Computer Methods and Programs in Biomedicine 259 (2025) 108506

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

journal homepage: www.sciencedirect.com/journal/computer-methods-and-programs-in-biomedicine

Deep learning techniques for automated Alzheimer's and mild cognitive impairment disease using EEG signals: A comprehensive review of the last decade (2013 - 2024)

Madhav Acharya [a,*], Ravinesh C Deo [a,*], Xiaohui Tao [b], Prabal Datta Barua [a,c,d,e,f,g,h,i,j,k], Aruna Devi [l], Anirudh Atmakuru [m], Ru-San Tan [n]

[a] School of Mathematics, Physics and Computing, University of Southern Queensland, Springfield Central, QLD 4300, Australia
[b] School of Mathematics, Physics and Computing, University of Southern Queensland, Toowoomba QLD 4350, Australia
[c] Cogninet AI, Sydney, NSW 2010, Australia
[d] Faculty of Engineering and Information Technology, University of Technology Sydney, Sydney, NSW 2007, Australia
[e] School of Computing and Information Science, Anglia Ruskin University Cambridge Campus, UK
[f] Australian International Institute of Higher Education, Sydney, NSW 2000, Australia
[g] School of Science Technology, University of New England, Australia
[h] School of Biosciences, Taylor's University, Malaysia
[i] School of Computing, SRM Institute of Science and Technology, India
[j] School of Science and Technology, Kumamoto University, Japan
[k] Sydney School of Education and Social Work, University of Sydney, Australia
[l] School of Education and Tertiary Access, University of the Sunshine Coast, Petrie, QLD 4502, Australia
[m] University of Massachusetts, Amherst, USA
[n] National Heart Centre Singapore, Singapore, Duke-NUS Medical School, Singapore

A R T I C L E I N F O

Keywords:
Artificial intelligence
Neurological diseases
EEG signals
Deep learning
Alzheimer's disease (AD)
Mild cognitive impairment (MCI)

A B S T R A C T

Background and Objectives: Mild Cognitive Impairment (MCI) and Alzheimer's Disease (AD) are progressive neurological disorders that significantly impair the cognitive functions, memory, and daily activities. They affect millions of individuals worldwide, posing a significant challenge for its diagnosis and management, leading to detrimental impacts on patients' quality of lives and increased burden on caregivers. Hence, early detection of MCI and AD is crucial for timely intervention and effective disease management.
Methods: This study presents a comprehensive systematic review focusing on the applications of deep learning in detecting MCI and AD using electroencephalogram (EEG) signals. Through a rigorous literature screening process based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines, the research has investigated 74 different papers in detail to analyze the different approaches used to detect MCI and AD neurological disorders.
Results: The findings of this study stand out as the first to deal with the classification of dual MCI and AD (MCI+AD) using EEG signals. This unique approach has enabled us to highlight the state-of-the-art high-performing models, specifically focusing on deep learning while examining their strengths and limitations in detecting the MCI, AD, and the MCI+AD comorbidity situations.
Conclusion: The present study has not only identified the current limitations in deep learning area for MCI and AD detection but also proposes specific future directions to address these neurological disorders by implement best practice deep learning approaches. Our main goal is to offer insights as references for future research encouraging the development of deep learning techniques in early detection and diagnosis of MCI and AD neurological disorders. By recommending the most effective deep learning tools, we have also provided a benchmark for future research, with clear implications for the practical use of these techniques in healthcare.

* Corresponding authors.
E-mail addresses: madhav.acharya@unisq.edu.au (M. Acharya), ravinesh.deo@unisq.edu.au (R.C. Deo).

https://doi.org/10.1016/j.cmpb.2024.108506
Received 13 August 2024; Received in revised form 29 October 2024; Accepted 6 November 2024
Available online 12 November 2024

1. Introduction

Dementia is a group of diseases that manifest with a decline in memory, reasoning, and the ability to perform daily activities. Globally, more than 55 million people suffer from dementia worldwide, with over 60 % in low-and middle-income countries [1,2], and 10 million new cases are diagnosed every year [3]. Mild cognitive impairment (MCI) is considered the intermediate stage between the cognitive changes seen in normal aging–e.g., occasional forgetfulness–and those associated with dementia [4]. Unlike dementia, MCI patients often retain functional capacity. The prevalence of MCI among adults aged 60 and above ranges from approximately 6.7 % to 25.2 % [5]. The risk decreases with age and higher levels of education, and it is more prevalent in men [6]. The most common cause of dementia is Alzheimer's disease (AD), a progressive neurodegenerative disease characterized by amyloid-beta peptide's accumulation in the brain [7]. The exact etiology of AD is unclear, although multiple factors have been implicated, including advanced age and gene [8]. Genetic variations in the APOE ε4 allele are associated with AD and MCI [9]–some MCI patients are pre-clinical AD who eventually progress to clinical AD with severe functional impairment–but the presence of a culprit gene may not always result in cognitive decline [10]. Finally, several chronic diseases and lifestyle factors have been linked to cognitive impairment, e.g., diabetes, hypertension, hypercholesterolemia, obesity, depression, smoking, lack of physical exercise, and low education levels [11].

There is no definitive test for Alzheimer's disease (AD). The clinical diagnosis requires comprehensive cognitive and neurological evaluations, neuroimaging such as magnetic resonance imaging (MRI) and positron emission tomography (PET), and electroencephalogram (EEG) [12]. In imaging studies, the hippocampus (the memory center) shows a gradual decrease in size with age, mild cognitive impairment (MCI), and AD (Fig. 1). Imaging tests are expensive, time-consuming, and require expert interpretation [13].

Fig. 1 demonstrates the progression of brain degeneration from normal aging to severe AD, highlighting the specific structural changes that occur in key brain areas responsible for memory and cognitive functions. The Fig. serves as a visual aid to understand how Alzheimer's disease manifests and worsens over time. EEG, which can continuously map the brain's surface electrical potentials via multichannel scalp electrodes, provides spatial and highly resolved temporal information about the functional activities of various brain regions and has been used to detect diverse neurocognitive disorders, including Alzheimer's Disease (AD) [14,15]. EEG signals are particularly useful in such applications as they capture both linear and non-linear dynamic information [16]. Disease-associated variations in EEG signals and their frequency spectra allow for diagnostic discrimination [17]. However, because EEG signals have small amplitudes (measured in microvolts), subtle changes in different channels are difficult to analyse, especially when the data set is large. Hence, there is a need for computer-aided diagnosis of Mild Cognitive Impairment (MCI) and AD in the presence of voluminous data. Indeed, artificial intelligence (AI) techniques, especially deep learning, have shown great potential to detect and diagnose MCI and AD based on EEG signals [18]. Instead of relying entirely on human intervention, which can be time-consuming and expensive, deep learning techniques self-optimise by analysing large amounts of data to detect intricate patterns in EEG signals automatically. In other words, they can detect subtle patterns in the data that are otherwise difficult for humans to discern [19]. The variations in EEG signals and their frequency spectra associated with diseases allow diagnostic discrimination [20]. As EEG signals have small amplitudes (microvolts), subtle changes in the different channels are difficult to analyse, especially when the data is huge [21].

Fig. 2 provides a visual summary of the landscape of systematic reviews in AD and MCI detection, highlighting the complementary focus of different studies. Each review covers a specific aspect of neurological detection (EEG, MRI, progression from MCI to AD), contributing to a holistic understanding of how deep learning and other techniques are advancing the detection of neurodegenerative diseases.

Hence, computer-aided diagnosis of MCI and AD is needed in the presence of voluminous data. Indeed, artificial intelligence (AI) techniques, especially deep learning, have shown great potential to detect and diagnose MCI and AD based on EEG signals [22]. Instead of relying entirely on human intervention, which could be time-consuming and expensive, deep learning techniques self-optimize by analyzing large amounts of data to identify intricate patterns in EEG signals automatically. In other words, they can detect subtle patterns in the data that are otherwise difficult for humans to discern [23].

In this review paper, we aim to perform a detailed systematic review of deep learning methods for diagnosing AD and MCI neurological disorders using EEG signals. The remainder of the paper is organized as follows: Section 2 details the information search methodology related to AD and MCI, Section 3 presents the results and synthesis of the findings, Section 4 discusses the results in further detail, and Section 5 presents the conclusions and recommendations for future research directions in adopting deep learning for AD and MCI neurological disorder detection.

2. Material and methods

2.1. Related reviews

Our literature search demonstrated that there currently appears to be a lack of systematic reviews focusing on deep-learning methods for EEG-based AD and MCI diagnosis. Therefore, using different combinations of the search words "EEG", "deep learning", "Alzheimer OR AD OR MCI OR Mild Cognitive Impairment" and "Review" on PubMed and Google Scholar, we found 14 review studies (Table A. 1) that were most relevant for AD and MCI neurological disorder detection. The search shows that several studies [24,25,26] have reviewed different techniques to identify the various neuropsychiatric disorders, including MCI, AD, Parkinson's diseases, bipolar disorder, depression, etc, but none of those studies have specifically focused on MCI or AD detection techniques (like deep learning), and utilising the particular data types (like EEG)

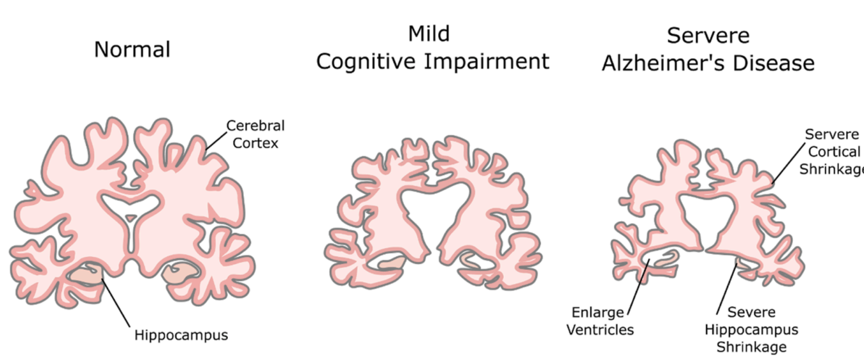

Fig. 1. Brain structure of the normal individuals compared to the MCI and AD patients.

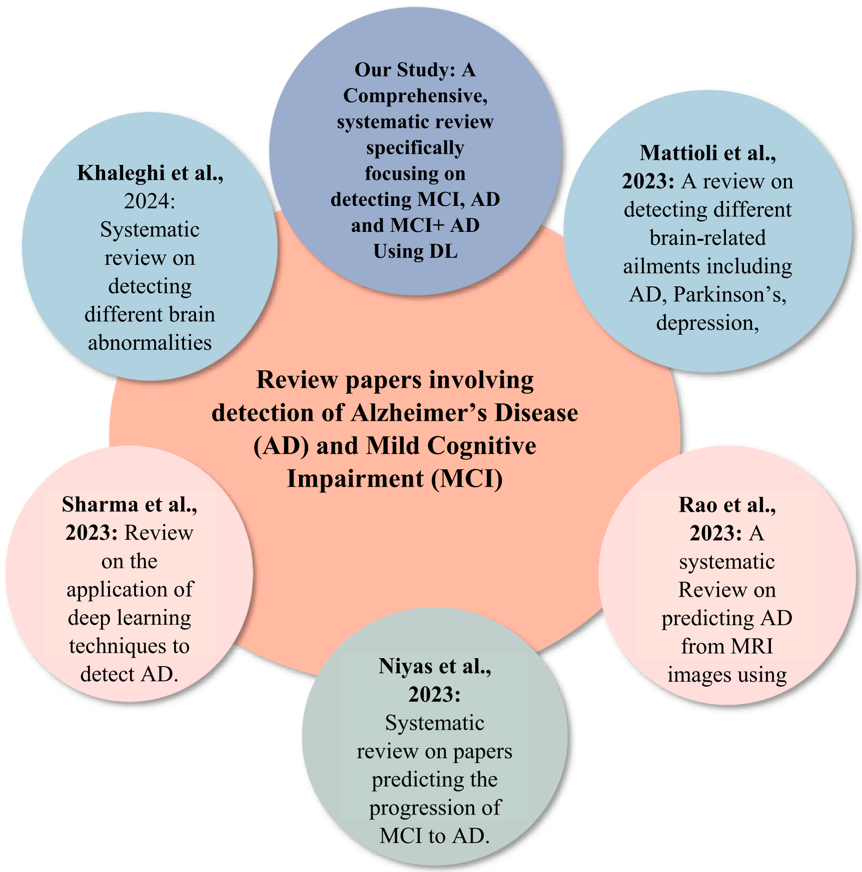

Fig. 2. Comparison of our review paper with other review papers developed for automated detection of AD and MCI.

considered in this research. Some papers reviewed deep learning, EEG signals, and MCI or AD in general, but they lacked a specific focus on EEG-based MCI/AD detection using AI techniques. [27] and [28] have also studied the application of machine learning (ML) to detect AD using EEG images. Still, these studies did not include the high-performing deep learning models that are commonly used today.

In Fig. 2, we depict how our carefully selected papers differed from five of the most recent related review studies in this field.

According to Fig. 3, there are three phases: Identification, Screening and Inclusion. Initially, 440 papers were collected, with 359 remaining after removing duplicates. In the Screening phase, 244 papers were excluded for irrelevance, leaving 115 for assessment. Finally, 41 more papers were excluded due to a lack of machine learning or deep learning results, resulting in 74 papers being included in the review. The process ensures a rigorous selection of relevant, high-quality studies.

2.2. Literature search

We conducted a comprehensive literature search in PubMed, Scopus, Web of Science, and IEEE repositories for articles published between 1st January 2014 and 30th June 2024, following the PRISMA guidelines. Using the Boolean string: (Alzheimer OR AD OR MCI OR mild cognitive impairment) AND (Deep Learning OR Transfer Learning OR Natural Language Processing) AND (EEG OR Electroencephalogram), the initial search yielded 440 results across the four databases. After removing 81 duplicates and excluding 244 irrelevant works, review studies and non-journal publications, 115 articles remained. Of these, 41 papers did not report any performance results of the deep learning models, leaving 74 papers for analysis and review (Fig. 3: Article search strategy based on the PRISMA guidelines.).

3. Results

Among the 74 papers published in the last decade, 45 papers (61 %) have been published from 2022 (Fig. 4), which mirrors the secular development of state-of-the-art deep learning models in the field as well as computing power, which significantly improved the accuracy and efficiency of analyzing EEG signals for detecting MCI and AD. For the analysis of results, we stratified the 74 papers into three groups based on the condition/s being classified: MCI (11 papers), AD (38 papers), or both (25 papers). The first two groups classify MCI patients versus healthy controls and AD patients versus healthy controls, respectively. The third group included papers that either classify MCI versus AD versus healthy controls or MCI versus AD patients. Among the last group are papers that studied the progression of MCI into AD.

