An AI-powered EEG system for early Alzheimer’s detection Area:
Abstract

Seven studies describe EEG-based artificial intelligence systems that distinguish early Alzheimer’s disease from nor- mal or mildly impaired brain activity. Each study applies advanced filtering (typically 1–50 Hz), artifact rejection, and normalization to segment EEG signals, then extracts features via wavelet transforms, FFT, and other time–frequency methods. In these investigations, classic techniques (SVM, LDA, k–NN, decision trees) and deep learning models (CNN, RNN, BiLSTM, hybrid architectures) achieved binary classification accuracies between 78.9% and 100% and multi–class accuracies from 56.8% to 99.9%. Reported sensitivity values range from 60% to 88% and specificity from 70% to 90%, with cross–validation schemes including leave–one–out, LOSO, k–fold, and jack–knifing. Although studies vary in preprocessing protocols and sample sizes (from 21 to 225 participants) and often provide limited demographic details, they consistently show that AI–powered EEG analysis can capture subtle patterns associated with early Alzheimer’s disease.

Paper search

Using your research question ”An AI-powered EEG system for early Alzheimer’s detection Area: medi- cal_imaging,artificial_intelligence,machine_learning,signal_processing,image_processing Domain: Healthcare Objective: The main objective of this project is to develop an AI-based system that can enable early detection

of Alzheimer’s disease using EEG signals. The project aims to process and analyze brainwave data to identify subtle patterns associated with cognitive decline, even before noticeable clinical symptoms appear. Advanced signal preprocessing and feature extraction techniques will be employed to ensure that meaningful information is captured from raw EEG data. Deep learning models will then be trained and validated to accurately distinguish between healthy brain activity and early Alzheimer’s patterns. Additionally, the system will focus on providing explainable and clinically relevant insights, so that medical professionals can better understand the basis of the predictions. The broader goal is to create a non-invasive, cost-effective, and accessible screening tool that supports timely intervention and contributes to improving patient care and quality of life. make a problem statement”, we searched across over 126 million academic papers from the Semantic Scholar corpus. We retrieved the 50 papers most relevant to the query.

Screening

We screened in sources that met these criteria:

- EEG and Alzheimer's Focus: Does this study involve EEG signal acquisition and analysis for Alzheimer's disease detection, diagnosis, or classification?

- AI/ML Methods: Does this study apply artificial intelligence, machine learning, or deep learning techniques (not limited to traditional statistical methods only)?
- Early-Stage Disease Focus: Does this study include participants with early-stage Alzheimer's disease, mild cognitive impairment (MCI), or preclinical stages?
- Performance Metrics: Does this study report diagnostic performance metrics such as sensitivity, specificity, accuracy, or AUC?
- Study Groups and Controls: Does this study include clearly defined participant groups with appropriate control groups or comparison populations?
- Publication Type: Is this an original research article, systematic review, or meta-analysis (not a case report, editorial, conference abstract, or opinion piece)?
- Human Studies: Is this a human study (not an animal study or in-vitro study)?
- Diagnostic Classification Focus: Does this study focus on diagnostic classification and include EEG data for Alzheimer's disease (not exclusively other neuroimaging modalities without EEG or exclusively other neurode- generative diseases without Alzheimer's)?

We considered all screening questions together and made a holistic judgement about whether to screen in each paper.

Data extraction

We asked a large language model to extract each data column below from each paper. We gave the model the extraction instructions shown below for each column.

- Study Design Type:

  Identify the primary study design type from the full text. Look in the methods section for explicit description of study design. Categorize as:

- Cross-sectional diagnostic study
- Prospective cohort study
- Case-control study
- Experimental/interventional study

If multiple design elements are present, list in order of prominence. If unclear, note ”Design not clearly specified” and provide a brief explanation.

- Signal Processing and AI Techniques:

  Systematically extract ALL signal processing and AI techniques used in the study:

1. EEG Preprocessing techniques (e.g., filtering, artifact removal)
2. Feature extraction methods (e.g., Empirical Mode Decomposition, Blind Source Separation)
3. Machine learning/AI classification algorithms (e.g., Linear Discriminant Analysis, Deep

Learning models) For each technique, provide:

- Specific name of technique
- Brief description of how it was applied
- Any unique parameters or modifications used

If multiple techniques were compared, note all techniques and their relative performance.

