Nie, B. Automatic method for individual parcellation of manganese-enhanced magnetic resonance imaging of rat brain. *Front. Neurosci.* **2022**, *16*, 954237. [[CrossRef](#)]
215. Zhang, X.; Feng, Y.; Chen, W.; Li, X.; Faria, A.V.; Feng, Q.; Mori, S. Linear Registration of Brain MRI Using Knowledge-Based Multiple Intermediator Libraries. *Front. Neurosci.* **2019**, *13*, 909. [[CrossRef](#)]
216. Dadar, M.; Manera, A.L.; Fonov, V.S.; Ducharme, S.; Collins, D.L. MNI-FTD templates, unbiased average templates of frontotemporal dementia variants. *Sci. Data* **2021**, *8*, 222. [[CrossRef](#)]
217. Giraldo, D.L.; Smith, R.E.; Struyfs, H.; Niemantsverdriet, E.; De Roeck, E.; Bjerke, M.; Engelborghs, S.; Romero, E.; Sijbers, J.; Jeurissen, B. Investigating Tissue-Specific Abnormalities in Alzheimer’s Disease with Multi-Shell Diffusion MRI. *J. Alzheimer’s Dis.* **2022**, *90*, 1771–1791. [[CrossRef](#)]
218. Zhang, X.; Liu, Y.; Zhang, Q.; Yuan, F. Multi-Modality Reconstruction Attention and Difference Enhancement Network for Brain MRI Image Segmentation. *IEEE Access* **2022**, *10*, 31058–31069. [[CrossRef](#)]

219. Jones, D.; Lowe, V.; Graff-Radford, J.; Botha, H.; Barnard, L.; Wiepert, D.; Murphy, M.C.; Murray, M.; Senjem, M.; Gunter, J.; et al. A computational model of neurodegeneration in Alzheimer's disease. *Nat. Commun.* **2022**, *13*, 1643. [[CrossRef](#)]
220. Wang, Z.; Albarghouthi, A.; Prakriya, G.; Jha, S. Interval universal approximation for neural networks. *Proc. ACM Program. Lang.* **2022**, *6*, 1–29. [[CrossRef](#)]
221. Pham, V.T.; Jang, Y.; Park, J.W.; Kim, D.J.; Kim, S.E. Cable damage identification of cable-stayed bridge using multi-layer perceptron and graph neural network. *Steel Compos. Struct.* **2022**, *44*, 227–240.
222. Rosenblatt, F. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychol. Rev.* **1958**, *65*, 386–408. [[CrossRef](#)] [[PubMed](#)]
223. Sabharwal, T.; Gupta, R. Deep facial recognition after medical alterations. *Multimed. Tools Appl.* **2022**, *81*, 25675–25706. [[CrossRef](#)]
224. Fattah, E.A.; Van Niekerk, J.; Rue, H. Smart Gradient—An adaptive technique for improving gradient estimation. *Found. Data Sci.* **2022**, *4*, 123. [[CrossRef](#)]
225. Ojha, V.; Nicosia, G. Backpropagation Neural Tree. *Neural Netw.* **2022**, *149*, 66–83. [[CrossRef](#)] [[PubMed](#)]
226. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Learning representations by back-propagating errors. *Nature* **1986**, *323*, 533–536. [[CrossRef](#)]
227. Goodfellow, I.; Bengio, Y.; Courville, A.; Bengio, Y. *Deep Learning*; MIT Press Cambridge: Cambridge, MA, USA, 2016; Volume 1.
228. Liu, X.; Faes, L.; Kale, A.U.; Wagner, S.K.; Fu, D.J.; Bruynseels, A.; Mahendiran, T.; Moraes, G.; Shamdas, M.; Kern, C.; et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: A systematic review and meta-analysis. *Lancet Digit. Health* **2019**, *1*, e271–e297. [[CrossRef](#)]
229. Dolph, C.V.; Alam, M.; Shboul, Z.; Samad, M.D.; Iftekharuddin, K.M. Deep learning of texture and structural features for multiclass Alzheimer's disease classification. In Proceedings of the 2017 International Joint Conference on Neural Networks (IJCNN), Anchorage, AK, USA, 14–19 May 2017.