Fig. 4 highlights a significant increase in research starting in 2018, with peaks in 2022 and 2023 due to advances in deep learning models and computational power. 2014-2017 saw low activity, likely reflecting the initial stages of applying AI to EEG analysis. A drop in 2020 is attributed to the pandemic, but there was a resurgence in 2021, aligning with improvements in machine learning technologies. The text notes a decline in publications in 2024, potentially indicating stabilization in research. Fig. 4 shows that 61 % of papers were published from 2022 onwards, emphasizing the field's rapid growth, driven by technological advancements and the need for early detection of neurodegenerative diseases.

3.1. MCI detection

In this section, we report the key papers that used deep learning methods for EEG-based MCI diagnosis (Table A. 2). Early detection of MCI is crucial so that intervention can be introduced to retard its

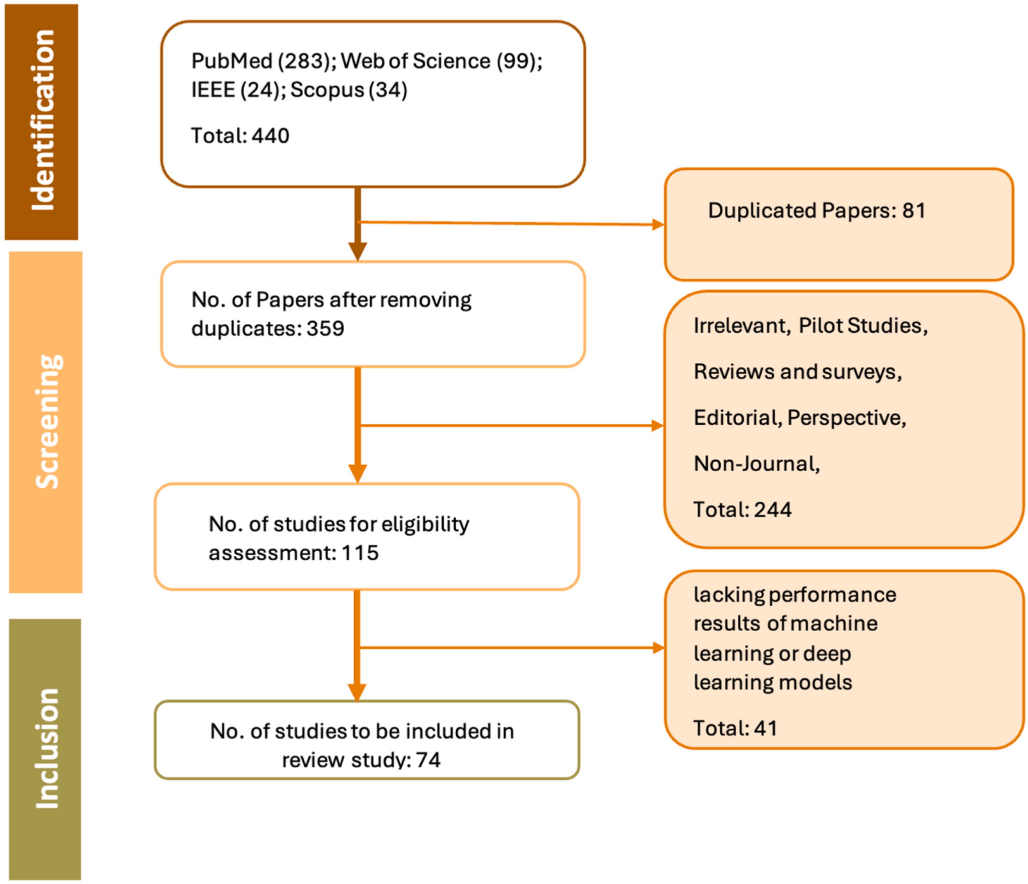

Fig. 3. Article search strategy based on the PRISMA guidelines.

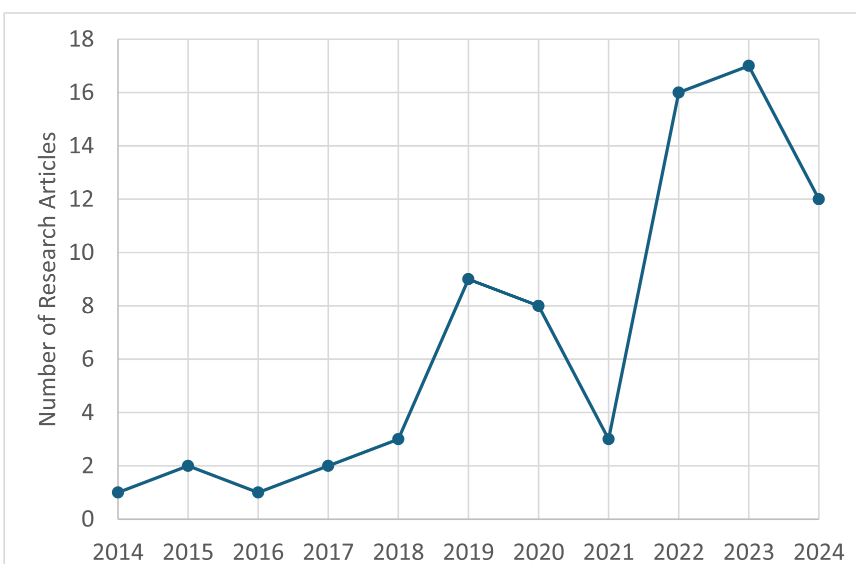

Fig. 4. Distribution of articles by the publication year.

progression to AD, which is associated with severe cognitive decline, loss of independence, and higher healthcare costs. Modifications in brain connectivity density can be detected on the EEG. Compared with MCI, AD patients exhibit increased amounts of permutation Jaccard distances (PJD) and reduction in network density across all sub-bands [29]. [30] investigated various techniques like tuned residue iteration decomposition (t-RIDE), residue iteration decomposition (RIDE), independent component analysis (ICA), and the grand average method on data acquired from subjects who had performed a P300 speller task. Their t-RIDE algorithm demonstrated high efficiency for MCI diagnosis versus controls, which opens up the possibility of AI-based techniques for automatic MCI screening and, via mobile devices, remote monitoring of neurocognitive function.

Later studies are broadly categorized into three groups: articles that used scarce data, transfer learning models, and EEG data in a multimodal setting. To overcome the challenge of obtaining large open-access EEG databases to train AI models, [31] has developed a model based on iterative amplitude adjusted Fourier transform (IAAFT) and bidirectional long short-term memory (BiLSTM) that could distinguish MCI patients versus healthy controls using small amounts of EEG signal data. Used along with data augmentation, IAAFT generates surrogate EEG data with similar characteristics to the original data, thereby decreasing the chances of overfitting and improving generalizability.

By additionally deploying BiLSTM to capture temporal dependencies in EEG signals, their model attained 97.20 % classification accuracy on a small dataset of 10 MCI patients and 10 healthy controls. Other studies relied on pre-processing steps to mitigate data scarcity, including denoising using techniques like stationary wavelet transformation, segmentation, and down-sampling of raw data to expand the sample size and achieve similar classification performance. The key lies in performing clever transformations to represent the available data in a more interpretable form. In [32] and [33], stationary wavelet and discrete wavelet transformations, respectively, were used to decompose EEG signals into sub-bands, from which features were extracted to perform MCI versus normal classifications, which attained accuracies exceeding 95 %. These studies demonstrate that by carefully selecting EEG channels and features, it is possible to achieve high accuracy in detecting MCI, even in the presence of limited data.

[34] used ResNet-18 to distinguish MCI patients versus healthy controls based on the frequency and spatial properties of EEG data. They observed that the frontal, left temporal, and parietal areas of the brain were most affected in MCI patients, which differed from healthy controls in the θ and low α bands on the EEG signal. From studying how MCI patients differed from mild AD patients, they observed that MCI patients had a larger affected right temporal area. Incorporating these findings into the model, they attained a high 98.33 % accuracy for the classification of MCI patients versus healthy controls. [35] reproduced similar results on a larger, publicly available database, which not only lent support to the results of [34], but also suggested a disrupted brain network organization in MCI patients characterized by higher local efficiency in the beta band on the EEG signal.

Integrating information from multimodal sources in individual patients can potentially provide a more complex and comprehensive capture of the disease for more a accurate diagnosis. However, it can come at the cost of increased difficulty in acquiring the different data. Several studies used such an approach to detect MCI, with EEG signals being part of the multimodal datasets. [36] They used eye movement data, information from neuropsychological assessments, and EEG signals. Their model extracted 40 features using logistic regression, among which five features were significantly related to MCI, yielding 81.51 % classification accuracy. [37] studied mindfulness impact in MCI using sleep-related information. In a double-blind randomized controlled trial, they assigned 75 patients with MCI and insomnia into two groups: the mindfulness group and the health education control group.

Compared to the control group, the mindfulness group showed significant improvements in sleep quality along with low levels of anxiety and stress. Analysis of EEG recordings of both these groups revealed changes in brain activity, indicating relaxation and alterations in frequency bands associated with attention during mindfulness practice. This suggests that mindfulness can enhance sleep and cognitive abilities in MCI patients, providing a low-cost, scalable intervention suitable for implementation. [38] studied robot-based training for improving working memory and cognitive function in older patients. Their numerical results indicated that the intervention led to an 8 % increase in cognitive scores, as measured by standardized assessments, compared to the control group. [39] developed an EEG-based MCI to assess cognitive workloads in MCI patients. They recorded EEG data from 124 brain areas of participants as they performed different cognitive tasks on a robot simulator. EEG-based MCI was found to be sensitive to changes in the subject's mental workload. Their study's results demonstrated that the EMCI index could effectively differentiate between MCI and HC groups, with accuracy rates of 89.09 % and an F1 score of 89.44 % in the beta frequency band.

3.2. AD detection

In this section, we report the papers that used deep learning methods for EEG-based AD diagnosis (Table A. 3). Deep learning techniques were used for both feature extraction [17] and classification [40,41,42,43]. [44] investigated various biomarkers, including EEG signals, prioritizing non-invasiveness, cost, and portability. Applying multiscale analysis and embedding space theory to EEG-based brain functional networks, they were able to classify AD patients versus healthy controls with 98 % accuracy using traditional classifiers. [45] developed a deep learning model that could distinguish AD from healthy aging adults based on signal inputs acquired using a two-lead ambulatory EEG system that recorded sleep-related data. Compared with healthy controls, AD patients spent less time in slow-wave sleep; other sleep stages were not significantly different between the two groups.

Among the deep learning architectures, convolutional neural networks (CNNs), transfer learning models, and generative adversarial networks (GANs) have gained popularity for AD diagnostic applications. CNNs are adept at image analysis and can automatically learn intricate patterns from EEG signals. Moreover, they can extract features and produce highly accurate results when combined with advanced graph-based networks [46,47,48,49]. [50] used CNNs to extract spatiotemporal features from multi-channel time series EEG signal data efficiently. However, they formatted the EEG signals collected from multiple brain regions into a 2D array. This allowed them to analyze the interrelationships between these areas like in an image. Their model attained 100 % classification accuracy for AD detection While such complex models perform well, the results generated are unclear. In developing their CNN-based AD detection model, [48] incorporated explainability and interpretability by using saliency maps to visually highlight components of the EEG signals (e.g., which frequency bands and modulations) that contributed the most to model classification. This combination of CNNs and saliency maps not only improved diagnostic accuracy but also provided a clearer understanding of the model's decision-making process. In a recent study, [51] used CNNs to classify AD patients into two groups based on severity. Their CNN model used multiple layers, including three convolutional and two fully connected layers, to process 2D matrices extracted from multi-channel EEG signals, as in [50]. From a database of more than 650 patients from five different hospitals, they classified AD patients into moderate and advanced with over 97 % accuracy.

Pre-trained on ImageNet data, transfer learning AD detection models do not need to be trained from scratch and only require fine-tuning to the specific EEG dataset. [52] employed AlexNet on resting-state EEG signals to classify AD versus MCI versus MC healthy aging, attaining over 98 % three-class classification accuracy. This study showcased the potential of transfer learning for AD detection, which paved the way for the development of more advanced transfer learning models. [53] used ResNet-50 to detect AD using EEG signals. In their model, raw EEG data were first pre-processed using principal component analysis to remove noise and then decomposed using wavelet transforms to extract statistical features. These features were converted into 2D plots, which then served as input to ResNet-50. To obviate the need for fine-tuning of pre-trained networks, some researchers proposed novel transfer learning architectures trained specifically for AD detection to address the specific needs of the application. [54] proposed EEGAlzheimer's Net, a transfer learning-based attention LSTM to handle non-linearities in EEG signals. The architecture combined spatial and temporal feature extraction using CNNS and recurrent neural networks (RNNs), respectively, with a transformer-based architecture to detect AD. Model accuracy was 99.85 %. To address the spontaneous and highly variable nature of EEG signals, [55] proposed Adazd-Net, which used adaptive flexible analytic wavelet transforms to capture subtle changes in EEG signals associated with AD. They then incorporated Shapely explanations, Morris sensitivity analysis, and local interpretable model-agnostic explanations (LIME) by depicting the contributions of each feature towards decision-making and providing explainability and interpretability to the model. Model accuracy was 99.85 %.

Several researchers have developed GANs and encoder-based models for EEG-based AD diagnosis (but less so for MCI diagnosis). To address limited and imbalanced EEG training data in their AD detection models, [56] and [57] used GANs to generate synthetic data points very similar to real data, which facilitated model training and improved model accuracy. To improve the quality of synthetic data and produce explainable predictions, [56] used Wasserstein GAN to minimize the earth-mover distance between real and synthetic data distributions. [57] combined GANs with a Marine Predator Algorithm to optimize feature extraction and improve accuracy. Unlike GANs, which synthesize data, encoder-based models extract features and are widely combined with CNNs, LSTMs, and transformers for AD detection. [58] proposed DICENet, a convolutional-transformer-based encoder architecture, for AD diagnosis. DICENet comprises two parallel convolutional blocks that perform dimensionality reduction, yielding outputs that are fed to a transformer-based model to make the final prediction. [59] used spatiotemporal autoencoders with CNNs and LSTMs to analyze brain dynamics as assessed by EEG in AD patients, attaining over 96 % AD classification accuracy. They observed that their model was robust against different instabilities of EEG signals and that the brain state trajectories of AD patients manifested as ring-manifolds, distinguishing them from controls.

