MJSAT

Malaysian Journal of Science and Advanced Technology

journal homepage: https://mjsat.com.my/

Machine Learning Techniques for the Early Detection of Alzheimer's Disease: A Systematic Review

Mohamad Al Saeed*[1], and N. H. R. Azamin[1]

[1]*School of Engineering and Computing, MILA University, Nilai, Malaysia.*

KEYWORDS

Alzheimer's Disease
Early Detection
Machine Learning
Neuroimaging
Systematic Review

ARTICLE HISTORY

Received 15 January 2025
Received in revised form 20 June 2025
Accepted 25 June 2025
Available online 29 June 2025

ABSTRACT

Alzheimer's disease (AD) is a progressive neurodegenerative disorder that poses a significant global health challenge. Early and accurate detection is crucial for timely intervention and for the development of new therapies. Machine learning (ML) has emerged as a powerful tool for analyzing the complex, high-dimensional data associated with AD. This systematic review was conducted by searching major academic databases for peer-reviewed literature published between 2020 and 2025. We identified studies that applied ML models to neuroimaging data for the early detection of AD. Information on ML models, datasets, and performance metrics was extracted and synthesized to provide a comprehensive overview of the field. A range of ML models are employed, from traditional supervised learning algorithms like Support Vector Machines (SVM) to more advanced ensemble (e.g., Random Forest) and deep learning methods (e.g., CNNs). Studies consistently show that ensemble and deep learning models achieve high performance (>90% accuracy in many cases), particularly in multiclass classification. However, the field faces persistent challenges, including severe class imbalance in common datasets, issues of data quality and anomalies, and the "black box" nature of complex models, which limits their interpretability and clinical trust. ML models show immense promise for the early and accurate detection of AD. However, for these tools to be successfully translated into clinical practice, future research must focus on developing robust, generalizable, and interpretable models that can effectively address the challenges of data imbalance and pathological heterogeneity.

1. INTRODUCTION

Representing a profound and escalating public health crisis, Alzheimer's disease (AD) stands as the most common cause of dementia—a progressive neurodegenerative disorder that gradually erodes cognitive function and robs individuals of their independence[1]. An estimated 7.2 million Americans age 65 and older are living with Alzheimer's dementia in 2025, and this number is projected to reach 13.8 million by 2060 without significant medical breakthroughs [2]. The disease is defined by a gradual and devastating decline in cognitive function, including memory, language, and problem-solving skills, which ultimately robs individuals of their independence[3]. At its core, the pathology of AD is characterized by two hallmark features: the extracellular aggregation of amyloid-beta (Aβ) peptides into amyloid plaques and the intracellular accumulation of hyperphosphorylated tau protein into neurofibrillary tangles (NFTs)[4, 5].

A crucial aspect of AD is its long, insidious preclinical phase, where these neuropathological changes begin to accumulate silently in the brain, often decades before the first clinical symptoms of memory loss become apparent [6]. This asymptomatic period presents a critical, but challenging, window of opportunity for intervention. The modern conceptualization of AD has evolved to recognize this disease continuum, shifting the focus of the research community away from managing late-stage symptoms and toward the paramount goal of early and accurate detection [7]. Identifying the disease in its earliest stages is essential for enrolling patients in clinical trials, developing effective disease-modifying therapies, and implementing strategies to slow its progression [8].

Traditional diagnostic methods, however, are often insufficient for this task. Clinical assessments and cognitive tests can be subjective and may lack the sensitivity to detect the subtle changes of early-stage AD[9, 10]. More definitive diagnostic tools, such as Positron Emission Tomography (PET) scans and cerebrospinal fluid (CSF) analysis, while valuable, are limited by their high cost, invasiveness, or lack of widespread accessibility [11]. This diagnostic gap highlights

*Corresponding author:
E-mail address: Mohamad Al Saeed < m.alsaid007@gmail.com >.
https://doi.org/10.56532/mjsat.v5i2.570

the urgent need for more objective, scalable, and cost-effective methods for early detection.