230. Liu, S.; Liu, S.; Cai, W.; Che, H.; Pujol, S.; Kikinis, R.; Feng, D.; Fulham, M.J. Multimodal neuroimaging feature learning for multiclass diagnosis of Alzheimer's disease. *IEEE Trans. Biomed. Eng.* **2014**, *62*, 1132–1140. [[CrossRef](#)]
231. Kobayashi, T. Optimistic reinforcement learning by forward Kullback–Leibler divergence optimization. *Neural Netw.* **2022**, *152*, 169–180. [[CrossRef](#)]
232. Ji, S.; Zhang, Z.; Ying, S.; Wang, L.; Zhao, X.; Gao, Y. Kullback–Leibler Divergence Metric Learning. *IEEE Trans. Cybern.* **2020**, *52*, 2047–2058. [[CrossRef](#)]
233. Nair, V.; Hinton, G. 3D object recognition with deep belief nets. In Proceedings of the Advances in Neural Information Processing Systems, NIPS 2009, Vancouver, BC, Canada, 7–10 December 2009.
234. Ju, R.; Hu, C.; Li, Q. Early diagnosis of Alzheimer's disease based on resting-state brain networks and deep learning. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2017**, *16*, 244–257. [[CrossRef](#)]
235. Ithapu, V.K.; Singh, V.; Okonkwo, O.C.; Chappell, R.J.; Dowling, N.M.; Johnson, S.C.; Initiative, A.D.N. Imaging-based enrichment criteria using deep learning algorithms for efficient clinical trials in mild cognitive impairment. *Alzheimer's Dement.* **2015**, *11*, 1489–1499. [[CrossRef](#)]
236. Bhatkoti, P.; Paul, M. Early diagnosis of Alzheimer's disease: A multi-class deep learning framework with modified k-sparse autoencoder classification. In Proceedings of the 2016 International Conference on Image and Vision Computing New Zealand (IVCNZ), Palmerston North, New Zealand, 21–22 November 2016.
237. Oh, K.; Chung, Y.-C.; Kim, K.W.; Kim, W.-S.; Oh, I.-S. Classification and Visualization of Alzheimer's Disease using Volumetric Convolutional Neural Network and Transfer Learning. *Sci. Rep.* **2019**, *9*, 1–16. [[CrossRef](#)]
238. Hosseini-Asl, E.; Keynton, R.; El-Baz, A. Alzheimer's disease diagnostics by adaptation of 3D convolutional network. In Proceedings of the 2016 IEEE International Conference on Image Processing (ICIP), Phoenix, AZ, USA, 25–28 September 2016.
239. Er, F.; Goularas, D. Predicting the Prognosis of MCI Patients Using Longitudinal MRI Data. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2020**, *18*, 1164–1173. [[CrossRef](#)]
240. Suk, H.-I.; Initiative, T.A.D.N.; Lee, S.-W.; Shen, D. Latent feature representation with stacked auto-encoder for AD/MCI diagnosis. *Anat. Embryol.* **2013**, *220*, 841–859. [[CrossRef](#)]
241. Shakeri, M.; Lombaert, H.; Tripathi, S.; Kadoury, S. Deep spectral-based shape features for Alzheimer's disease classification. In Proceedings of the International Workshop on Spectral and Shape Analysis in Medical Imaging, Athens, Greece, 21 October 2016; Springer: Berlin/Heidelberg, Germany, 2016.
242. Jiao, Z.; Ji, Y.; Gao, P.; Wang, S.-H. Extraction and analysis of brain functional statuses for early mild cognitive impairment using variational auto-encoder. *J. Ambient Intell. Humaniz. Comput.* **2020**, *14*, 5439–5450. [[CrossRef](#)]
243. Basu, S. Early prediction of alzheimer's disease progression using variational autoencoders. 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.
244. Biffi, C.; Cerrolaza, J.J.; Tarroni, G.; Bai, W.; de Marvao, A.; Oktay, O.; Ledig, C.; Le Folgoc, L.; Kamnitsas, K.; Doumou, G.; et al. Explainable Anatomical Shape Analysis Through Deep Hierarchical Generative Models. *IEEE Trans. Med. Imaging* **2020**, *39*, 2088–2099. [[CrossRef](#)]
245. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial nets. *Adv. Neural Inf. Process. Syst.* **2014**, *27*, 2672–2680.