Like in the case of MCI detection, several recent studies [60,61,62] have explored the use of multimodal data, information fusion, and model fusion for EEG-based AD detection. These approaches typically have high robustness and generalizability and produce better results than models that use data from one source alone. [63] analyzed EEG data in combination with functional near-infrared spectroscopy (fNIRS), which observed significantly better results in AD versus healthy classification performance than models trained on either EEG or fNIRS data alone. [64] studied the impacts of AD on spontaneous brain activity by using a thermodynamics-based framework to map the asymmetry of brain dynamics with time using a multimodal dataset of functional magnetic resonance imaging (fMRI) and EEG signals collected from AD patients and healthy controls. They found that AD correlated with a breakdown of temporal irreversibility at global, local, and network levels, and across multiple oscillatory frequency bands. The limbic, frontoparietal, default mode and salience networks were particularly affected, while temporal irreversibility was linked to cognitive decline in AD and gray matter volume in healthy controls. [65] combined multimodal data with ensemble deep learning models. From EEG signals and fNIRS collected concurrently during cognitive tasks, features derived from both modalities were optimized using a Pearson correlation coefficient-based feature selection strategy. Using a hybrid system that combined deep learning methods and decision analysis techniques, they accurately classified subjects into healthy, MCI, and two levels of AD severity, showcasing the potential of integrated data to refine medical assessments. [66] built an ensemble model combining multiple 2D-CNNs to capture intricate patterns in EEG images. Each CNN model serves as an individual classifier, and the predictions from all CNNs are aggregated to form the final decision. In so doing, the model mitigates inherent model biases, attaining 97.9 % accuracy, which outperformed the individual models.

[67] used the state-of-the-art vision transformers to screen for AD patients. EEG images are first converted into scalograms and fed to the vision transformer model. These models treat each segment of the EEG scalogram as part of a larger picture, assessing not only local features but also how these features relate to others across the entire image, allowing for comprehensive extraction of spatial and temporal patterns. Further, the self-attention block allows the capture of complex long-range dependencies in data.

3.3. MCI+AD (MCI and SD) detection

In this section, we describe the studies that used EEG recordings and deep learning for MCI+AD detection (Table A. 4), which encompassed three-class classification of healthy versus MCI versus AD, binary classification of MCI versus, or progression of MCI to AD. Deep learning setups like artificial neural networks can help eliminate noise and identify invariant features within these signals [68]. In addition, deep learning extracts and selects features automatically, which can either be classified by the deep model or be fed to standard ML classifiers. In their deep model, [68] used traditional ML algorithms like naïve Bayes for classification, attaining 98.25 % accuracy for three-class classification. Other researchers have also adopted this approach [69,70,71,72]. [69] used fast Fourier and continuous wavelet transforms to identify specific bands in the EEG signals that capture the most important features that discriminate AD versus MCI versus healthy controls.

Inputting deep learning-extracted features to a k-nearest neighbors (KNN) classifier, which is simple yet able to handle non-linear data effectively, they attained 99 % model classification accuracy. [72] also used KNNs for classification, but performed the feature selection differently. They introduced a technique called iterative filtering decomposition to decompose EEG signals into intrinsic mode functions. From here, four crucial features–power spectral density, Tsallis entropy, variance, and fractal dimension–are extracted. Including other cognitive tests to enrich these features, they attained 92 % accuracy. The importance of efficient feature selection increases as the number of classes in the classification task increases. Apart from neural networks and Fourier transformations, ANOVA and regressions have also been used for feature selection in such cases [70]. In their model, the EEG features obtained using ANOVA and Ridge regression reflected phase, spectral, and temporal characteristics during rest and memory-encoding states. With such comprehensive feature extraction, they were able to successfully classify the participants into four groups–AD, amnestic MCI, non-amnestic MCI, and subjective cognitive decline–with 93.1 % accuracy.

While the works above involved deep learning for feature extraction/selection only, recent works have also applied deep learning for the classification of MCI and AD from EEG signal data [73,74,75,76,77]. [78] built an attention-based technique to distinguish MCI from subjective cognitive decline using resting state EEG signals. Owing to its multi-head attention, the transformer architecture is particularly efficient at handling temporal dependencies. By pre-processing EEG signals and extracting relevant frequency bands such as delta, theta, alpha, and beta, the model achieved an area-under-curve (AUC) of 0.807, demonstrating the potential of deep learning models for MCI detection. Deep learning and transfer learning models often outperform ML classifiers. [75] built a deep neural network model called CEEDNet that consisted of different state-of-the-art models like VGG, ResNet, and vision transformers to detect MCI versus AD patients using spatial and temporal EEG signals and attained an AUC score of 0.9. While these models are extremely powerful, data augmentation and extensive pre-processing are often required to prevent this. [79] investigated the importance of data augmentation and its impact on transfer learning approaches. By performing data augmentation on their data, they were able to improve the classification accuracy of transfer learning models like ResNet by over 5 %.

Among diverse deep learning architectures, CNNs are the most common for medical image analysis, often attaining classification accuracies exceeding 90 % [80,81,82]. [77] used CNNs to distinguish AD from healthy cognitive aging using EEG data and attained 92.5 % model accuracy. Of note, their work exploited CNNs in two ways. First, the CNN model allowed them to bypass the complexities of traditional feature engineering, directly identifying patterns in EEG signals associated with different dementia subtypes, such as AD, dementia with Lewy bodies, and idiopathic normal-pressure hydrocephalus. Second, the model's temporal resolution and ability to learn from short, overlapping EEG segments enable it to capture transient neural dynamics disrupted in dementia.

Recent papers have also taken advantage of CNNs by using them in combination with other models. [83] proposed a hybrid model that used CNN and bi-directional gated recurrent units to detect MCI. In this work, CNNs are cleverly used to extract both spatial and temporal data associated with EEG recordings. Specifically, the model uses small and large temporal CNNs to independently capture temporal aspects of the EEG data, which are then combined. A spatial CNN further processes these combined features to exploit spatial relationships between different brain regions. Such a model is highly robust and [83] was able to detect MCI patients with 99 % accuracy despite using images that had not been pre-processed. [84] and [85] used graph convolutional networks (GCNs) to construct brain functional networks from EEG data for MCI and AD detection. GCNs effectively capture topological structures and neural interactions within EEG-derived functional connectivity, enabling the extraction of significant features and patterns associated with different stages of dementia. Here, the use of graph theory metrics enhances diagnostic precision, highlighting the potential of GCNs in clinical applications for early and accurate MCI, AD, and dementia diagnosis.

The performance of deep models can potentially improve with the incorporation of multimodal data. Notably, deep architectures can process large amounts of data and extract useful features from all different sources. As a result, building functional connectivity frameworks, using information fusion and developing diagnostic tools to analyze EEG signals automatically, have become popular [86,87,88]. [89] explored the use of complex tensor factorization with the PARAFAC2 model for estimating brain connectivity from EEG data. Their EEG model was built to effectively decompose EEG data into meaningful scalp components that are described by spatial, spectral, and complex trial profiles. They derived a new connectivity metric based on the complex trial profiles of the extracted components and showed that PARAFAC2 outperformed other traditional tensor analysis methods like PARAFAC and MVAR-ICA.

While dealing with feature fusion and multimodal data, one needs to be wary of incomplete data as it is extremely difficult to obtain complete data from multiple sources for the same subject. [76] proposed a method to handle incomplete multimodal data using a disease-image-specific deep learning framework that integrates image synthesis and disease diagnosis into a unified process. It comprises a disease-image-specific network (DSNet) for modelling disease-image specificity and a feature-consistency generative adversarial network (FGAN) for imputing missing images. While DSNet captures disease-relevant information from whole-brain scans, the FGAN module is used for synthesizing missing images while ensuring feature consistency. By training DSNet and FGAN together, they generated synthetic diagnosis-oriented images that achieved state-of-the-art performance in detecting AD as well as the progression of MCI to AD. [90] used a combination of EEG, eye tracking, and behavioral data to provide a cost-effective and noninvasive diagnostic alternative to traditional clinical methods, which are often expensive and require specialized expertise. Using domain adversarial neural networks and data augmentation, they attained accuracies of 88.81 % and 100 % for MCI and AD diagnoses, respectively.

4. Discussion

Deep learning has burgeoned due to recent exponential improvements in computational power, data availability, and algorithmic innovations. This has a direct impact on its applications in detecting MCI and AD based on EEG recordings over the last couple of years (Figure). In the following sections, we will describe some of the best models in each of the three domains (MCI detection, AD detection, and MCI+AD detection) and highlight pressing aspects like multimodality and explainability that have gained the spotlight over the last couple of years.

Fig. 5 presents the scalable workflow for diagnosing cognitive impairments. By combining EEG data with deep learning and cloud technology, this system enables automated and accessible detection of normal aging, mild cognitive impairment, and Alzheimer's disease, potentially leading to earlier and more accurate diagnoses.

In the following sections, we will describe some of the best models across three key areas: MCI detection, AD detection, and MCI + AD detection. Each domain addresses different aspects of neurodegenerative disease identification, providing unique insights into how deep learning can be applied to EEG data. We will also discuss two critical areas that have gained prominence recently: multimodality and explainability, both of which have substantially impacted the development of more reliable and transparent AI models.

Multimodality refers to integrating multiple data sources, such as EEG, MRI, and PET scans, to enhance the diagnostic accuracy of deep learning models. While EEG provides valuable temporal information about brain function, combining it with structural data from MRI or PET scans offers a more comprehensive view of the brain, leading to improved detection of MCI and AD. This approach has become increasingly popular as it helps to overcome the limitations of single-modality analysis, providing a richer understanding of neurodegeneration.

Explainability, on the other hand, has emerged as a crucial factor in building trust with clinicians and patients. Traditional deep learning models are often considered "black boxes" that deliver predictions without offering insight into how those predictions were made. Recent advances in explainable AI (XAI) techniques, such as SHAP, LIME, and GradCAM, now allow researchers to interpret the decision-making process of these models. By highlighting which EEG features or brain regions contribute most to a prediction, explainable models can provide clinicians with clearer insights into the workings of the model, making them more likely to be adopted in clinical settings.

In the following sections, we will explore the top-performing models in each of the three domains and how these models incorporate multimodal data and explainability techniques to improve the detection of MCI and AD.

4.1. *Key findings and discussion of the best models*

From our review of 74 papers, we found that CNNs are widely adaptable and perform exceedingly well in different tasks ranging from feature extraction to classification. They can also be used as ensembles with other CNNs or deep learning architectures. Transfer learning models like ResNet have performed the best in detecting MCI and AD and require minimal training time. Coming to MCI+AD detection, feature extraction is the most crucial step. This is because it might be difficult to distinguish between patterns of MCI and AD patients compared to MCI/AD patients versus healthy controls. In all three cases, multimodal models outperformed all others and demonstrated the potential to use information from different sources to detect these diseases at an early stage.

In the context of detecting MCI, four papers reported accuracy scores (Fig. 6). Among these, the ResNet model by [34] performed the best. EEG recordings were first processed to obtain low-order functional connectivity (LOFC) scores, which helped quantify the interactions between different brain regions. The LOFC measurements from four frequency bands were combined to create a multi-channel input for the ResNet model. Because of ResNet's ability to capture both local and hierarchical features in the data, it identified patterns specific to MCI and analyze how different regions of the brain connected and communicated across various frequency bands. In so doing, the model attained accuracy of up to 100 %.

In AD detection, we have observed that different authors quantified their results using different metrics. These results are summarized in Figs 7- 9 in terms of accuracy, AUC, and F1 scores. Due to their well-established ability to extract features and perform classification based

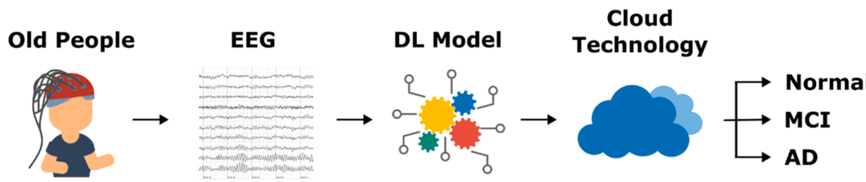

Fig. 5. Pictorial summary of MCI/AD/MCI+AD detection using deep learning and EEG signals.

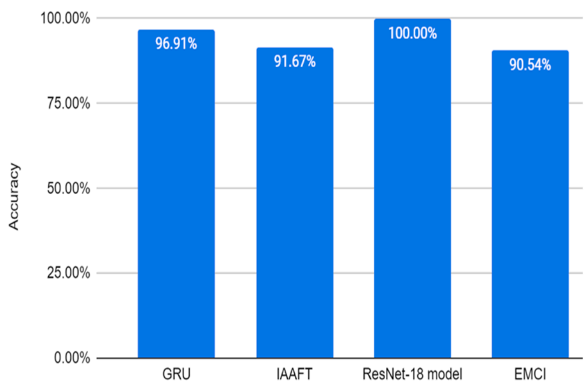

Fig. 6. Accuracy scores for models used in detecting MCI.

on image data, many studies have used CNN-based models to detect AD. Their work highlights the predictive ability of deep learning models in an ensemble setup. The paper utilizes an ensemble approach where they train five 2D-CNNs independently on the same EEG dataset after pre-processing and noise removal. Building an ensemble of five independent CNNs allows each constituent model to independently identify and learn from different features, potentially capturing a broader spectrum of diagnostic signals. To make the final prediction, the outputs of each individual CNN are aggregated to form a single, more accurate prediction. This reduces the chances of overfitting and mitigates the risk of any model introducing its bias into the prediction, thus guaranteeing increased generalizability and robustness.

With respect to MCI+AD detection, several machine learning models performed exceedingly well, achieving accuracies of over 90 % (Fig. 10). Once again, different metrics were used to quantify model performances (Figs. 11 and 12) . As discussed earlier, [69,90], and [83] are a few examples, all of whom achieved over 99 % predictive accuracy. [86] produced a similar work by developing a tool called a lacsogram to characterize MCI and different stages of AD. Their work relies on elaborate signal processing. EEG signals are first decomposed using discrete wavelet transforms into the delta, theta, alpha, beta, and gamma sub-bands since each of these might determine AD differently. Lapstral and cepstral analyses are conducted on these signals and different

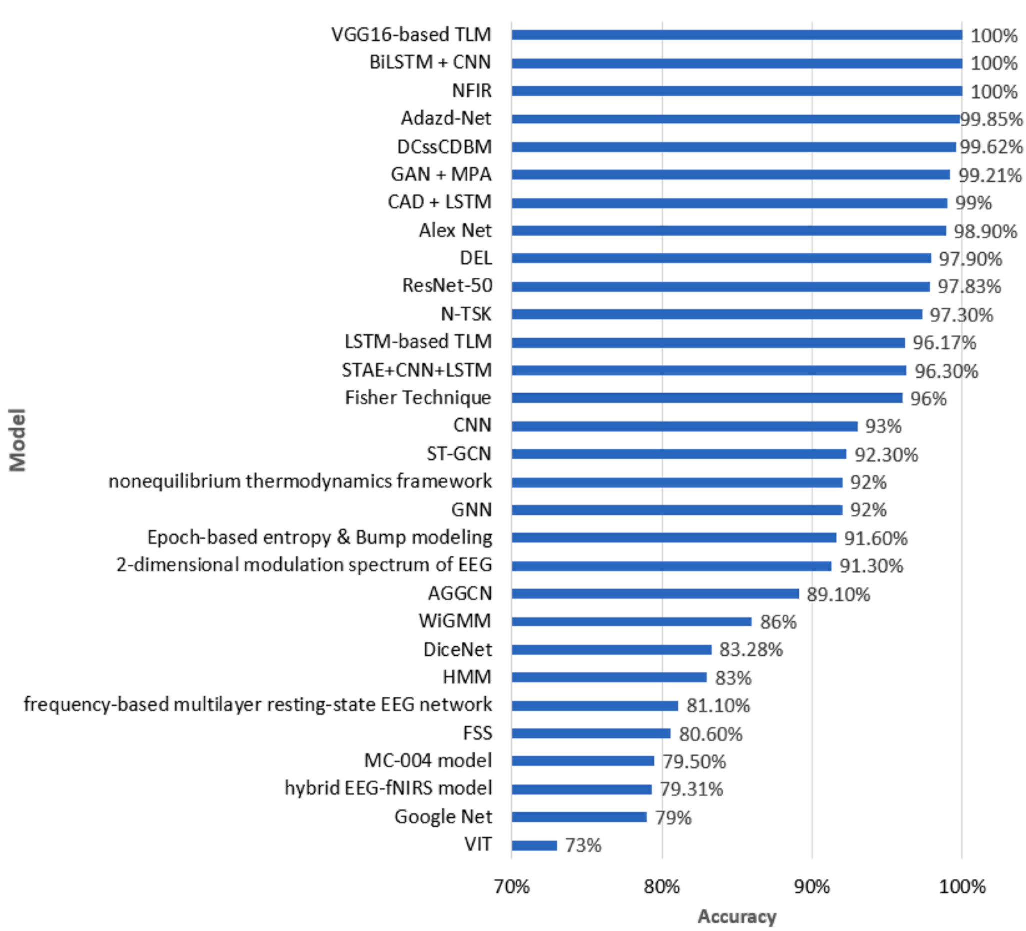

Fig. 7. Accuracy score for the models used in detecting AD.