The advent of machine learning (ML) offers a transformative approach to this challenge. As a computational paradigm, ML is uniquely suited to analyze the complex, high-dimensional data derived from neuroimaging and clinical assessments [12]. ML models have the potential to identify the subtle, intricate patterns that are indicative of early-stage AD but are often imperceptible to human observers, thereby providing a more accurate, accessible, and non-invasive pathway for early diagnosis [13].

The primary aim of this systematic review is to survey and synthesize the current landscape of machine learning techniques applied to the early detection of Alzheimer's disease. This review addresses the following research questions:

RQ1: What are the standard datasets and feature types used in ML-based AD research?

RQ2: What are the most common machine learning models applied to AD classification, and what is their reported performance?

RQ3: What are the primary methodological challenges and research gaps limiting the clinical utility of these models?

This paper is structured as follows. Section 2 outlines the methodology used for this systematic review. Section 3 presents the analysis and synthesis of the literature, structured around the key research questions. Finally, Section 4 offers a conclusion and outlines future research directions.

2. SCOPE AND METHODOLOGY

This systematic literature review follows established guidelines for identifying, evaluating, and synthesizing relevant scientific literature.

2.1 Search Strategy

The literature search was conducted across three major academic databases: PubMed, IEEE Xplore, and Scopus. The search strategy employed a combination of keywords and Boolean operators to identify relevant studies, as detailed in Table 1.

Table 1. Search Terms and Boolean Operators.

Category	Search Terms
Disease Focus	"Alzheimer's Disease" OR "Mild Cognitive Impairment"
Methodology	"Machine Learning" OR "Deep Learning" OR "Ensemble Learning" OR "Support Vector Machine" OR "Random Forest" OR "CNN"
Data Modality	"Neuroimaging" OR "MRI"
Dataset	"OASIS" OR "ADNI"
Combined Query	(Disease Focus) AND (Methodology) AND (Data Modality)

2.2 Inclusion and Exclusion Criteria

The selection of studies was guided by a formal set of inclusion and exclusion criteria, as outlined in Table 2, to ensure the quality and relevance of the literature included in this review.

Table 2. Inclusion and Exclusion Criteria.

Inclusion Criteria	Exclusion Criteria
Published between January 2020 and June 2025	Published outside the specified date range
Peer-reviewed journal or conference publication	Not peer-reviewed (e.g., preprints, theses)
Written in English	Written in a language other than English
Focus on ML/DL for AD or MCI detection	Not relevant to the application of ML to AD
Utilizes neuroimaging data (MRI, PET, etc.)	Does not use neuroimaging data
Provides detailed methodology and clear performance metrics	Insufficient methodological detail or performance metrics
Uses standard, accessible datasets (e.g., ADNI, OASIS)	Uses private, inaccessible datasets

2.3 Study Selection and Data Extraction

The study selection process followed a systematic approach illustrated in the PRISMA flowchart (Figure 1). Our comprehensive database search across PubMed, IEEE Xplore, Scopus, and Google Scholar targeted Alzheimer's disease terminology, cognitive impairment indicators, and machine learning techniques including deep learning, ensemble methods, SVM, random forest, and CNN approaches applied to neuroimaging datasets.

Following deduplication, a rigorous two-stage screening protocol was implemented. The first stage screened titles and abstracts against predefined exclusion criteria, removing studies lacking AD/MCI relevance, neuroimaging data, falling outside the 2020-2025 publication window, or non-peer-reviewed publications. This initial screening significantly reduced the candidate pool while ensuring focus on recent, high-quality advances.

The second stage employed full-text review with stringent quality assessment. Studies were excluded for insufficient methodological rigor, missing critical performance metrics, inadequate sample sizes, or poor study quality compromising synthesis reliability. This dual-stage approach ensured only methodologically sound studies with robust experimental designs reached final inclusion.

The final cohort of selected studies underwent systematic data extraction using a standardized form designed to capture essential methodological components: ML model architectures, dataset characteristics, input modalities, key performance indicators, classification tasks, preprocessing methodologies, and validation approaches. This extracted information was then categorized by study type, dataset utilized, and core techniques to enable comprehensive synthesis and identification of methodological trends, performance benchmarks, and emerging challenges in AI-driven early detection of Alzheimer's disease.