246. Yi, X.; Walia, E.; Babyn, P. Generative adversarial network in medical imaging: A review. *Med. Image Anal.* **2019**, *58*, 101552. [[CrossRef](#)]
247. Islam, J.; Zhang, Y. GAN-based synthetic brain PET image generation. *Brain Inform.* **2020**, *7*, 1–12. [[CrossRef](#)]
248. Liu, Y.; Pan, Y.; Yang, W.; Ning, Z.; Yue, L.; Liu, M.; Shen, D. Joint Neuroimage Synthesis and Representation Learning for Conversion Prediction of Subjective Cognitive Decline. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Virtual, 4–6 October 2020; Springer: Berlin/Heidelberg, Germany, 2020.
249. Roychowdhury, S.; Roychowdhury, S. A Modular Framework to Predict Alzheimer’s Disease Progression Using Conditional Generative Adversarial Networks. In Proceedings of the 2020 International Joint Conference on Neural Networks (IJCNN), Glasgow, UK, 19–24 July 2020.
250. Baumgartner, C.F.; Koch, L.M.; Tezcan, K.C.; Ang, J.X.; Konukoglu, E. Visual feature attribution using wasserstein gans. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018.
251. Kim, H.W.; Lee, H.E.; Lee, S.; Oh, K.T.; Yun, M.; Yoo, S.K. Slice-selective learning for Alzheimer’s disease classification using a generative adversarial network: A feasibility study of external validation. *Eur. J. Nucl. Med.* **2020**, *47*, 2197–2206. [[CrossRef](#)]
252. Rachmadi, M.F.; Valdés-Hernández, M.d.C.; Makin, S.; Wardlaw, J.; Komura, T. Automatic spatial estimation of white matter hyperintensities evolution in brain MRI using disease evolution predictor deep neural networks. *Med. Image Anal.* **2020**, *63*, 101712. [[CrossRef](#)] [[PubMed](#)]
253. Sun, H.; Mehta, R.; Zhou, H.; Huang, Z.; Johnson, S.; Prabhakaran, V.; Singh, V. Dual-glow: Conditional flow-based generative model for modality transfer. In Proceedings of the IEEE International Conference on Computer Vision, Seoul, Republic of Korea, 27 October–2 November 2019. [[CrossRef](#)]
254. 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. [[CrossRef](#)] [[PubMed](#)]
255. 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. [[CrossRef](#)] [[PubMed](#)]
256. Fisher, C.K.; Smith, A.M.; Walsh, J.R. Machine learning for comprehensive forecasting of Alzheimer’s Disease progression. *Sci. Rep.* **2019**, *9*, 13622. [[CrossRef](#)] [[PubMed](#)]
257. Razavi, F.; Tarokh, M.J.; Alborzi, M. An intelligent Alzheimer’s disease diagnosis method using unsupervised feature learning. *J. Big Data* **2019**, *6*, 32. [[CrossRef](#)]
258. 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.
259. 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.
260. Cheng, J.; Dalca, A.V.; Fischl, B.; Zöllei, L. Cortical surface registration using unsupervised learning. *arXiv* **2020**, arXiv:2004.04617. [[CrossRef](#)]
261. Imen, W.; Amna, M.; Fatma, B.; Ezahra, S.F.; Masmoudi, N. Fast HEVC intra-CU decision partition algorithm with modified LeNet-5 and AlexNet. *Signal Image Video Process.* **2022**, *16*, 1811–1819. [[CrossRef](#)]
262. Lu, S.; Wang, S.-H.; Zhang, Y.-D. Detection of abnormal brain in MRI via improved AlexNet and ELM optimized by chaotic bat algorithm. *Neural Comput. Appl.* **2021**, *33*, 10799–10811. [[CrossRef](#)]
263. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. *Commun. ACM* **2017**, *60*, 84–90. [[CrossRef](#)]
264. Soffer, S.; Ben-Cohen, A.; Shimon, O.; Amitai, M.M.; Greenspan, H.; Klang, E. Convolutional Neural Networks for Radiologic Images: A Radiologist’s Guide. *Radiology* **2019**, *290*, 590–606. [[CrossRef](#)]
265. 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. [[CrossRef](#)]
266. Oseledets, I.V. Tensor-train decomposition. *SIAM J. Sci. Comput.* **2011**, *33*, 2295–2317. [[CrossRef](#)]
267. 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. [[CrossRef](#)]
268. 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. [[CrossRef](#)]
269. 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.