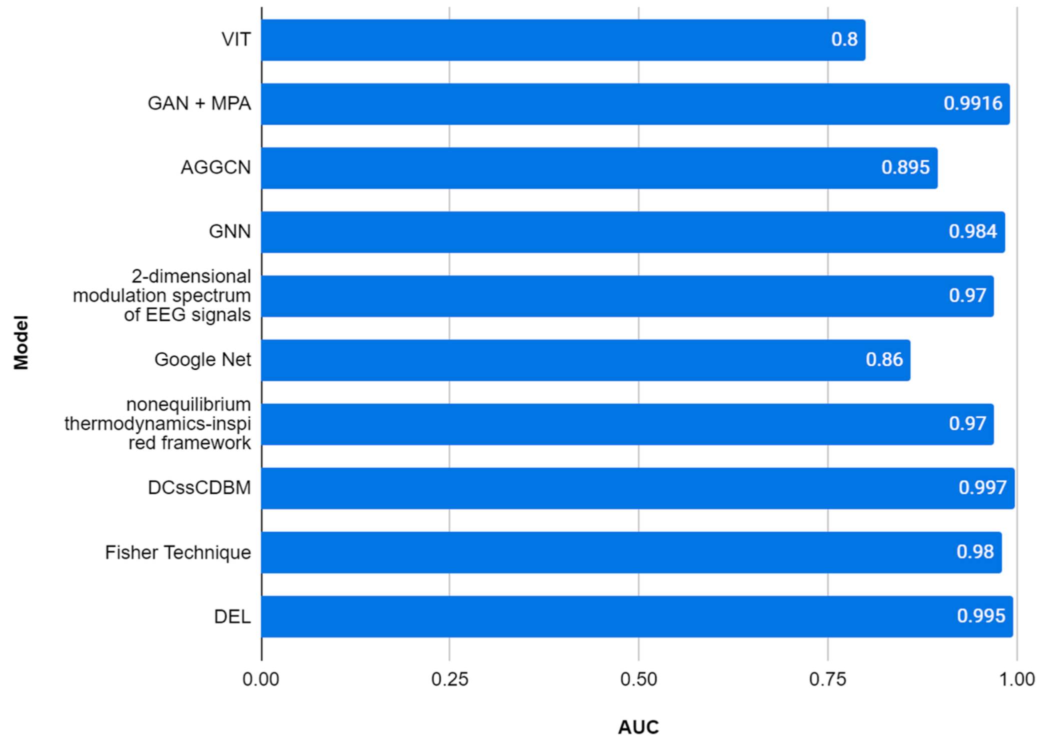

Fig. 8. AUC scores for the models used in detecting AD.

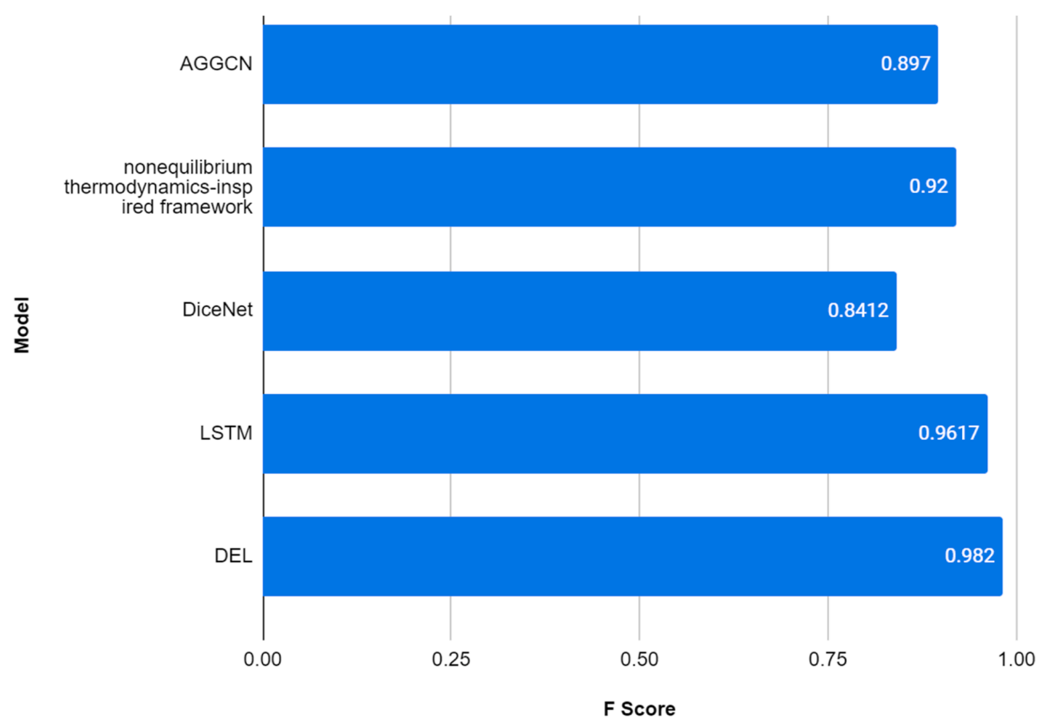

Fig. 9. F1 scores for models used in detecting AD.

distance measures are used between these lapstral and cepstral coefficients of different sub-bands. These distance measures then help statistical and ML models quantify the differences between the EEG patterns of healthy controls and patients with different stages of AD.

While these results highlight the high performance of deep learning models, collectively these works suggest that multimodal and ensemble models always stand out and perform significantly better than their counterparts and that performing elaborate signal processing, data augmentation, and feature selection can help build reliable, robust models, in all the cases of MCI, AD, and MCI+AD detection. In particular, ResNet has been the best-performing model for MCI classification, while an ensemble of CNNs has been best suited for AD detection. In the more complex case of MCI+AD detection that typically deals with a three-class classification problem, feature extraction is key, and the

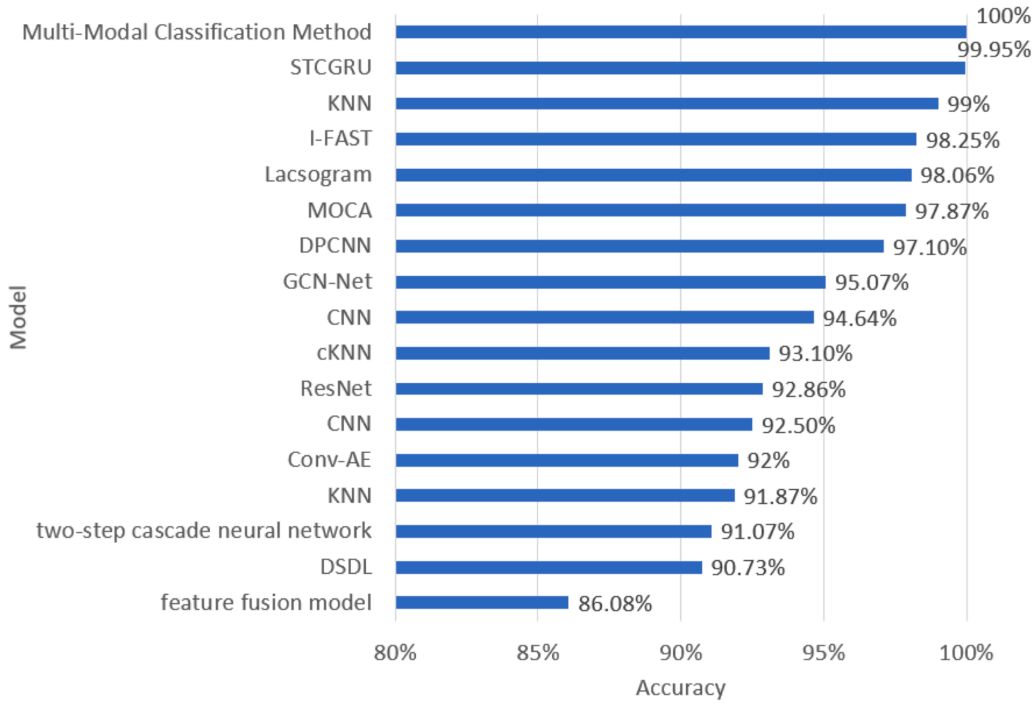

Fig. 10. Accuracy scores for the models used in detecting MCI+AD.

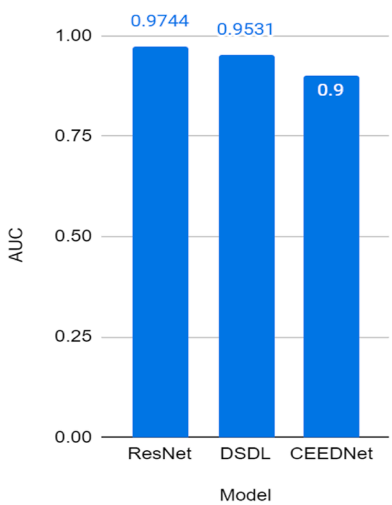

Fig. 11. AUC scores for models used in detecting MCI+AD.

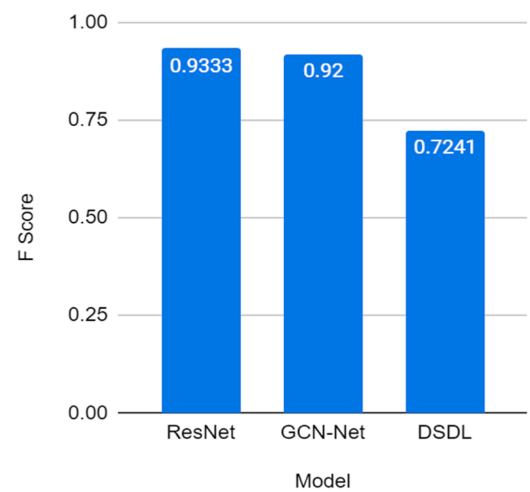

Fig. 12. F1 scores for models used in detecting MCI+AD.

model's predictive ability depends on how well it distinguishes between MCI and AD features which are likely to be more similar as compared to that of healthy controls. Further, while accuracy and AUC scores have been commonly used metrics to report model performance, sensitivity and specificity scores have also been used in a few works (Fig. 13). Moving ahead, we urge the research community to adopt a single reliable metric, such as the accuracy of F1 scores, to allow for model comparison and reproduction of results.

The Multi-Modal Classification Method achieves the highest sensitivity (100 %), meaning it correctly identifies all AD cases. STCGRU and KNN models also perform excellently, with sensitivity scores close to 99 %. Lacsogram, MOCA, and other models also show very high sensitivity scores (>97 %), suggesting robust performance in identifying AD patients. The Feature Fusion Model has the lowest sensitivity (86.08 %), indicating it misses more AD cases than other models.

Adazd-Net shows perfect specificity (100 %), meaning it successfully identifies all non-AD patients. DCssCDBM model, GAN + MPA, and DEL also exhibit very high specificity (>99 %), making them excellent for avoiding false positives. The ViT model shows the lowest specificity (57.10 %), indicating it struggles with correctly identifying non-AD patients, resulting in a high rate of false positives.

MOCA achieves perfect sensitivity (100 %), correctly identifying all MCI+AD cases, making it the best-performing model for combined detection. ResNet (93.33 %) and DSDL (91.05 %) also perform well, but models like PCA + FBCSP show slightly lower sensitivity (87 %).

MOCA again leads in specificity (97.38 %), meaning it avoids false positives better than the other models. ResNet and DSDL also perform well with specificity scores above 88 %, but PCA + FBCSP shows the lowest specificity (80 %), meaning it is less effective in identifying non-MCI/AD patients compared to the top models.

The findings of the reviewed models according to Figs 13 -- 16 are:

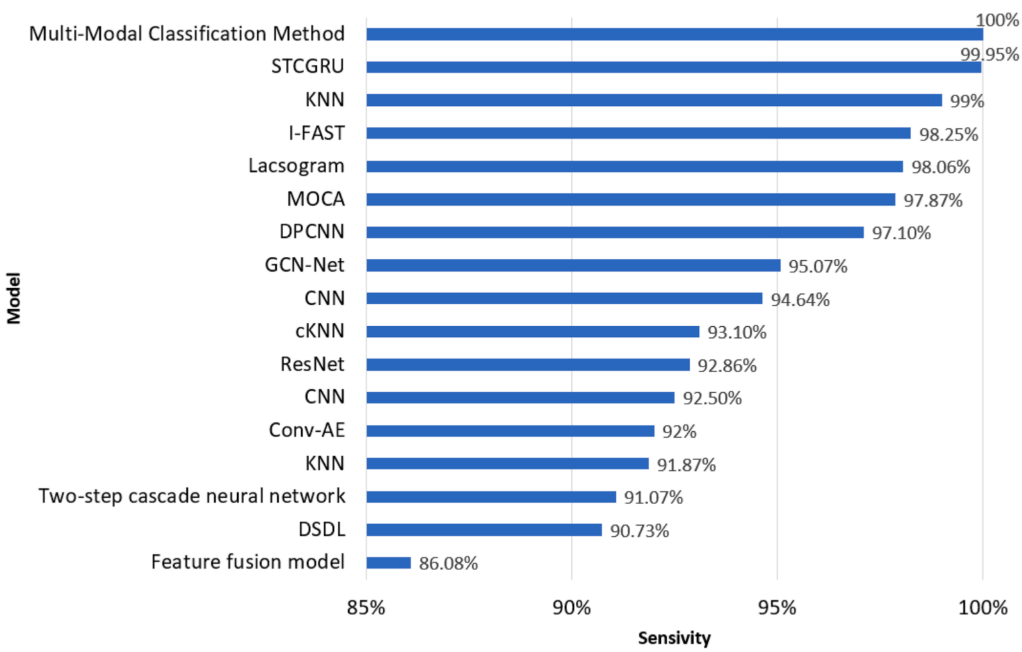

Fig. 13. Sensitivity scores for models used in detecting AD.