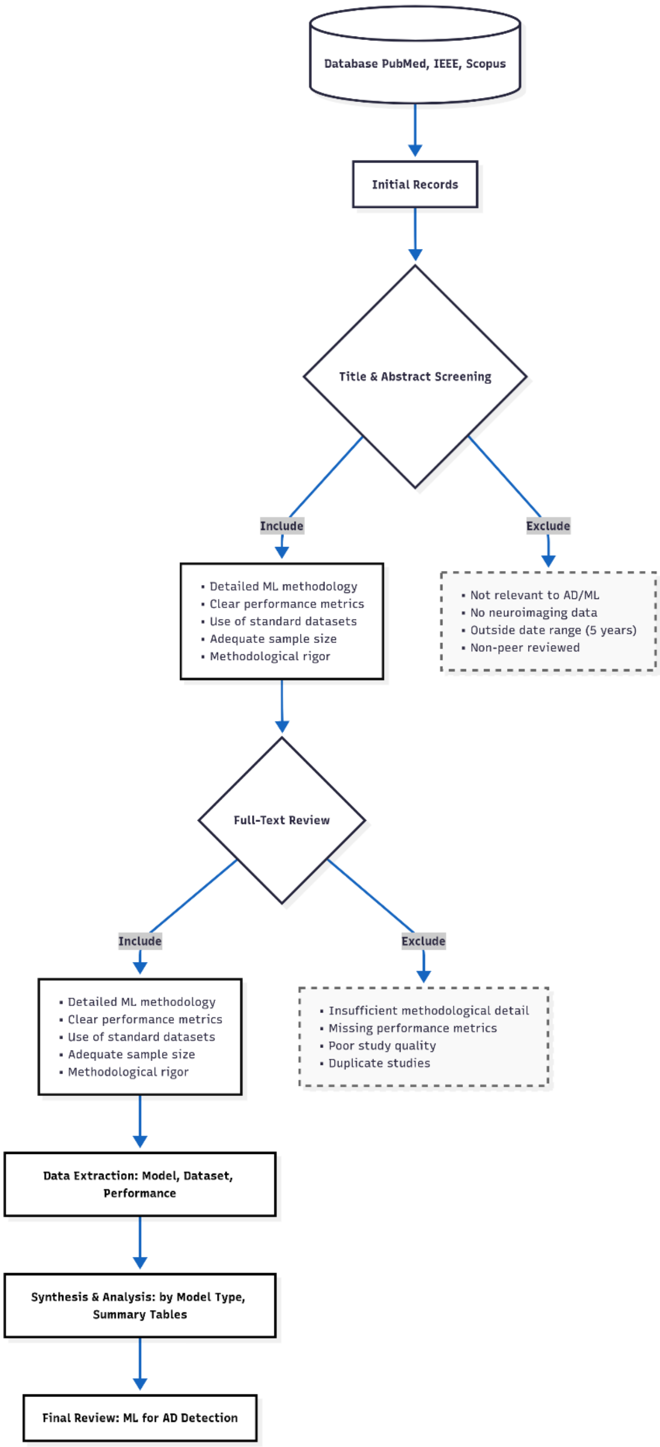

Fig. 1. Flowchart of the Study Selection Process.

3. Analysis and Synthesis of the Literature

This section synthesizes the findings from the selected literature, structured to directly address the research questions delineated in the introduction, providing a comprehensive overview of the current state of ML-based AD detection.

3.1 RQ1: What are the standard datasets and feature types used in ML-based AD research?

The performance of ML models heavily relies on large, well-curated datasets and the features extracted from them. Two of the most prominent public datasets in AD research are the OASIS (Open Access Series of Imaging Studies) [14] and the ADNI (Alzheimer's Disease Neuroimaging Initiative) [15]. These repositories provide a wealth of data, including structural MRI (sMRI), functional MRI (fMRI), PET, genetic, and clinical biomarker information.