270. 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. [[CrossRef](#)]
271. 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. [[CrossRef](#)]
272. Islam, J.; Zhang, Y. Understanding 3D CNN Behavior for Alzheimer’s Disease Diagnosis from Brain PET Scan. *arXiv* **2019**, arXiv:1912.04563.

273. 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. [[CrossRef](#)]
274. 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. [[CrossRef](#)]
275. 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. [[CrossRef](#)] [[PubMed](#)]
276. 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. [[CrossRef](#)] [[PubMed](#)]
277. Basaia, S.; Agosta, F.; Wagner, L.; Canu, E.; Magnani, G.; Santangelo, R.; Filippi, M. Studying the manifold structure of Alzheimer's Disease: A deep learning approach using convolutional autoencoders. *IEEE J. Biomed. Health Inform.* **2019**, *24*, 17–26.
278. Payan, A.; Montana, G. Predicting Alzheimer's disease: A neuroimaging study with 3D convolutional neural networks. *arXiv* **2015**, arXiv:1502.02506.
279. 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.
280. 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.
281. 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.
282. 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. [[CrossRef](#)]
283. 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. [[CrossRef](#)]
284. 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. [[CrossRef](#)]
285. 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. [[CrossRef](#)]
286. 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.
287. 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. 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. 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.
290. 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.
291. Wang, M.; Lian, C.; Yao, D.; Zhang, D.; Liu, M.; Shen, D. Spatial-Temporal Dependency Modeling and Network Hub Detection for Functional MRI Analysis via Convolutional-Recurrent Network. *IEEE Trans. Biomed. Eng.* **2019**, *67*, 2241–2252. [[CrossRef](#)]
292. Lee, G.; Nho, K.; Kang, B.; Sohn, K.-A.; Kim, D.; Weiner, M.W.; Aisen, P.; Petersen, R.; Jack, C.R.; Jagust, W.; et al. Predicting Alzheimer's disease progression using multi-modal deep learning approach. *Sci. Rep.* **2019**, *9*, 1952. [[CrossRef](#)]
293. Li, H.; Fan, Y. Early prediction of Alzheimer's disease dementia based on baseline hippocampal MRI and 1-year follow-up cognitive measures using deep recurrent neural networks. In Proceedings of the 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), Venice, Italy, 8–11 April 2019.
294. 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. [[CrossRef](#)]
295. Bronstein, M.M.; Bruna, J.; LeCun, Y.; Szlam, A.; Vandergheynst, P. Geometric Deep Learning: Going beyond Euclidean data. *IEEE Signal Process. Mag.* **2017**, *34*, 18–42. [[CrossRef](#)]
296. Ma, X.; Wu, G.; Kim, W.H. Enriching Statistical Inferences on Brain Connectivity for Alzheimer's Disease Analysis via Latent Space Graph Embedding. In Proceedings of the 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), Iowa City, IA, USA, 3–7 April 2020.

297. Wu, Z.; Pan, S.; Chen, F.; Long, G.; Zhang, C.; Philip, S.Y. A comprehensive survey on graph neural networks. *IEEE Trans. Neural Netw. Learn. Syst.* **2020**, *32*, 4–24. [[CrossRef](#)]
298. Song, T.-A.; Chowdhury, S.R.; Yang, F.; Jacobs, H.; El Fakhri, G.; Li, Q.; Johnson, K.; Dutta, J. Graph Convolutional Neural Networks for Alzheimer’s Disease Classification. In Proceedings of the 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), Venice, Italy, 8–11 April 2019.
299. Song, T.-A.; Chowdhury, S.R.; Yang, F.; Jacobs, H.I.L.; Sepulcre, J.; Wedeen, V.J.; Johnson, K.A.; Dutta, J. A Physics-Informed Geometric Learning Model for Pathological Tau Spread in Alzheimer’s Disease. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Lima, Peru, 4–8 October 2020; Springer: Berlin/Heidelberg, Germany, 2020.