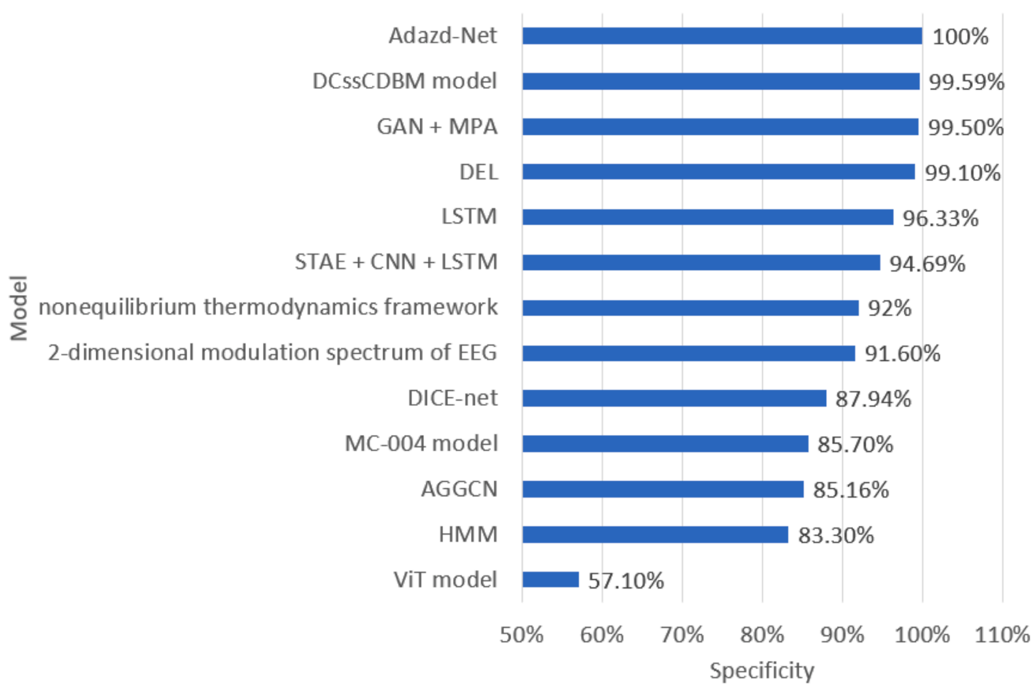

Fig. 14. Specificity scores for models used in detecting AD.

- Models like MOCA, Multi-Modal Classification Method, Adazd-Net, and STCGRU show consistently high sensitivity and specificity, making them the best choices for both AD and MCI+AD detection.
- While models like the Multi-Modal Classification Method and MOCA excel in sensitivity, others like Adazd-Net and GAN + MPA stand out for their high specificity.
- Lower-performing models such as the Feature Fusion Model and ViT model may require further refinement to enhance both sensitivity and specificity, ensuring they can reliably detect AD and MCI cases without producing false positives.

4.2. Benefits and challenges of multi-modality

With the deep learning models processing large amounts of data easily and identifying the crucial features from all of them, multimodal data have gained popularity recently. Since these datasets involve integrating information from different sources about the same person, they provide a holistic view of the person's condition, thus enabling more reliable predictions. Further, since information is available at different levels about the same individual, extracting different features from each of these would not only enable early detection of the disease, but also open possibilities for personalized treatments. One such

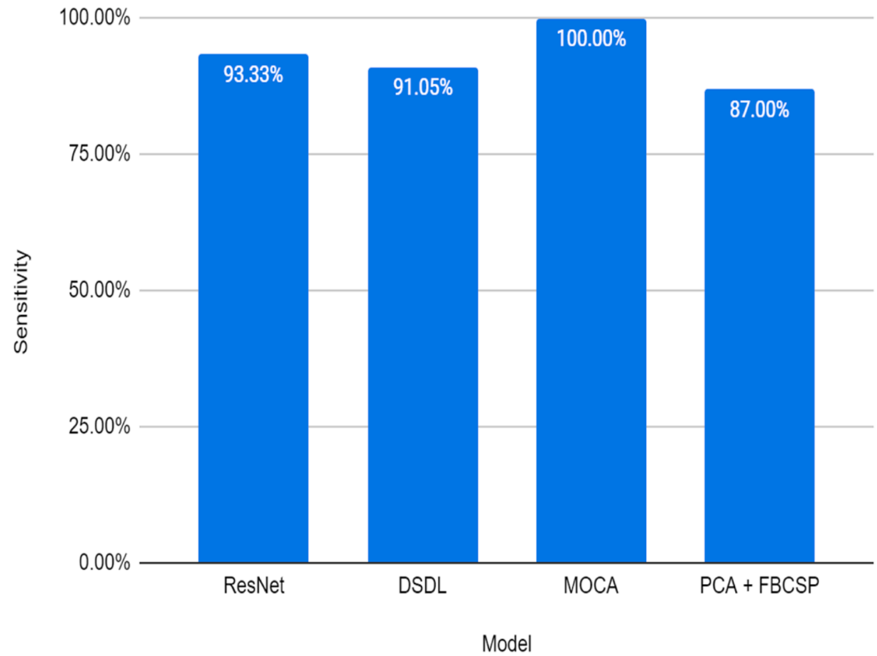

Fig. 15. Sensitivity scores for models used in detecting MCI+AD.

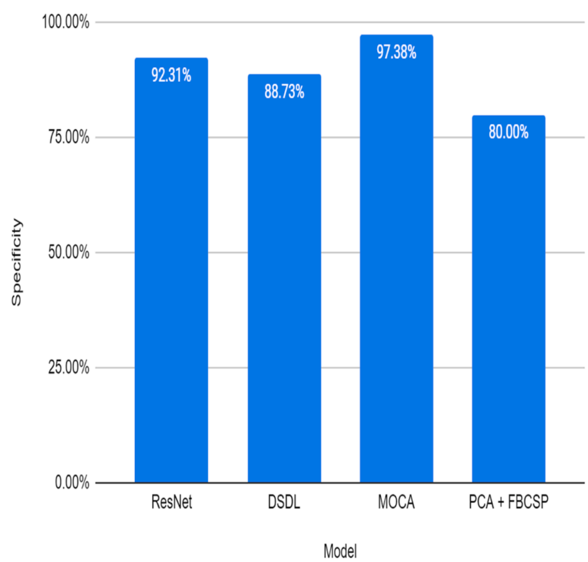

Fig. 16. Specificity scores for models used in detecting MCI+AD.

example is the study by [62]. They used EEG recordings in combination with PET scans, sleep measurement techniques, and cognitive testing to study MCI and AD patients.

Among the 74 studies reviewed in this work, nine of them used multimodal data, among which five were published in 2022 or later. Among these studies, we found that EEG data was often used in combination with other neuroimaging data like MRI or eye movement data. One such example is that of [64], who used EEG in combination with fMRI data to study the temporal irreversibility of brain dynamics in AD. In this process, while fMRI provided spatial details about brain activity, they exploited EEG signals to get insights into temporal changes. By studying the irreversibility of time series signals across the two modalities, they found that AD patients showed significantly lower levels of complexity and temporal asymmetry in brain activity, which could be a crucial feature in differentiating them from healthy individuals. While [62] and [64] used different neuroimaging recordings, [36] and [90] used eye-movement data along with EEG recordings to detect MCI and AD. Tracking the movement of eyes can give insights into the individual's cognitive load and response accuracy, which may be diminished in the presence of MCI/AD. For example, fixation and saccade tasks can help measure how quickly someone can redirect their gaze. Difficulty in performing such a task can be a sign of execution dysfunction common in MCI and AD patients. As a result, both [36] and [90] noted that tracking eye movement in the presence of visual stimulus can help ameliorate the performance of models that use EEG recordings alone.

While multimodal data allows for the training of more robust and generalizable models that are not overfitted to any setting, they come with a separate set of challenges. Integrating data from multiple modalities into one deep learning model is a complex task and requires technical expertise. Next, the complexity of multimodal models also requires more computational and storage resources to process all the information carefully. Moreover, obtaining information about the same individual from multiple data sources is extremely difficult, tedious, and time-consuming. In such cases, incomplete data should be dealt with carefully [76]. Finally, interpreting the results obtained from such models is extremely difficult because it is very challenging to identify which features from which data source were instrumental in making a prediction.

4.3. Explainable AI and interpretability of deep learning models

In the context of medical image analysis, deep learning models appear to lack interpretability and explainability aspects because their decision-making process involves complex, nonlinear computations across multiple layers, making it difficult to trace how different inputs and features affect prediction. Several works have highlighted the importance of building explainable and interpretable deep learning models in medical image analysis [91,92]. It is extremely crucial to build such transparent models to gain medical practitioners' trust and ensure the application of deep learning in practice [93]. By highlighting the most discriminative features and ranking them in terms of their importance toward the final decision-making, such efforts add a layer of interpretability and support clinicians in decision-making [56]. This would not only provide enhanced diagnostic accuracy but improve scalability and facilitate personalized treatment.

In the context of explainability and interpretability, visualization is key. Visualization helps researchers understand how a model has made a particular prediction and what features were given importance during this decision-making process. For example, [55] used Shapley explanations, Morris sensitivity analysis, and LIME to provide interpretability. While Morris sensitivity analysis allows one to identify the input variables that are most influential, LIME and Shapley scores help understand which features contribute the most to decision-making. They do so by quantifying the contribution of each feature to the prediction of a specific instance. The visual attention mechanism adopted by [48] works similarly. By incorporating visual attention within neural networks, one can highlight those regions of the input EEG data that are most influential in making a prediction. Highlighting such features as specific time points or electrode readings provides relevant understanding for clinicians using them. [94] show a different approach to achieve interpretability. They represent EEG signals as a graph, where each node refers to an EEG electrode and each edge denotes functional connections and interactions of the brain. Following this, they build a gated graph convolutional model that allows for dynamic weighting of the edges and nodes in the graph. Interpretability is achieved by studying these weights since each edge corresponds to the brain's functional connections and priority is given to those nodes that carry more predictive power in predicting if a patient is healthy or affected by MCI or AD.

Several studies have attempted to make deep learning models more explainable by identifying specific visual features or input components that contribute most to the decision-making process in EEG-based MCI and AD detection. These features are critical in improving the diagnostic accuracy and usability of models in clinical practice.

The visual and EEG features that have been found to be most indicative of class discrimination between healthy controls, MCI patients, and AD patients include:

- Specific electrode readings from temporal, parietal, prefrontal, and hippocampal regions.
- Changes in functional brain connectivity observed through graph-based models.
- Important time points in EEG recordings, often in the alpha and beta bands.
- Frequency bands such as theta and delta, which have shown to correlate with cognitive decline.
- Spatial EEG topography, particularly in the occipital and frontal lobes, which helps distinguish AD patients.

By identifying and highlighting these discriminative features using explainable AI techniques like Shapley values, LIME, and Grad-CAM, researchers have made deep learning models more transparent. This interpretability is essential in clinical settings, where understanding which specific features contribute to diagnoses allows clinicians to make more informed decisions, ultimately enhancing the models' adoption in practice.

Among the 74 reviewed studies, considering that only four papers [55,94,48,56] have worked on building explainable and interpretable models, the working of all other papers remains a black box. We thus urge future researchers to expand on such works to bridge the gap between using deep learning for MCI and AD detection in theory and practice.

Fig. 17 highlights the need for more public dataset usage and disclosure in the field of EEG-based MCI and AD detection. Coupled with the call for more explainable models, this figure suggests a need for greater openness and accessibility in future research efforts to foster collaboration and innovation.

4.4. Databases used

In this section, we discuss the databases used by the reviewed studies. Out of the 74 studies reviewed, 67 papers have explicitly cited the source of their data. Of these 67 studies, 49 used private datasets, while 18 used publicly available datasets. Fig. 17 and Fig. 18 depict the percentage of papers that used public and private datasets in each category.

Fig. 18 showcases the reliance on private datasets across AD and MCI detection studies. This trend points to a critical need for more publicly accessible datasets to support transparency, reproducibility, and innovation in this domain. The scarcity of public datasets limits the research community's ability to build on existing studies, highlighting a gap that future research should aim to address. It is crucial to highlight that compared to studies that have used publicly available datasets, almost 2.5 times the number of studies have used private datasets. This indicates the lack of large public datasets easily accessible to researchers

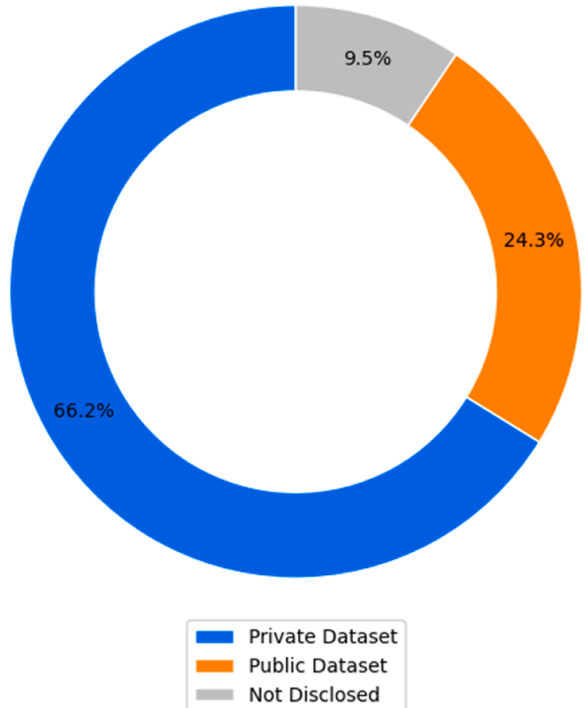

Fig. 17. Distribution of private and public databased used.