Before features are used for model training, raw neuroimaging data is typically processed through automated pipelines to perform tasks such as segmentation, registration, and normalization. These pipelines produce quantitative, handcrafted features that are then used as inputs for ML models. A summary of the most common pipelines is provided in Table 3.

Table 3. Common Automated Pipelines for Neuroimaging Preprocessing.

Pipeline	Primary Function	Key Features Extracted	Typical Modality
FreeSurfer	Cortical surface reconstruction and parcellation	Cortical thickness, subcortical volumes, surface area	sMRI
SPM (Statistical Parametric Mapping)	Voxel-based morphometry (VBM)	Regional grey matter volume and density	sMRI, PET
FSL (FMRIB Software Library)	Voxel-wise statistical analysis of brain imaging data	VBM, tract-based spatial statistics (TBSS)	sMRI, fMRI, DTI
ANTs (Advanced Normalization Tools)	High-dimensional image registration and segmentation	Morphometric features, deformation fields	sMRI

The features extracted from these pipelines, along with clinical and demographic data, form the basis for most classification tasks. Common features include:

- Structural MRI features: Estimated Total Intracranial Volume (eTIV), normalized Whole Brain Volume (nWBV), and Atlas Scaling Factor (ASF).
- Clinical and demographic features: Age, years of education (EDUC), socioeconomic status (SES), Mini-Mental State Examination (MMSE) scores, and Clinical Dementia Rating (CDR).

3.2 RQ2: What are the most common machine learning models applied to AD classification, and what is their reported performance?

In response to this question, this section surveys the predominant machine learning models found in the literature, which are broadly categorized as supervised, ensemble, and deep learning approaches. A synthesis of recent studies is presented in Table 4.

- **Supervised Learning Models**: These models learn from labeled data to make predictions. Support Vector Machines (SVM) are frequently used due to their effectiveness in high-dimensional spaces [16], for instance, demonstrated the utility of an optimized SVM, achieving 92% accuracy on the OASIS dataset through rigorous hyperparameter tuning. Other models like

Naive Bayes (NB) have also been used; [17] showed that combining NB with k-means clustering could significantly improve performance, achieving a prediction accuracy of 91.89%. A comparative analysis by [18] investigated the utility of various classical ML algorithms on clinical data derived from the OASIS-2 dataset. The study employed correlation analysis for feature selection, identifying a core set of predictive clinical and demographic variables. Their findings indicated that a k-Nearest Neighbors (k-NN) model yielded superior performance, achieving 92.13% accuracy in a multiclass classification task (Demented, Non-Demented, Converted). This result underscores the efficacy of well-established, less complex algorithms for dementia classification, particularly when applied to structured clinical data.

- **Ensemble Learning Models**: These methods combine multiple models to achieve greater predictive performance. Boosting algorithms like AdaBoost are prominent, with one study by [19] reporting a 96.07% accuracy by focusing on non-MRI clinical data, suggesting a viable path for low-cost diagnostics. Random Forest (RF), an ensemble method first proposed by [20], is another popular choice, valued for its robustness. [13] effectively paired an RF model with the Synthetic Minority Over-sampling Technique (SMOTE) to achieve a high recall of 98% in identifying dementia cases. An anatomically informed approach was proposed by [21], who developed a methodology centered on the hippocampus, a brain structure integral to memory and significantly impacted by AD. Their framework involved the segmentation of the hippocampus from structural MRI scans to generate targeted anatomical features. An XGBoost model was subsequently trained on these features. To address the inherent class imbalance in their multiclass problem (Demented, Non-Demented, Converted), the authors applied the SMOTE algorithm. This integration of domain-specific biological information with advanced ensemble learning resulted in a classification accuracy of 94%, highlighting the strategic value of incorporating anatomical biomarkers into diagnostic models.
- **Deep Learning Models**: Deep learning has become a dominant force in medical image analysis. Convolutional Neural Networks (CNNs) are particularly adept at learning features directly from MRI and PET scans. [22] developed a CNN-based model combined with k-means clustering for noise reduction. Other studies have explored more advanced architectures, such as the modified capsule network (MCapNet) proposed by [23], which achieved an accuracy of 92.39% by better preserving the spatial relationships in brain imaging data. To address the dual challenges of class imbalance and hyperparameter optimization, [24] developed an integrated deep learning framework. The methodology utilized CNN to extract salient features from sMRI data. The training dataset was then balanced using SMOTE to prepare it for a 4-stage multiclass classification task. A key contribution of this work was the novel application of Spider Monkey Optimization (SMO), a bio-inspired metaheuristic, for fine-tuning the model's parameters to enhance classification precision. This comprehensive approach yielded a final accuracy of 91%, demonstrating the viability of hybrid frameworks that pair CNNs with advanced optimization algorithms for more robust AD detection.