300. Song, X.; Frangi, A.; Xiao, X.; Cao, J.; Wang, T.; Lei, B. Integrating similarity awareness and adaptive calibration in graph convolution network to predict disease. In *Medical Image Computing and Computer Assisted Intervention—MICCAI 2020: Proceedings of the 23rd International Conference, Lima, Peru, 4–8 October 2020*; Proceedings, Part VII 23; Springer: Berlin/Heidelberg, Germany, 2020.
301. Yang, J.; Zheng, W.-S.; Yang, Q.; Chen, Y.-C.; Tian, Q. Spatial-Temporal Graph Convolutional Network for Video-Based Person Re-Identification. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020.
302. Mirakhorli, J.; Mirakhorli, M. Graph-Based Method for Anomaly Detection in Functional Brain Network using Variational Autoencoder. *bioRxiv* **2019**, 616367. [[CrossRef](#)]
303. Zhu, W.; Razavian, N. Graph Neural Network on Electronic Health Records for Predicting Alzheimer’s Disease. *arXiv* **2019**, arXiv:1912.03761.
304. Ma, J.; Zhu, X.; Yang, D.; Chen, J.; Wu, G. Attention-Guided Deep Graph Neural Network for Longitudinal Alzheimer’s Disease Analysis. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Lima, Peru, 4–8 October 2020; Springer: Berlin/Heidelberg, Germany, 2020.
305. Bruna, J.; Zaremba, W.; Szlam, A.; LeCun, Y. Spectral networks and locally connected networks on graphs. *arXiv* **2013**, arXiv:1312.6203.
306. Wee, C.-Y.; Liu, C.; Lee, A.; Poh, J.S.; Ji, H.; Qiu, A. Cortical graph neural network for AD and MCI diagnosis and transfer learning across populations. *NeuroImage Clin.* **2019**, *23*, 101929. [[CrossRef](#)]
307. Zhao, X.; Zhou, F.; Ou-Yang, L.; Wang, T.; Lei, B. Graph Convolutional Network Analysis for Mild Cognitive Impairment Prediction. In Proceedings of the 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), Venice, Italy, 8–11 April 2019.
308. Kazi, A.; Shekarforoush, S.; Krishna, S.A.; Burwinkel, H.; Vivar, G.; Wiestler, B.; Kortüm, K.; Ahmadi, S.-A.; Albarqouni, S.; Navab, N. Graph convolution based attention model for personalized disease prediction. 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.
309. Huang, Y.; Chung, A.C. Edge-Variational Graph Convolutional Networks for Uncertainty-Aware Disease Prediction. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Lima, Peru, 4–8 October 2020; Springer: Berlin/Heidelberg, Germany, 2020.
310. Yu, S.; Wang, S.; Xiao, X.; Cao, J.; Yue, G.; Liu, D.; Wang, T.; Xu, Y.; Lei, B. Multi-scale Enhanced Graph Convolutional Network for Early Mild Cognitive Impairment Detection. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Lima, Peru, 4–8 October 2020; Springer: Berlin/Heidelberg, Germany, 2020.
311. Chakraborty, R.; Zhen, X.; Vogt, N.; Bendlin, B.; Singh, V. Dilated convolutional neural networks for sequential manifold-valued data. In Proceedings of the IEEE International Conference on Computer Vision, Seoul, Republic of Korea, 27–28 October 2019.
312. You, Z.; Zeng, R.; Lan, X.; Ren, H.; You, Z.; Shi, X.; Zhao, S.; Guo, Y.; Jiang, X.; Hu, X. Alzheimer’s Disease Classification with a Cascade Neural Network. *Front. Public Health* **2020**, *8*, 584387. [[CrossRef](#)]
313. Gadgil, S.; Zhao, Q.; Pfefferbaum, A.; Sullivan, E.V.; Adeli, E.; Pohl, K.M. Spatio-Temporal Graph Convolution for Resting-State fMRI Analysis. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Lima, Peru, 4–8 October 2020; Springer: Berlin/Heidelberg, Germany, 2020.
314. Pfau, D.; Vinyals, O. Connecting generative adversarial networks and actor-critic methods. *arXiv* **2016**, arXiv:1610.01945.
315. Capecchi, E.; Dobarjeh, 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.
316. Suk, H.-I.; Shen, D. Deep learning-based feature representation for AD/MCI classification. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Nagoya, Japan, 22–26 October 2016; Springer: Berlin/Heidelberg, Germany, 2013.