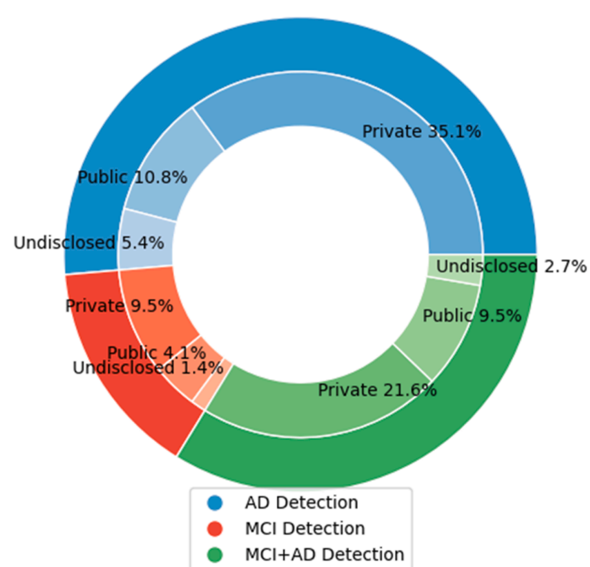

Fig. 18. Distribution of databases based on the detection task.

in this domain. Analyzing this further, Fig. 19 sheds more light on this limitation. Among the 67 studies that cited their data sources and statistics, 21 works used datasets consisting of 50 or fewer participants, and more than half the studies were based on datasets consisting of 100 or fewer participants. To build models that are robust, generalizable, and capable of catering to different individuals in the real world, it is extremely important to release publicly available datasets consisting of a large number of data points.

Fig. 19 showcases the variability in sample sizes across studies in AD, MCI, and combined detection (AD + MCI) categories. While AD detection studies are more likely to use larger datasets, MCI detection studies often rely on smaller sample sizes. The figure underscores the need for larger, more diverse datasets, especially in MCI detection, to improve the generalizability and robustness of findings in this important area of neurodegenerative research.

Table 1: Details of publicly available datasets for MCI/AD detection using EEG recordings. Among these publicly available datasets, some of the largest ones include the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset, the Sina and Nour Hospital dataset [95], and the Florida State University dataset [96]. The Florida State University dataset records EEG signals from healthy controls in two settings. The 96 participants are divided into two categories of 46 individuals each. The first category consists of 24 AD patients and 24 healthy controls whose EEG signals are recorded with their eyes open. The second category consists of a similar split of participants, but the EEG signals are recorded with their eyes closed. Recent studies like [53] and [66] have used this dataset to detect AD. The Sina and Nour Hospital dataset [95] has also been widely used over the last couple of years [33,32,82,83]. It is a relatively smaller dataset consisting of EEG recordings of 27 subjects (16 normal and 11 MCI participants) aged between 60 and 77. All EEG signals were recorded continuously using 19 electrodes in the morning with the participants resting comfortably in a quiet room with closed eyes. Because of the targeted age group of the participants, this dataset can be used to train models to distinguish between naturally aging individuals and MCI patients. Although used in several research works, these databases are limited by the availability of EEG recordings and cater only to the study of MCI or AD. On the other hand, the ADNI dataset consists of a huge EEG dataset along with MRI and PET scans for studying AD and MCI. The EEG images in the dataset are used to assess brain activity patterns that may differentiate between normal aging, MCI, and AD. This multimodal approach facilitates comprehensive research in studying potential biomarkers of both AD and MCI. It not only provides multimodal information, but also facilitates the study of MCI, AD, and healthy controls simultaneously.

Table 1 provides a comprehensive overview of publicly available EEG datasets used for MCI and AD detection, which are critical for advancing research in this field. These datasets offer a range of sample sizes, conditions (MCI, AD, dementia), and multimodal features (e.g., MRI, PET), providing valuable resources for training and validating machine learning models aimed at early detection and progression monitoring of neurodegenerative diseases. The availability of these datasets fosters collaboration, transparency, and reproducibility in the research community.

Moreover, the deep learning models comparisons have been given in Table 2.

4.5. Future research work

To overcome the different limitations discussed earlier, we propose the following future research works:

1. *Availability of a huge public database*: It may be noted from this work that most of the research has been carried out using smaller private or public databases. We propose to have more public databases for researchers to develop accurate, robust, and faster DL models. The public databases developed using data collected from various countries and centers can aid in creating a robust model.
2. *Explainable AI and uncertainty quantification*: AI models perform like black boxes by diagnosing input data. It does not explain the process involved in obtaining output. Hence, explainable AI (XAI) can be employed to develop the confidence of clinicians and researchers. Such techniques can be employed in hospitals for the detection of various mental disorders and treatment. Techniques such as LIME, SHapley Additive exPlanations (SHAP), and Gradcam (Gradient-weighted Class Activation Mapping) have been developed to address the model explainability to explain the working of the generated model [106].
3. Uncertainty Quantification (UQ): Most developed AI models perform well using small databases. Their performances vary when subjected to real-world scenarios in the presence of noises due to changes in the

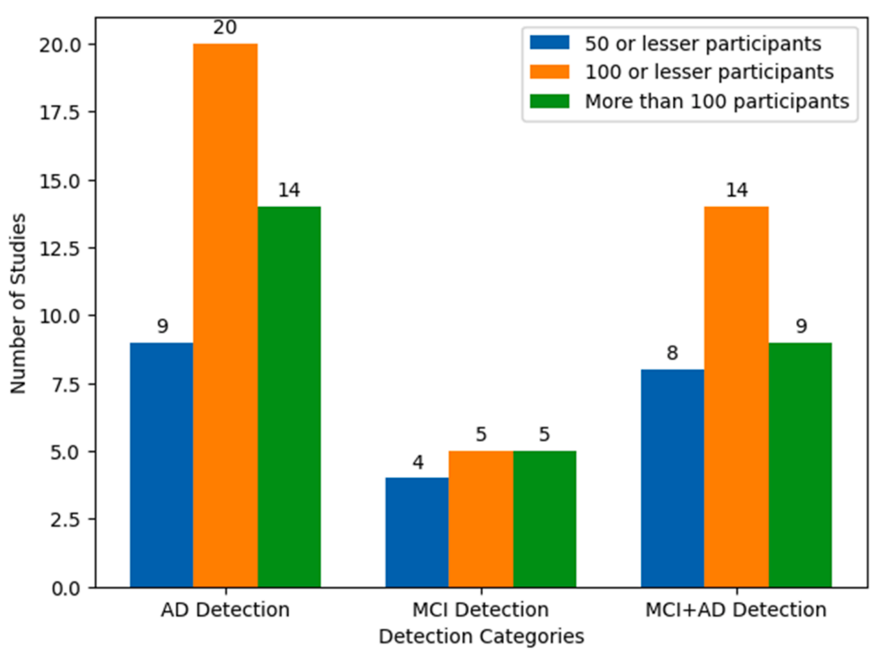

Fig. 19. Studies stratified based on the size of the dataset.

Table 1
Details of publicly available datasets for MCI/AD detection using EEG recordings.

Database Name/Citation/ Link	Studies using the database	Important characteristics
ADNI dataset (https://adni.loni.usc.edu/)	[76,35,64]	ADNI primarily focuses on MRI and PET imaging, but it includes EEG data for multimodal MCI and AD research.
Sina and Nour Hospital, Isfahan, Iran [95] (https://misp.mui.ac.ir/en/eeg-data-0)	[32,82,33,83]	EEG dataset of 27 subjects - 16 healthy controls and 11 MCI patients, all of whom are aged between 60 and 77 years.
[97]	[82]	Consists of 109 subjects - 7 MCI patients and 102 healthy controls
Chung-Ang University Hospital EEG – CAUEEG Dataset [75]	[75]	1379 EEG recordings from 1155 patients including normal (459), MCI (417), Dementia (311) classes.
The MCI and mild AD dataset – [98,99]	[79]	MCI Dataset - consists of 22 MCI subjects and 38 healthy controls. Mild AD Dataset - consists of 17 mild AD subjects and 24 healthy controls
VSTMBT memory task EEG Data [100]	[89]	EEG data collected based on the VSTMBT memory task from 23 MCI patients and 24 healthy controls
OpenNeuro Dataset – [101]	[102]	EEG dataset of 36 AD patients, 23 Dementia patients, and 29 healthy controls.
Scalp EEG dataset (https://osf.io/download/yhg9w/)	[50]	Information from 19 channels for 24 healthy individuals and 24 AD patients.
Hospital das Clínicas in São Paulo, Brazil [17]	[52,103]	EEG recordings from 31 mild AD patients, 20 moderate AD patients, and 35 healthy individuals.
RFGHCPLA Dataset [104]	[105]	EEG recordings from 39 subjects (20 healthy controls and 19 aMCI patients).
Florida State University dataset [96]	[53,66]	EEG recordings taken with the eyes of individuals opened and closed - 48 healthy controls and 48 AD patients.
Matouš Cejnek dataset - (https://figshare.com/articles/dataset/dataset_zip/5450293/1)	[53]	EEG recordings of 7 MCI patients, 59 AD patients, and 102 healthy controls.
AD Classification dataset – (https://github.com/tsyoshihara/Alzheimer-s-Classification-EEG/tree/master/data)	[54]	EEG data consisting of MCI patients, AD patients, and healthy controls.

data or tuning parameters. In such cases, UQ can be employed to quantify and mitigate the uncertainty in the data and model in the clinical environment [107,108,109].

4. *Information Fusion*: The performance obtained using EEG signals and DL models has been considered in this review. Physiological signals such as electrocardiogram (ECG), heart rate variability (HRV), photoplethysmography (PPG), or/and speech signals can also be considered with brain images. Such data fusion methods can improve the performance of the DL models with huge databases [110].
5. *Developing new DL models*: New DL models, such as deep attention mechanism models (Hafiz et al., [111]), graph convolutional neural networks (GCNN) [112], deep multi-task learning models (DMTLM) [113], federated learning (FL) models [114], and deep mutual learning models (DMLM) [115,116], can be explored to obtain accurate performance with huge databases. The attention mechanism helps to focus on the important portion of the input. The graph models are the structures fed as input to the DL models. DMTLM performs multiple trained tasks simultaneously. FL is a technology that obtains information about the data for new AI models without touching it. DMLM improves each other's performance among many networks.
6. *Data Standardization*: During the data collection of EEG and other physiological signals from various centers, due to the variation in the acquisition protocols and types of equipment, there is a possibility of changes in the magnitude of input data to the AI system. Such inconsistencies introduce errors in the AI systems. To overcome these problems, data standardization needs to be done before feeding to the AI system [117].
7. *Incorporation of Transfer Learning:* Transfer learning can be explored to address the issue of limited data availability in smaller databases. This method allows leveraging pre-trained models developed on larger, related datasets to improve model performance on smaller datasets. Transfer learning has shown promise in various fields, including medical imaging, where models trained on larger datasets can be fine-tuned for specific tasks like MCI and AD detection using EEG data.
8. *Personalized AI Models:* Developing personalized AI models tailored to individual differences in brain activity can significantly enhance the accuracy and reliability of MCI and AD detection. EEG signals exhibit high variability between individuals, so AI models that account for personalized baselines and patterns may yield better results than generalized models.
9. *Real-Time Detection and Monitoring:* Future research can focus on developing AI systems capable of real-time EEG analysis for continuous monitoring of MCI and AD progression.

5. Conclusion

This review paper has analysed the various models employed for AD, MCI, and (MCI+AD) categories from (2013-2024) using EEG signals. We have observed that the ensemble CNN model yields the highest for automated detection of AD, ResNet is effective for detecting MCI, and efficient feature extraction using CNNs is extremely crucial for (MCI+AD) detection. Further, we observed that multimodal datasets help build robust, generalizable, and high-performing models in all three cases.

The limitation of this work is that most of the studies have used smaller databases for the automated detection of classes based on EEG signals. We need to use large databases from many countries and various centers to overcome this limitation. Also, HRV or PPG signals can be extracted using wearable devices to develop DL models and can be used in home-based environments.

XAI and UQ must be employed to use the developed model in the clinical environment. The deep learning model needs to be developed using a huge, diverse data population belonging to various races and age groups.

Ethical approval

Not applicable.

Funding

The first author has received the University of Southern Queensland (UniSQ) Domestic PhD Research Scholarship (2023-2026) and the Research and Training Program (RTP) Scholarship funded by the Australian Government, both of which are greatly acknowledged.

CRediT authorship contribution statement

Madhav Acharya: Writing – original draft, Visualization, Validation, Methodology, Investigation, Funding acquisition, Formal analysis,

Table 2
Comparative results of the deep learning models.

Model	Architecture features	Strengths	Common applications	Performance highlights
CNN (Convolutional Neural Network)	Convolutional and pooling layers, fully connected layers	Strong feature extraction from spatial data, effective for image-like EEG inputs	Feature extraction and classification of EEG signals	Commonly used in AD detection, shows high performance in ensemble learning (e.g., 99 % sensitivity)
ResNet (Residual Neural Network)	Residual connections, deep architecture	Handles very deep networks without vanishing gradient issues, captures both local and global features	MCI detection, EEG data with multi-channel input	Achieved up to 100 % accuracy in MCI detection
LSTM (Long Short-Term Memory Network)	Recurrent architecture with memory cells and gates	Captures temporal dependencies, retains information across sequences	Sequential EEG data, temporal analysis	Paired with CNN for high specificity (96.33 %) in AD detection
GCN (Graph Convolutional Network)	Graph convolution layers, handles graph-structured data	Models spatial relationships between brain regions	EEG as brain connectivity graphs, MCI detection	High accuracy due to effective modeling of brain connectivity
Autoencoders (Conv-AE)	Encoder-decoder structure, unsupervised feature learning	Learns compressed representations of EEG signals	Feature extraction and dimensionality reduction	Achieved high sensitivity (92 %) in AD detection
Attention-based Models	Attention layers, context vectors	Focuses on important parts of the input data, enhances interpretability	MCI and AD detection, explainable AI	Models like STCGRU achieve sensitivity close to 99 %
Federated Learning (FL)	Distributed local models, central aggregation	Protects data privacy, enables training across multiple sources	EEG data from decentralized datasets	Promising for privacy-sensitive EEG data, reduces data sharing concerns
Multi-modal Models	Parallel input streams, fusion layers for multiple data types (e.g., EEG + MRI)	Combines complementary information from multiple sources	Combining EEG with neuroimaging, AD detection	Improved diagnostic accuracy through modality integration (e.g., 99 % predictive accuracy in combined MCI+AD detection)

Data curation, Conceptualization. **Ravinesh C Deo:** Writing – review & editing, Supervision, Resources, Project administration. **Xiaohui Tao:** Writing – review & editing, Supervision. **Prabal Datta Barua:** Writing – review & editing, Supervision, Resources, Project administration. **Aruna Devi:** Writing – review & editing, Supervision, Project administration. **Anirudh Atmakuru:** Writing – review & editing, Resources, Methodology, Investigation. **Ru-San Tan:** Writing – review & editing, Supervision, Resources, Methodology.

Declaration of competing interest

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

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.cmpb.2024.108506.