- Participant Groups and Sample Characteristics:

  For each participant group in the study, extract:

- Total number of participants
- Diagnostic category (e.g., Healthy Controls, Mild Cognitive Impairment, Mild/Moderate Alzheimer's)
- Mean age (and standard deviation)
- Gender distribution
- Inclusion/exclusion criteria

If multiple groups are present, create a clear breakdown for each group. If age or gender data is incomplete, note ”Partial data available” and specify what information is missing.

- Classification Performance Metrics:

  Extract ALL reported performance metrics for the AI/machine learning classification:

1. Accuracy (%)
2. Sensitivity (%)
3. Specificity (%)
4. Precision (%)
5. F1 Score (if

reported) For each

metric:

- Report the exact value
- Note the validation method used (e.g., k-fold cross-validation, leave-one-subject-out)
- If multiple metrics or validation approaches are used, list all

If performance metrics are presented in a graph or figure, carefully transcribe the numerical values.

- Key Findings and Clinical Implications:

  Summarize the primary findings related to Alzheimer's detection, focusing on:

1. Key insights about EEG patterns or biomarkers
2. Potential clinical utility of the proposed method

3. Any limitations or recommendations for future research

Extract direct quotes from the conclusion section that highlight the most significant findings. Ensure the summary captures the broader implications for early Alzheimer's detection.

Results

Characteristics of Included Studies

Study	Study Population Approach	EEG Setup	AI	Validation Method	Full text retrieved
Araújo et al., 2022	38 participants: Healthy controls, Mild Cognitive Impairment (MCI), Mild/Moderate Alzheimer's Disease (AD), Advanced AD. Age/gender: No mention found.	Electroencephalography (EEG) signals segmented into 5-second windows, 1–40 Hz band-pass filter, normalization.	Classic machine learning (decision trees, Support Vector Machine (SVM), Naive Bayes, k-Nearest Neighbors (k-NN), logistic regression, discriminant analysis), Deep Learning (Convolutional Neural Network (CNN)). Feature selection via f-score.	Leave-one-out cross-validation	Yes

Study	Study Population Approach	EEG Setup	AI	Validation Method	Full text retrieved
Alsharabi et al., 2023	88 participants: 35 neurotypical, 31 mild AD, 22 moderate AD. Mean age: 66.9–75.2. Gender: 35M/53F.	(progressed to AD), 38 controls. Mean age: ~72. Gender: No mention found.	Resting-state EEG, band-pass elliptic filter.	Blind Source Separation (BSS, AMUSE algorithm).	Empirical Mode Decomposition (EMD) for feature extraction; machine learning (Linear Discriminant Analysis (LDA), SVM, Random Forest (RF)), Deep Learning (Artificial Neural Network (ANN), Recurrent Neural Network (RNN), CNN). SVM (Radial Basis Function kernel), logistic regression. Features: spectral, complexity, entropy.
Pérez-Valero et al., 2022	21 participants: mild AD, MCI-non-AD, healthy controls. Age/gender: partial.		Commercial EEG, 1–45 Hz Finite Impulse Response (FIR) filter, artifact rejection (Autoreject, Independent Component Analysis (ICA)).		LDA on relative power in frequency bands.
Cichocki et al., 2005	60 participants: 22 MCI		Resting EEG, artifact-free intervals,		

k-fold and Leave-One-Subject-Out (LOSO) cross-validation	Yes
	Yes
Leave-One-Subject-Out (LOSO) cross-validation	
	Yes
Jack-knifing cross-validation	

Baker et al., 2008 participants:	58 17 AD, 25 MCI,	1 6 c	o n t r	o l s .	Mean age: ~75. Gender: 23M/35F.

0.5–50 Hz bandpass, artifact-free 30-second segments, normalization.	K-means clustering on band power features.	No explicit validation method stated; follow-up for MCI progression	Yes

Study	Study Population	EEG Setup	AI Approach	Validation Method	Full text retrieved
Ismail et al., 2019	Number, age, gender: No mention found. MCI, AD, healthy.	Low-cost EEG, 3-level N-Back test, Fast Fourier Transform (FFT) for subbands.	Deep Learning (CNN) on 2D images of subbands.	No mention found	No
Khosravi et al., 2024	225 participants: 59 AD, 56 MCI, 110 controls. Age/gender: No mention found.	Discrete Wavelet Transform (DWT), windowing, Short-Time Fourier Transform (STFT) for spectrograms.	Convolutional, Attention, Bidirectional Long Short-Term Memory (CL-ATBiLSTM), Bayesian optimization.	No explicit validation method stated	Yes