317. Suk, H.-I.; Lee, S.-W.; Shen, D. Deep ensemble learning of sparse regression models for brain disease diagnosis. *Med. Image Anal.* **2017**, *37*, 101–113. [[CrossRef](#)]
318. Shi, J.; Zheng, X.; Li, Y.; Zhang, Q.; Ying, S. Multimodal neuroimaging feature learning with multimodal stacked deep polynomial networks for diagnosis of Alzheimer’s disease. *IEEE J. Biomed. Health Inform.* **2017**, *22*, 173–183. [[CrossRef](#)] [[PubMed](#)]
319. Lu, D.; Popuri, K.; Ding, G.W.; Balachandar, R.; Beg, M.F. Multiscale deep neural network based analysis of FDG-PET images for the early diagnosis of Alzheimer’s disease. *Med. Image Anal.* **2018**, *46*, 26–34. [[CrossRef](#)] [[PubMed](#)]

320. Ning, K.; Chen, B.; Sun, F.; Hobel, Z.; Zhao, L.; Matloff, W.; Toga, A.W. Classifying Alzheimer's disease with brain imaging and genetic data using a neural network framework. *Neurobiol. Aging* **2018**, *68*, 151–158. [[CrossRef](#)] [[PubMed](#)]
321. Lee, E.; Choi, J.-S.; Kim, M.; Suk, H.-I. Toward an interpretable Alzheimer's disease diagnostic model with regional abnormality representation via deep learning. *Neuroimage* **2019**, *202*, 116113. [[CrossRef](#)] [[PubMed](#)]
322. Bashyam, V.M.; Erus, G.; Doshi, J.; Habes, M.; Nasrallah, I.M.; Truelove-Hill, M.; Srinivasan, D.; Mamourian, L.; Pomponio, R.; Fan, Y.; et al. MRI signatures of brain age and disease over the lifespan based on a deep brain network and 14,468 individuals worldwide. *Brain* **2020**, *143*, 2312–2324. [[CrossRef](#)]
323. Lundervold, A.S.; Lundervold, A. An overview of deep learning in medical imaging focusing on MRI. *Z. Med. Phys.* **2019**, *29*, 102–127. [[CrossRef](#)]
324. 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. [[CrossRef](#)]
325. Deng, J.; Dong, W.; Socher, R.; Li, L.-J.; Li, K.; Fei-Fei, L. Imagenet: A large-scale hierarchical image database. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009.
326. 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. [[CrossRef](#)]
327. 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. [[CrossRef](#)]
328. 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. [[CrossRef](#)]
329. 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. [[CrossRef](#)]
330. 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. [[CrossRef](#)]
331. 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.
332. Bae, J.B.; Lee, S.; Jung, W.; Park, S.; Kim, W.; Oh, H.; Han, J.W.; Kim, G.E.; Kim, J.S.; Kim, J.H. Identification of Alzheimer's disease using a convolutional neural network model based on T1-weighted magnetic resonance imaging. *Sci. Rep.* **2020**, *10*, 22252. [[CrossRef](#)]
333. Lin, L.; Wu, Y.; Wu, X.; Wu, S. APOE- ϵ 4 allele load modifies the brain aging process in cognitively normal late middle aged and older adults. In Proceedings of the International Conference on Artificial Intelligence, Information Processing and Cloud Computing, Sanya, China, 21–22 June 2019.
334. Cheng, B.; Initiative, T.A.D.N.; Liu, M.; Shen, D.; Li, Z.; Zhang, D. Multi-Domain Transfer Learning for Early Diagnosis of Alzheimer's Disease. *Neuroinformatics* **2016**, *15*, 115–132. [[CrossRef](#)]
335. Sagi, O.; Rokach, L. Ensemble learning: A survey. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2018**, *8*, e1249. [[CrossRef](#)]
336. Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: A simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* **2014**, *15*, 1929–1958.
337. Geras, K.J.; Wolfson, S.; Shen, Y.; Wu, N.; Kim, S.; Kim, E.; Heacock, L.; Parikh, U.; Moy, L.; Cho, K. High-resolution breast cancer screening with multi-view deep convolutional neural networks. *arXiv* **2017**, arXiv:1703.07047.