References

[1] M. Abd Al-Alim, R. Mubarak, N.M. Salem, I Sadek, A machine-learning approach for stress detection using wearable sensors in free-living environments, Comput. Biol. Med. 179 (2024) 108918.
[2] J. Hugo, M. Ganguli, Dementia and cognitive impairment: epidemiology, diagnosis, and treatment, Clin. Geriatr. Med. 30 (2014) 421–442.
[3] A. Wimo, B. Winblad, H. Aguero-Torres, E. von Strauss, The magnitude of dementia occurrence in the world, Alzheimer Dis. Assoc. Disorders 17 (2003) 63–67.
[4] R.C. Petersen, Mild cognitive impairment, Continuum: lifelong Learn. Neurol. 22 (2016) 404–418.
[5] S. Jongsiriyanyong, P. Limpawattana, Mild cognitive impairment in clinical practice: a review article, Am. J. Alzheimer's Dis. Other Dementias® 33 (2018) 500–507.
[6] A.M. McGrattan, Y. Zhu, C.D. Richardson, D. Mohan, Y.C. Soh, A. Sajjad, C. Van Aller, S. Chen, S.-M. Paddick, M. Prina, Prevalence and risk of mild cognitive impairment in low and middle-income countries: a systematic review, J. Alzheimer's Dis. 79 (2021) 743–762.
[7] P. dos Santos, C. Leide, P.F. Ozela, M. de Fatima de Brito Brito, A.A. Pinheiro, E. C. Padilha, F.S. Braga, d.S. de Paula, H. Carlos, C.B.R dos Santos, Alzheimer's disease: a review from the pathophysiology to diagnosis, new perspectives for pharmacological treatment, Curr. Med. Chem. 25 (2018) 3141–3159.
[8] A.C. Naj, G.D. Schellenberg, A.s.D.G. Consortium, Genomic variants, genes, and pathways of Alzheimer's disease: An overview, Am. J. Med. Genet. Part B: Neuropsychiatr. Genet. 174 (2017) 5–26.
[9] S. Cervantes, L. Samaranch, J.M. Vidal-Taboada, I. Lamet, M.J. Bullido, A. Frank-García, F. Coria, A. Lleó, J. Clarimón, E. Lorenzo, Genetic variation in APOE cluster region and Alzheimer's disease risk, Neurobiol. Aging 32 (2011) 2107, e2107-2107. e2117.
[10] E. Arnáiz, O. Almkvist, Neuropsychological features of mild cognitive impairment and preclinical Alzheimer's disease, Acta Neurol. Scand. 107 (2003) 34–41.
[11] E. Shatil, Does combined cognitive training and physical activity training enhance cognitive abilities more than either alone? A four-condition randomized controlled trial among healthy older adults, Front. Aging Neurosci. 5 (2013) 8.
[12] Y.F. Tai, P. Piccini, Applications of positron emission tomography (PET) in neurology, J. Neurol. Neurosurg. Psychiatry 75 (2004) 669–676.
[13] F. Massa, M. Bauckneht, E. Peira, C. Lapucci, A. Picco, S. Capitanio, D. Arnaldi, L. Roccatagliata, A. Chincarini, F. Nobili, Neuroimaging findings in mild cognitive impairment, PET and SPECT in Neurol. (2021) 367–425.
[14] J. Dauwels, F. Vialatte, A. Cichocki, Diagnosis of Alzheimer's disease from EEG signals: where are we standing? Curr. Alzheimer. Res. 7 (2010) 487–505.
[15] C.M. Michel, M.M. Murray, Towards the utilization of EEG as a brain imaging tool, Neuroimage 61 (2012) 371–385.
[16] G. Rodriguez-Bermudez, P.J. Garcia-Laencina, Analysis of EEG signals using nonlinear dynamics and chaos: a review, Appl. Math. Inf. Sci. 9 (2015) 2309.
[17] R. Cassani, T.H. Falk, Alzheimer's disease diagnosis and severity level detection based on electroencephalography modulation spectral "patch" features, IEEE J. Biomed. Health Inform. 24 (2019) 1982–1993.
[18] L. Lazli, M. Boukadoum, O.A. Mohamed, A survey on computer-aided diagnosis of brain disorders through MRI based on machine learning and data mining methodologies with an emphasis on Alzheimer disease diagnosis and the contribution of the multimodal fusion, Appl. Sci. 10 (2020) 1894.
[19] D.T. Nguyen, W. Li, P.O. Ogunbona, Human detection from images and videos: A survey, Pattern. Recognit. 51 (2016) 148–175.
[20] S. Siuly, Y. Li, Y. Zhang, EEG signal analysis and classification, IEEE Trans. Neural Syst. Rehabilit. Eng. 11 (2016) 141–144.
[21] A. Puce, M.S. Hämäläinen, A review of issues related to data acquisition and analysis in EEG/MEG studies, Brain Sci. 7 (2017) 58.
[22] S.-A. Sadegh-Zadeh, E. Fakhri, M. Bahrami, E. Bagheri, R. Khamsehashari, M. Noroozian, A.M. Hajiyavand, An approach toward artificial intelligence Alzheimer's disease diagnosis using brain signals, diagnostics 13 (2023) 477.
[23] P. Matthew, M. Anderson, Novel approaches to developing multimodal biometric systems with autonomic liveness detection characteristics, in: Intelligent Systems for Science and Information: Extended and Selected Results from the Science and Information Conference 2013, Springer, 2014, pp. 121–138.
[24] S. Chen, T.-L. Liu, Y. Jia, J. Li, Recent advances in bio-integrated electrochemical sensors for neuroengineering, Fundam. Res. (2023).
[25] F. Mattioli, V. Maglianella, S. D'Antonio, E. Trimarco, D. Caligiore, Non-invasive brain stimulation for patients and healthy subjects: Current challenges and future perspectives, J. Neurol. Sci. (2023) 122825.
[26] M.J. Rivera, M.A. Teruel, A. Mate, J. Trujillo, Diagnosis and prognosis of mental disorders by means of EEG and deep learning: a systematic mapping study, Artif. Intell. Rev. (2022) 1–43.
[27] S. Yang, J.M.S. Bornot, K. Wong-Lin, G. Prasad, M/EEG-based bio-markers to predict the MCI and Alzheimer's disease: a review from the ML perspective, IEEE Trans. Biomed. Eng. 66 (2019) 2924–2935.
[28] K.D. Tzimourta, V. Christou, A.T. Tzallas, N. Giannakeas, L.G. Astrakas, P. Angelidis, D. Tsalikakis, M.G. Tsipouras, Machine learning algorithms and statistical approaches for Alzheimer's disease analysis based on resting-state EEG recordings: A systematic review, Int. J. Neural Syst. 31 (2021) 2130002.
[29] N. Mammone, C. Ieracitano, H. Adeli, A. Bramanti, F.C. Morabito, Permutation Jaccard distance-based hierarchical clustering to estimate EEG network density modifications in MCI subjects, IEEE Trans. Neural Netw. Learn. Syst. 29 (2018) 5122–5135.

[30] D. De Venuto, V.F. Annese, G. Mezzina, Remote neuro-cognitive impairment sensing based on P300 spatio-temporal monitoring, IEEE Sens. J. 16 (2016) 8348–8356.
[31] X. Li, H. Zhou, R. Su, J. Kang, Y. Sun, Y. Yuan, Y. Han, X. Chen, P. Xie, Y. Wang, A mild cognitive impairment diagnostic model based on IAAFT and BiLSTM, Biomed. Signal. Process. Control 80 (2023) 104349.
[32] A.M. Alvi, S. Siuly, H. Wang, K. Wang, F. Whittaker, A deep learning based framework for diagnosis of mild cognitive impairment, Knowl. Based. Syst. 248 (2022) 108815.
[33] M. Aljalal, S.A. Aldosari, M. Molinas, F.A. Alturki, Selecting EEG channels and features using multi-objective optimization for accurate MCI detection: validation using leave-one-subject-out strategy, Sci. Rep. 14 (2024) 12483.
[34] F. Duan, Z. Huang, Z. Sun, Y. Zhang, Q. Zhao, A. Cichocki, Z. Yang, J. Sole-Casals, Topological network analysis of early Alzheimer's disease based on resting-state EEG, IEEE Trans. Neural Syst. Rehabil. Eng. 28 (2020) 2164–2172.
[35] H. Wu, L. Kong, Y. Zeng, H. Bao, Resting-state brain connectivity via multivariate EMD in mild cognitive impairment, IEEE Trans. Cogn. Dev. Syst. 14 (2021) 552–564.
[36] J. Jiang, Z. Yan, C. Sheng, M. Wang, Q. Guan, Z. Yu, Y. Han, J. Jiang, A novel detection tool for mild cognitive impairment patients based on eye movement and electroencephalogram, J. Alzheimer's Dis. 72 (2019) 389–399.
[37] Z.-Z. Cai, R. Lin, X.-X. Wang, Y.-J. Yan, H. Li, Effects of mindfulness in patients with mild cognitive impairment with insomnia: A double-blind randomized controlled trial, Geriatr. Nurs. (Minneap) 47 (2022) 239–246.
[38] D. Stogl, O. Armbruster, M. Mende, B. Hein, X. Wang, P. Meyer, Robot-based training for people with mild cognitive impairment, IEEE Robot. Autom. Lett. 4 (2019) 1916–1923.
[39] H. Zeng, X. Fang, Y. Zhao, J. Wu, M. Li, H. Zheng, F. Xu, D. Pan, G. Dai, EMCI: a novel EEG-based mental workload assessment index of mild cognitive impairment, IEEE Trans. Biomed. Circuits. Syst. 16 (2022) 902–914.
[40] X. Bi, H. Wang, Early Alzheimer's disease diagnosis based on EEG spectral images using deep learning, Neural Netw. 114 (2019) 119–135.
[41] C. Porcaro, F. Vecchio, F. Miraglia, G. Zito, P.M. Rossini, Dynamics of the "cognitive" brain wave P3b at rest for Alzheimer dementia prediction in mild cognitive impairment, Int. J. Neural Syst. 32 (2022) 2250022.
[42] E.Q. Wu, X.-Y. Peng, S.-D. Chen, X.-Y. Zhao, Z.-R. Tang, Detecting Alzheimer's dementia degree, IEEE Trans. Cogn. Dev. Syst. 14 (2020) 116–125.
[43] H. Yu, X. Lei, Z. Song, C. Liu, J. Wang, Supervised network-based fuzzy learning of EEG signals for Alzheimer's disease identification, IEEE Trans. Fuzzy Syst. 28 (2019) 60–71.
[44] Z. Song, B. Deng, J. Wang, R. Wang, Biomarkers for Alzheimer's disease defined by a novel brain functional network measure, IEEE Trans. Biomed. Eng. 66 (2018) 41–49.
[45] B.A. Kent, A.A. Casciola, S.K. Carlucci, M. Chen, S. Stager, M.S. Mirian, P. Slack, J. Valerio, M.J. McKeown, H.H. Feldman, Home EEG sleep assessment shows reduced slow-wave sleep in mild–moderate Alzheimer's disease, Alzheimer's Dementia: Transl. Res. Clin. Interv. 8 (2022) e12347.
[46] S. Afshari, M. Jalili, Directed functional networks in Alzheimer's disease: disruption of global and local connectivity measures, IEEE J. Biomed. Health Inform. 21 (2016) 949–955.
[47] D. Klepl, F. He, M. Wu, D.J. Blackburn, P. Sarrigiannis, EEG-based graph neural network classification of Alzheimer's disease: An empirical evaluation of functional connectivity methods, IEEE Trans. Neural Syst. Rehabil. Eng. 30 (2022) 2651–2660.
[48] M. Lopes, R. Cassani, T.H. Falk, Using CNN Saliency Maps and EEG Modulation Spectra for Improved and More Interpretable Machine Learning-Based Alzheimer's Disease Diagnosis, Comput. Intell. Neurosci. 2023 (2023) 3198066.
[49] X. Shan, J. Cao, S. Huo, L. Chen, P.G. Sarrigiannis, Y. Zhao, Spatial–temporal graph convolutional network for Alzheimer classification based on brain functional connectivity imaging of electroencephalogram, Hum. Brain Mapp. 43 (2022) 5194–5209.
[50] M. Imani, Alzheimer's diseases diagnosis using fusion of high informative BiLSTM and CNN features of EEG signal, Biomed. Signal. Process. Control 86 (2023) 105298.
[51] C. Roncero-Parra, A. Parreño-Torres, R. Sánchez-Reolid, J. Mateo-Sotos, A. L. Borja, Inter-hospital moderate and advanced Alzheimer's disease detection through convolutional neural networks, Heliyon (2024) 10.
[52] C.J. Huggins, J. Escudero, M.A. Parra, B. Scally, R. Anghinah, A.V.L. De Araújo, L. F. Basile, D. Abasolo, Deep learning of resting-state electroencephalogram signals for three-class classification of Alzheimer's disease, mild cognitive impairment and healthy ageing, J. Neural Eng. 18 (2021) 046087.
[53] I.A. Fouad, F.E.-Z.M. Labib, Identification of Alzheimer's disease from central lobe EEG signals utilizing machine learning and residual neural network, Biomed. Signal. Process. Control 86 (2023) 105266.
[54] D. kumar Ravikanti, S. Saravanan, EEGAlzheimer'sNet: Development of transformer-based attention long short term memory network for detecting Alzheimer disease using EEG signal, Biomed. Signal. Process. Control 86 (2023) 105318.
[55] S.K. Khare, U.R. Acharya, Adazd-Net: Automated adaptive and explainable Alzheimer's disease detection system using EEG signals, Knowl. Based. Syst. 278 (2023) 110858.
[56] Z. Song, B. Deng, J. Wang, G. Yi, An EEG-based systematic explainable detection framework for probing and localizing abnormal patterns in Alzheimer's disease, J. Neural Eng. 19 (2022) 036007.
[57] J. Sekhar, C. Rajyalakshmi, S. Nagaraj, S. Sankar, R. Saturi, A. Harshavardhan, Deep generative adversarial networks with marine predators algorithm for classification of Alzheimer's disease using electroencephalogram, J. King Saud Univ. Comput. Inf. Sci. 35 (2023) 101848.
[58] A. Miltiadous, E. Gionanidis, K.D. Tzimourta, N. Giannakeas, A.T. Tzallas, DICE-net: a novel convolution-transformer architecture for Alzheimer detection in EEG signals, IEEE Access (2023).
[59] L. Wu, Q. Zhao, J. Liu, H. Yu, Efficient identification of Alzheimer's brain dynamics with Spatial-Temporal Autoencoder: A deep learning approach for diagnosing brain disorders, Biomed. Signal. Process. Control 86 (2023) 104917.
[60] N. Houmani, F. Vialatte, E. Gallego-Jutglà, G. Dreyfus, V.-H. Nguyen-Michel, J. Mariani, K. Kinugawa, Diagnosis of Alzheimer's disease with Electroencephalography in a differential framework, PLoS One 13 (2018) e0193607.
[61] Y. Suzuki, M. Suzuki, K. Shigenobu, K. Shinosaki, Y. Aoki, H. Kikuchi, T. Baba, M. Hashimoto, T. Araki, K. Johnsen, A prospective multicenter validation study of a machine learning algorithm classifier on quantitative electroencephalogram for differentiating between dementia with Lewy bodies and Alzheimer's dementia, PLoS One 17 (2022) e0265484.
[62] Z. Zavecz, V.D. Shah, O.G. Murillo, R. Vallat, B.A. Mander, J.R. Winer, W. J. Jagust, M.P. Walker, NREM sleep as a novel protective cognitive reserve factor in the face of Alzheimer's disease pathology, BMC Med. 21 (2023) 156.
[63] P.A. Cicalese, R. Li, M.B. Ahmadi, C. Wang, J.T. Francis, S. Selvaraj, P.E. Schulz, Y. Zhang, An EEG-fNIRS hybridization technique in the four-class classification of alzheimer's disease, J. Neurosci. Methods 336 (2020) 108618.
[64] J. Cruzat, R. Herzog, P. Prado, Y. Sanz-Perl, R. Gonzalez-Gomez, S. Moguilner, M. L. Kringelbach, G. Deco, E. Tagliazucchi, A. Ibañez, Temporal irreversibility of large-scale brain dynamics in Alzheimer's disease, J. Neurosci. 43 (2023) 1643–1656.
[65] L.C. Nunes, P.R. Pinheiro, M.C.D. Pinheiro, M. Simão Filho, R.E.C. Nunes, P.G.C. D Pinheiro, Automatic detection and diagnosis of neurologic diseases, IEEE Access 7 (2019) 29924–29941.
[66] M. Nour, U. Senturk, K. Polat, A novel hybrid model in the diagnosis and classification of Alzheimer's disease using EEG signals: Deep ensemble learning (DEL) approach, Biomed. Signal. Process. Control 89 (2024) 105751.
[67] M. Hata, Y. Miyazaki, K. Mori, K. Yoshiyama, S. Akamine, H. Kanemoto, S. Gotoh, H. Omori, A. Hirashima, Y. Satake, Utilizing portable electroencephalography to screen for pathology of Alzheimer's disease: a methodological advancement in diagnosis of neurodegenerative diseases, Front. Psychiatry 15 (2024) 1392158.
[68] M. Buscema, F. Vernieri, G. Massini, F. Scrascia, M. Breda, P.M. Rossini, E. Grossi, An improved I-FAST system for the diagnosis of Alzheimer's disease from unprocessed electroencephalograms by using robust invariant features, Artif. Intell. Med. 64 (2015) 59–74.
[69] P. Durongbhan, Y. Zhao, L. Chen, P. Zis, M. De Marco, Z.C. Unwin, A. Venneri, X. He, S. Li, Y. Zhao, A dementia classification framework using frequency and time-frequency features based on EEG signals, IEEE Trans. Neural Syst. Rehabil. Eng. 27 (2019) 826–835.
[70] S.-K. Kim, H. Kim, S.H. Kim, J.B. Kim, L. Kim, Electroencephalography-based classification of Alzheimer's disease spectrum during computer-based cognitive testing, Sci. Rep. 14 (2024) 5252.
[71] S. Kongwudhikunakorn, S. Kiatthaveephong, K. Thanontip, P. Leelaarporn, M. Piriyajitakonkij, T. Charoenpattarawut, P. Autthasan, R. Chaisaen, P. Dujada, T. Sudhawiyangkul, A pilot study on visually stimulated cognitive tasks for EEG-based dementia recognition, IEEe Trans. Instrum. Meas. 70 (2021) 1–10.
[72] N. Sharma, M.H. Kolekar, K. Jha, Iterative filtering decomposition based early dementia diagnosis using EEG with cognitive tests, IEEE Trans. Neural Syst. Rehabil. Eng. 28 (2020) 1890–1898.
[73] H. Azami, S.E. Arnold, S. Sanei, Z. Chang, G. Sapiro, J. Escudero, A.S. Gupta, Multiscale fluctuation-based dispersion entropy and its applications to neurological diseases, IEEE Access 7 (2019) 68718–68733.
[74] Y. Chen, C. Hu, B. Hu, L. Hu, H. Yu, C. Miao, Inferring cognitive wellness from motor patterns, IEEE Trans. Knowl. Data Eng. 30 (2018) 2340–2353.
[75] M.-j. Kim, Y.C. Youn, J Paik, Deep learning-based EEG analysis to classify normal, mild cognitive impairment, and dementia: Algorithms and dataset, Neuroimage 272 (2023) 120054.
[76] Y. Pan, M. Liu, Y. Xia, D. Shen, Disease-image-specific learning for diagnosis-oriented neuroimage synthesis with incomplete multi-modality data, IEEE Trans. Pattern. Anal. Mach. Intell. 44 (2021) 6839–6853.
[77] Y. Watanabe, Y. Miyazaki, M. Hata, R. Fukuma, Y. Aoki, H. Kazui, T. Araki, D. Taomoto, Y. Satake, T. Suehiro, A deep learning model for the detection of various dementia and MCI pathologies based on resting-state electroencephalography data: A retrospective multicentre study, Neural Netw. 171 (2024) 242–250.
[78] E. Sibilano, A. Brunetti, D. Buongiorno, M. Lassi, A. Grippo, V. Bessi, S. Micera, A. Mazzoni, V. Bevilacqua, An attention-based deep learning approach for the classification of subjective cognitive decline and mild cognitive impairment using resting-state EEG, J. Neural Eng. 20 (2023) 016048.
[79] H. Jia, Z. Huang, C.F. Caiafa, F. Duan, Y. Zhang, Z. Sun, J. Solé-Casals, Assessing the Potential of Data Augmentation in EEG Functional Connectivity for Early Detection of Alzheimer's Disease, Cognit. Comput. 16 (2024) 229–242.
[80] S. Fouladi, A.A. Safaei, N. Mammone, F. Ghaderi, M.J. Ebadi, Efficient deep neural networks for classification of Alzheimer's disease and mild cognitive impairment from scalp EEG recordings, Cognit. Comput. 14 (2022) 1247–1268.
[81] C. Ieracitano, N. Mammone, A. Bramanti, A. Hussain, F.C. Morabito, A convolutional neural network approach for classification of dementia stages based on 2D-spectral representation of EEG recordings, Neurocomputing 323 (2019) 96–107.