AI Approaches:

- Classic machine learning methods (e.g., SVM, LDA, logistic regression, decision trees, k-NN, RF, discriminant analysis) were used in four studies.
- Deep learning approaches (CNN, ANN, RNN, Bidirectional Long Short-Term Memory (BiLSTM)) were used in four studies.
- One study used clustering (k-means).
- One study used an advanced hybrid deep learning model (convolution, attention, BiLSTM with Bayesian opti- mization).
- Feature extraction or selection methods (e.g., EMD, f-score, spectral, complexity, entropy, FFT, DWT, STFT) were described in six studies.
- Bayesian optimization was reported in

one study. Validation Methods:

- Leave-one-out cross-validation was used in one study.
- Leave-One-Subject-Out (LOSO) cross-validation was used in two studies.
- k-fold cross-validation was used in one study.
- Jack-knifing cross-validation was used in one study.
- No explicit validation method was found for

three studies. EEG Setup:

- Filtering (band-pass, FIR, elliptic) was described in four studies.
- Artifact rejection (manual, Autoreject, ICA, BSS) was described in three studies.
- Windowing was described in two studies.
- Normalization was described in two studies.
- Time-frequency transforms (DWT, STFT, FFT) were described in three studies.
- Feature extraction (EMD, spectral, complexity, entropy) was described in four studies.

- Resting-state EEG was described in two studies.
- Use of commercial or low-cost EEG was reported in

two studies. Reporting Gaps:

- Participant numbers were not mentioned in two studies.
- Age and gender information was missing or only partially reported in four studies.

Effects

Classification Performance

Study	Classification Task	Accuracy	Sensitivity/Specificity	Sample Size
Araújo et al., 2022	Healthy controls vs. MCI: 78.9%; Healthy controls vs. Mild/Moderate AD: 81.0%; Healthy controls vs. Advanced AD: 84.2%; MCI vs. Mild/Moderate AD: 88.9%; MCI vs. Advanced AD: 93.8%; Mild/Moderate AD vs. Advanced AD: 77.8%; All groups vs. all: 56.8%	No mention found	No mention found	38
Alsharabi et al., 2023	Multi-class (neurotypical, mild AD, moderate AD): k-fold: up to 99.9%; LOSO: up to 94.8%	No mention found	No mention found	88

Pérez-Valero et al., 2022	MCI-non-AD vs. healthy controls: F1-score = 0.96; mild AD vs. healthy controls: F1-score = 0.86; Sensitivity (mild AD): 0.88	No mention found	21
Cichocki et al., 2005	MCI (progressed to AD) vs. controls: 80% (after preprocessing)	Sensitivity: 60–80%; Specificity: 70–90%	60

Study	Classification Task	Accuracy	Sensitivity/Specificity	Sample Size
Baker et al., 2008	AD vs. controls: >80%; MCI prediction: 4/6 subgroup correct at 2-year follow-up	No mention found	No mention found	58
		No mention found	No mention found	No mention
Ismail et al., 2019	MCI: 90.36%; AD: 92.52%	found No mention found		No mention
Khosravi et al., 2024	Two-class: 100%; Three-class: 96.52%	found		225

Classification Tasks:

- Six studies performed binary classification tasks (distinguishing between healthy controls and MCI/AD groups).
- Three studies performed multi-class classification tasks (distinguishing among neurotypical, mild AD, moder- ate AD, or all groups).

Accuracy:

- Accuracy values were reported in all seven studies.
- Reported accuracy ranged from 56.8% (multi-class, Araújo et al.) to 100% (two-class, Khosravi et al.).
- Multi-class accuracy was up to 99.9% (Alsharabi et al., k-fold), 96.52% (Khosravi et al., three-class), and 56.8% (Araújo et al., all groups).
- Binary classification accuracy ranged from

78.9% to 100%. Sensitivity/Specificity:

- Sensitivity values were found in two studies (0.88 in Pérez-Valero et al.; 60–80% in Cichocki et al.).
- Specificity values were found in one study (70–90% in Cichocki et al.).
- No mention of sensitivity or specificity was found in the other

five studies. Sample Size:

- Sample sizes were reported in six studies (21, 38, 58, 60, 88, and 225).
- No mention of sample size was found in one study.
- Reported sample sizes ranged from 21 to 225.