338. Seeley, M.; Clement, M.; Giraud-Carrier, C.; Snell, Q.; Bodily, P.; Fujimoto, S. A structured approach to ensemble learning for Alzheimer's disease prediction. In Proceedings of the 5th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics, Newport Beach, CA, USA, 20–23 September 2014.
339. Zhou, T.; Thung, K.-H.; Liu, M.; Shi, F.; Zhang, C.; Shen, D. Multi-modal latent space inducing ensemble SVM classifier for early dementia diagnosis with neuroimaging data. *Med. Image Anal.* **2020**, *60*, 101630. [[CrossRef](#)]
340. Liu, S.; Cai, W.; Liu, S.; Zhang, F.; Fulham, M.; Feng, D.; Pujol, S.; Kikinis, R. Multimodal neuroimaging computing: A review of the applications in neuropsychiatric disorders. *Brain Inform.* **2015**, *2*, 167. [[CrossRef](#)]
341. Qiu, S.; Chang, G.H.; Panagia, M.; Gopal, D.M.; Au, R.; Kolachalama, V.B. Fusion of deep learning models of MRI scans, Mini-Mental State Examination, and logical memory test enhances diagnosis of mild cognitive impairment. *Alzheimer's Dement. Diagn. Assess. Dis. Monit.* **2018**, *10*, 737–749. [[CrossRef](#)]
342. Zheng, X.; Shi, J.; Li, Y.; Liu, X.; Zhang, Q. Multi-modality stacked deep polynomial network based feature learning for Alzheimer's disease diagnosis. In Proceedings of the 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), Prague, Czech Republic, 13–16 April 2016.
343. Zhou, T.; Thung, K.-H.; Zhu, X.; Shen, D. Effective feature learning and fusion of multimodality data using stage-wise deep neural network for dementia diagnosis. *Hum. Brain Mapp.* **2018**, *40*, 1001–1016. [[CrossRef](#)] [[PubMed](#)]
344. Saribudak, A.; Subick, A.A.; Kim, N.H.; Rutta, J.A.; Uyar, M.Ü. Gene Expressions, Hippocampal Volume Loss and MMSE Scores in Computation of Progression and Pharmacologic Therapy Effects for Alzheimer's Disease. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2018**, *17*, 608–622. [[CrossRef](#)] [[PubMed](#)]

345. Liu, J.; Shang, S.; Zheng, K.; Wen, J.-R. Multi-view ensemble learning for dementia diagnosis from neuroimaging: An artificial neural network approach. *Neurocomputing* **2016**, *195*, 112–116. [[CrossRef](#)]
346. Lu, D.; Popuri, K.; Ding, G.W.; Balachandar, R.; Beg, M.F.; Weiner, M.; Aisen, P.; Petersen, R.; Jack, C.; Jagust, W.; et al. Multimodal and Multiscale Deep Neural Networks for the Early Diagnosis of Alzheimer’s Disease using structural MR and FDG-PET images. *Sci. Rep.* **2018**, *8*, 5697. [[CrossRef](#)]
347. Senanayake, U.; Sowmya, A.; Dawes, L. Deep fusion pipeline for mild cognitive impairment diagnosis. In Proceedings of the 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), Washington, DC, USA, 4–7 April 2018.