[82] M.N.A. Tawhid, S. Siuly, E. Kabir, Y. Li, Exploring frequency band-based biomarkers of EEG signals for mild cognitive impairment detection, IEEE Trans. Neural Syst. Rehabil. Eng. (2023).
[83] H. Zhou, L. Yin, R. Su, Y. Zhang, Y. Yuan, P. Xie, X. Li, STCGRU: A hybrid model based on CNN and BiGRU for mild cognitive impairment diagnosis, Comput. Methods Programs Biomed. 248 (2024) 108123.
[84] Z. You, R. Zeng, X. Lan, H. Ren, Z. You, X. Shi, S. Zhao, Y. Guo, X. Jiang, X. Hu, Alzheimer's disease classification with a cascade neural network, Front. Public Health 8 (2020) 584387.
[85] A.T. Adebisi, H.-W. Lee, K.C. Veluvolu, EEG-based brain functional network analysis for differential identification of dementia-related disorders and their onset, IEEE Trans. Neural Syst. Rehabil. Eng. (2024).
[86] P.M. Rodrigues, B.C. Bispo, C. Garrett, D. Alves, J.P. Teixeira, D. Freitas, Lacsogram: A new EEG tool to diagnose Alzheimer's disease, IEEE J. Biomed. Health Inform. 25 (2021) 3384–3395.
[87] Y. Shi, Q. Ma, C. Feng, M. Wang, H. Wang, B. Li, J. Fang, S. Ma, X. Guo, T. Li, Microstate feature fusion for distinguishing AD from MCI, Health Inf. Sci. Syst. 10 (2022) 16.
[88] B. Wallace, F. Knoefel, R. Goubran, R.A.L. Zunini, Z. Ren, A. Maccosham, EEG/ERP: Within episodic assessment framework for cognition, IEEE Trans. Instrum. Meas. 66 (2017) 2525–2534.
[89] L. Spyrou, M. Parra, J. Escudero, Complex tensor factorization with PARAFAC2 for the estimation of brain connectivity from the EEG, IEEE Trans. Neural Syst. Rehabil. Eng. 27 (2018) 1–12.
[90] S. Chen, C. Zhang, H. Yang, L. Peng, H. Xie, Z. Lv, Z.-G. Hou, A Multi-Modal Classification Method for Early Diagnosis of Mild Cognitive Impairment and Alzheimer's Disease Using Three Paradigms With Various Task Difficulties, IEEE Trans. Neural Syst. Rehabil. Eng. 32 (2024) 1456–1465.
[91] T. Dhar, N. Dey, S. Borra, R.S. Sherratt, Challenges of deep learning in medical image analysis—improving explainability and trust, IEEE Trans. Technol. Soc. 4 (2023) 68–75.
[92] Q. Teng, Z. Liu, Y. Song, K. Han, Y. Lu, A survey on the interpretability of deep learning in medical diagnosis, Multimed. Syst. 28 (2022) 2335–2355.
[93] A. Atmakuru, S. Chakraborty, O. Faust, M. Salvi, P.D. Barua, F. Molinari, U. Acharya, N. Homaira, Deep learning in radiology for lung cancer diagnostics: A systematic review of classification, segmentation, and predictive modeling techniques, Expert. Syst. Appl. (2024) 124665.
[94] D. Klepl, F. He, M. Wu, D.J. Blackburn, P. Sarrigiannis, Adaptive gated graph convolutional network for explainable diagnosis of Alzheimer's disease using EEG data, IEEE Tran. Neural Syst. Rehabil. Eng. (2023).
[95] M. Kashefpoor, H. Rabbani, M. Barekatain, Automatic diagnosis of mild cognitive impairment using electroencephalogram spectral features, J. Med. Signals Sens. 6 (2016) 25–32.
[96] A.M. Pineda, F.M. Ramos, L.E. Betting, A.S. Campanharo, Quantile graphs for EEG-based diagnosis of Alzheimer's disease, PLoS One 15 (2020) e0231169.
[97] M. Cejnek, O. Vysata, M. Valis, I. Bukovsky, Novelty detection-based approach for Alzheimer's disease and mild cognitive impairment diagnosis from EEG, Med. Biol. Eng. Comput. 59 (2021) 2287–2296.
[98] E. Gallego-Jutglà, J. Solé-Casals, F.-B. Vialatte, M. Elgendi, A. Cichocki, J. Dauwels, A hybrid feature selection approach for the early diagnosis of Alzheimer's disease, J. Neural Eng. 12 (2015) 016018.
[99] T. Musha, T. Asada, F. Yamashita, T. Kinoshita, Z. Chen, H. Matsuda, M. Uno, W. R. Shankle, A new EEG method for estimating cortical neuronal impairment that is sensitive to early stage Alzheimer's disease, Clin. Neurophysiol. 113 (2002) 1052–1058.
[100] M. Pietto, M.A. Parra, N. Trujillo, F. Flores, A.M. García, J. Bustin, P. Richly, F. Manes, F. Lopera, A. Ibáñez, Behavioral and electrophysiological correlates of memory binding deficits in patients at different risk levels for Alzheimer's disease, J. Alzheimer's Dis. 53 (2016) 1325–1340.
[101] A. Miltiadous, K.D. Tzimourta, T. Afrantou, P. Ioannidis, N. Grigoriadis, D. G. Tsalikakis, P. Angelidis, M.G. Tsipouras, E. Glavas, N. Giannakeas, A dataset of EEG recordings from: Alzheimer's disease, Frontotemporal dementia and Healthy subjects, OpenNeuro 1 (2023) 88.
[102] Y. Si, R. He, L. Jiang, D. Yao, H. Zhang, P. Xu, X. Ma, L. Yu, F. Li, Differentiating between Alzheimer's Disease and Frontotemporal Dementia Based on the Resting-State Multilayer EEG Network, IEEE Trans. Neural Syst. Rehabil. Eng. (2023).
[103] S. Siuly, Ö.F. Alçin, H. Wang, Y. Li, P. Wen, Exploring rhythms and channels-based EEG biomarkers for early detection of alzheimer's disease, IEEE Trans. Emerg. Top. Comput. Intell. (2024).
[104] D. Wen, Z. Bian, Q. Li, L. Wang, C. Lu, X. Li, Resting-state EEG coupling analysis of amnestic mild cognitive impairment with type 2 diabetes mellitus by using permutation conditional mutual information, Clin. Neurophysiol. 127 (2016) 335–348.
[105] D. Wen, P. Li, Y. Zhou, Y. Sun, J. Xu, Y. Liu, X. Li, J. Li, Z. Bian, L. Wang, Feature classification method of resting-state EEG signals from amnestic mild cognitive impairment with type 2 diabetes mellitus based on multi-view convolutional neural network, IEEE Trans. Neural Syst. Rehabil. Eng. 28 (2020) 1702–1709.
[106] H.W. Loh, C.P. Ooi, S. Seoni, P.D. Barua, F. Molinari, U.R. Acharya, Application of explainable artificial intelligence for healthcare: A systematic review of the last decade (2011–2022), Comput. Methods Programs Biomed. 226 (2022) 107161.
[107] M. Abdar, F. Pourpanah, S. Hussain, D. Rezazadegan, L. Liu, M. Ghavamzadeh, P. Fieguth, X. Cao, A. Khosravi, U.R. Acharya, A review of uncertainty quantification in deep learning: Techniques, applications and challenges, Inf. Fusion 76 (2021) 243–297.
[108] S. Flügge, S. Zimmer, U. Petersohn, arXiv preprint arXiv: 1909.08549, 2019.
[109] V. Jahmunah, E.Y.K. Ng, R.-S. Tan, S.L. Oh, U.R. Acharya, Uncertainty quantification in DenseNet model using myocardial infarction ECG signals, Comput. Methods Programs Biomed. 229 (2023) 107308.
[110] Y. Wang, X. Liu, C. Yu, Assisted diagnosis of alzheimer's disease based on deep learning and multimodal feature fusion, Complexity 2021 (2021) 6626728.
[111] A.M. Hafiz, S.A. Parah, R.U. Bhat. Attention mechanisms and deep learning for machine vision: A survey of the state of the art, 2021. http://www.doi.org.10.21203/rs.3.rs-510910/v1.
[112] S. Georgousis, M.P. Kenning, X. Xie, Graph deep learning: State of the art and challenges, IEEe Access 9 (2021) 22106–22140.
[113] P. Liu, X. Qiu, X. Huang, arXiv preprint arXiv: 1609.07222, 2016.
[114] P.M. Mammen, arXiv preprint arXiv: 2101.05428, 2021.
[115] B. Yuan, S. Ge, W. Xing, arXiv preprint arXiv: 2005.05083, 2020.
[116] H. Zhang, W. Liang, C. Li, Q. Xiong, H. Shi, L. Hu, G. Li, DCML: Deep contrastive mutual learning for COVID-19 recognition, Biomed. Signal. Process. Control 77 (2022) 103770.
[117] S. Seoni, A. Shahini, K.M. Meiburger, F. Marzola, G. Rotunno, U.R. Acharya, F. Molinari, M. Salvi, All you need is data preparation: A systematic review of image harmonization techniques in Multi-center/device studies for medical support systems, Comput. Methods Programs Biomed. (2024) 108200.