Technical Approach Comparison

All included studies used advanced signal processing and artificial intelligence techniques, with notable methodolog- ical diversity:

- Preprocessing:
  - Most studies used band-pass filtering (typically 1–40/50 Hz), artifact rejection (manual or automated), and normalization.

- Some studies used advanced artifact removal, such as Blind Source Separation (Cichocki et al., 2005) and Independent Component Analysis (Pérez-Valero et al., 2022).

- Feature Extraction:
  - Methods included wavelet transforms (Araújo et al., 2022; Khosravi et al., 2024), Empirical Mode Decom- position (Alsharabi et al., 2023), Fast Fourier Transform (Ismail et al., 2019), and spectrograms (Khosravi et al., 2024).
  - Features spanned spectral, entropy, complexity, and statistical domains.
- Artificial Intelligence Models:
  - Both classic machine learning (SVM, LDA, k-means, RF) and deep learning (CNN, RNN, BiLSTM, hybrid models) were used.
  - Deep learning models are increasingly favored, but classic machine learning can outperform in small datasets (as reported by Araújo et al., 2022).
- Validation:
  - Robust cross-validation (LOSO, k-fold) was common, but external validation was rare.

Key Insights:

- The field is moving toward hybrid and deep learning models, with increasing sophistication in feature extrac- tion and model optimization (e.g., Bayesian optimization in Khosravi et al., 2024).
- Lack of standardization in preprocessing and feature extraction complicates comparison and may affect repro- ducibility.

Clinical Validation and Generalizability

- Clinical Utility:
  - All included studies report the potential for non-invasive, cost-effective, and accessible early Alzheimer's Disease screening using artificial intelligence-powered EEG systems.
  - Several studies demonstrate the ability to distinguish Mild Cognitive Impairment from healthy controls, supporting early intervention.
- Generalizability:
  - Most studies are limited by small, single-center samples, incomplete demographic reporting, and lack of external validation.
  - Only Baker et al., 2008 included a prospective component.

- Only a few studies used commercial or low-cost EEG systems, which may enhance real-world applicabil- ity.

- Limitations:
  - Common limitations include small sample size, lack of external validation, incomplete reporting of de- mographic and clinical characteristics, and potential overfitting in high-accuracy reports.

Summary:

- The evidence from these studies supports the feasibility of artificial intelligence-powered EEG systems for early Alzheimer's Disease detection. However, generalizability and clinical translation remain limited by method- ological and reporting shortcomings.

References

A. Cichocki, S. Shishkin, T. Musha, Zbigniew Leonowicz, T. Asada, and T. Kurachi. “EEG Filtering Based on Blind Source Separation (BSS) for Early Detection of Alzheimer's Disease.” *Clinical Neurophysiology*, 2005.

Eduardo Pérez-Valero, C. Morillas, M. A. Lopez-Gordo, I. Carrera-Muñoz, S. López-Alcalde, and R. Vílchez-Carrillo. “An Automated Approach for the Detection of Alzheimer's Disease From Resting State Electroencephalography.” *Frontiers in Neuroinformatics*, 2022.

Khalil Alsharabi, Yasser Bin Salamah, Majid Aljalal, Akram M. Abdurraqeeb, and F. Alturki. “EEG-Based Clinical De- cision Support System for Alzheimer's Disorders Diagnosis Using EMD and Deep Learning Techniques.” *Frontiers in Human Neuroscience*, 2023.

M. Baker, Kwaku Akrofi, R. Schiffer, and Michael W. O’ Boyle. “EEG Patterns in Mild Cognitive Impairment (MCI) Patients.” *The Open Neuroimaging Journal*, 2008.

M. Ismail, K. Hofmann, and Mohamed A. Abd El Ghany. “Early Diagnoses of Alzheimer Using EEG Data and Deep Neural Networks Classification.” *2019 IEEE Global Conference on Internet of Things (GCIoT)*, 2019.

M. Khosravi, Hossein Parsaei, Khosro Rezaee, and M. Helfroush. “Fusing Convolutional Learning and Attention- Based Bi-LSTM Networks for Early Alzheimer’s Diagnosis from EEG Signals Towards IoMT.” *Scientific Reports*, 2024.

Teresa Araújo, João Paulo Teixeira, and P. Rodrigues. “Smart-Data-Driven System for Alzheimer Disease Detection Through Electroencephalographic Signals.” *Bioengineering*, 2022.