348. Moscovich, A.; Rosset, S. On the cross-validation bias due to unsupervised preprocessing. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **2022**, *84*, 1474–1502. [[CrossRef](#)]
349. Kenett, R.S.; Gotwalt, C.; Freeman, L.; Deng, X. Self-supervised cross validation using data generation structure. *Appl. Stoch. Model. Bus. Ind.* **2022**, *38*, 750–765. [[CrossRef](#)]
350. Nayak, G.; Padhy, N.; Mishra, T.K. 2D-DOST for seizure identification from brain MRI during pregnancy using KRVFL. *Health Technol.* **2022**, *12*, 757–764. [[CrossRef](#)]
351. Mila, C.; Mateu, J.; Pebesma, E.; Meyer, H. Nearest neighbour distance matching Leave-One-Out Cross-Validation for map validation. *Methods Ecol. Evol.* **2022**, *13*, 1304–1316. [[CrossRef](#)]
352. Wang, B.; Zou, H. Fast and exact leave-one-out analysis of large-margin classifiers. *Technometrics* **2022**, *64*, 291–298. [[CrossRef](#)]
353. Kim, J.; Basak, J.M.; Holtzman, D.M. The role of apolipoprotein E in Alzheimer’s disease. *Neuron* **2009**, *63*, 287–303. [[CrossRef](#)]
354. Mårtensson, G.; Ferreira, D.; Granberg, T.; Cavallin, L.; Oppedal, K.; Padovani, A.; Rektorova, I.; Bonanni, L.; Pardini, M.; Kramerberger, 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. [[CrossRef](#)]
355. Li, H.; Habes, M.; Wolk, D.A.; Fan, Y. A deep learning model for early prediction of Alzheimer’s disease dementia based on hippocampal magnetic resonance imaging data. *Alzheimer’s Dement.* **2019**, *15*, 1059–1070. [[CrossRef](#)]
356. DeLong, E.R.; DeLong, D.M.; Clarke-Pearson, D.L. Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach. *Biometrics* **1988**, *44*, 837–845. [[CrossRef](#)]
357. Bäckström, K.; Nazari, M.; Gu, I.Y.-H.; Jakola, A.S. An efficient 3D deep convolutional network for Alzheimer’s disease diagnosis using MR images. In Proceedings of the 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), Washington, DC, USA, 4–7 April 2018.
358. Abadi, M.; Barham, P.; Chen, J.; Chen, Z.; Davis, A.; Dean, J.; Devin, M.; Ghemawat, S.; Irving, G.; Isard, M. Tensorflow: A system for large-scale machine learning. In Proceedings of the 12th {USENIX} Symposium on Operating Systems Design and Implementation ({OSDI} 16), Savannah, GA, USA, 2–4 November 2016.
359. Jia, Y.; Shelhamer, E.; Donahue, J.; Karayev, S.; Long, J.; Girshick, R.; Guadarrama, S.; Darrell, T. Caffe: Convolutional architecture for fast feature embedding. In Proceedings of the 22nd ACM International Conference on Multimedia, Orlando, FL, USA, 3–7 November 2014.
360. Team, T.T.D.; Al-Rfou, R.; Alain, G.; Almahairi, A.; Angermueller, C.; Bahdanau, D.; Ballas, N.; Bastien, F.; Bayer, J.; Belikov, A. Theano: A Python framework for fast computation of mathematical expressions. *arXiv* **2016**, arXiv:1605.02688.
361. Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L. Pytorch: An imperative style, high-performance deep learning library. In Proceedings of the Advances in Neural Information Processing Systems, NIPS 2019, Vancouver, BC, Canada, 8–14 December 2019.
362. Nigri, E.; Ziviani, N.; Cappabianco, F.; Antunes, A.; Veloso, A. Explainable Deep CNNs for MRI-Based Diagnosis of Alzheimer’s Disease. *arXiv* **2020**, arXiv:2004.12204.
363. Lian, C.; Liu, M.; Wang, L.; Shen, D. End-to-end dementia status prediction from brain mri using multi-task weakly-supervised attention network. 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.
364. 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.
365. 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.
366. Kawaguchi, K.; Kaelbling, L.; Bengio, Y. Generalization in deep learning. *arXiv* **2017**, arXiv:1710.05468.
367. 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](#)]
368. 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.
369. Wu, B.; Sun, X.; Hu, L.; Wang, Y. Learning with unsure data for medical image diagnosis. In Proceedings of the IEEE International Conference on Computer Vision, Seoul, Republic of Korea, 27 October–2 November 2019.

-
370. Isaacs, J.D.; Boenink, M. Biomarkers for dementia: Too soon for routine clinical use. *Lancet Neurol.* **2020**, *19*, 884–885. [[CrossRef](#)]
371. Jack, C.R., Jr.; Wiste, H.J.; Weigand, S.D.; Rocca, W.A.; Knopman, D.S.; Mielke, M.M.; Lowe, V.J.; Senjem, M.L.; Gunter, J.L.; Preboske, G.M. Age-specific population frequencies of cerebral β -amyloidosis and neurodegeneration among people with normal cognitive function aged 50–89 years: A cross-sectional study. *Lancet Neurol.* **2014**, *13*, 997–1005. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.