Table 4. Summary of Performance for Key Machine Learning Models in AD Classification.

Reference	Model	Input Modality	Key Technique / Features	Dataset	Classification Task	Reported Accuracy	Reported AUC
[13]	Random Forest	Clinical	t-SNE for feature representation, SMOTE.	OASIS	Binary	94.00%	N/A
[16]	SVM	Clinical	Hyperparameter optimization (Grid Search).	OASIS	Binary	92.00%	0.91
[17]	Naive Bayes & k-means	Clinical	k-means clustering data sharing.	OASIS	Multiclass	91.89%	N/A
[18]	k-NN	Clinical	Correlation Analysis.	OASIS	Multiclass	92.13%	N/A
[19]	AdaBoost	Clinical	Manual feature selection, non-MRI data.	OASIS	Binary	96.07%	0.95
[21]	XGBoost	sMRI	Hippocampus Segmentation, SMOTE.	OASIS	Multiclass	94.00%	N/A
[22]	CNN & k-means	sMRI	Noise reduction via clustering.	OASIS	Binary	83.95%	0.90
[23]	MCapNet	Clinical	Kernel PCA for feature extraction	OASIS	Binary	92.39%	N/A
[24]	CNN	sMRI	SMOTE, Spider Monkey Optimization (SMO).	OASIS	Multiclass	91.00%	N/A

A synthesis of recent studies applying these models is presented in Table 4 which provides a comparative overview of the models, the input modalities and classification tasks they address, and their key performance metrics. While accuracy is a common metric, the inclusion of the Area Under the Receiver Operating Characteristic Curve (AUC) in Table 4 is crucial for a more nuanced assessment. AUC measures a model's ability to distinguish between classes across all possible decision thresholds. This is particularly important in a clinical context with imbalanced datasets, as a high AUC score indicates strong diagnostic ability that is not skewed by the prevalence of the majority class.

3.3 RQ3: What are the primary methodological challenges and research gaps limiting the clinical utility of these models?

Addressing the primary methodological challenges and research gaps is pivotal for translating machine learning models from theoretical constructs into clinically viable diagnostic tools. Despite the high classification accuracies reported in the literature, a critical examination reveals that the field confronts several fundamental obstacles that impede this transition. Our synthesis of the selected studies identifies three principal domains of challenges that consistently emerge: (1) data-related issues, which encompass the quality, accessibility, and inherent properties of the datasets used for training and validation; (2)

model-related limitations, concerning the interpretability, generalizability, and optimization of the algorithms themselves; and (3) biological complexities, which involve the pathological heterogeneity of dementia. This section delineates these challenges, providing a structured analysis of the key barriers to the clinical implementation of ML-based AD diagnostics, as summarized in Table 5.

Table 5. Key Challenges in ML for AD Detection.

Challenge Category	Specific Challenge	Common Approaches / Solutions
Data-Related	Class Imbalance	Data-level (e.g., SMOTE), Algorithm-level (e.g., Weighted Ensembles)
	Data Quality & Anomalies	Density-based clustering (e.g., DBSCAN) for outlier removal
Model-Related	Interpretability	Explainable AI (XAI) techniques (e.g., SHAP)
	Hyperparameter Optimization	Grid Search, Random Search
Biological	Pathological Heterogeneity	Development of models for patient stratification, Multimodal data fusion

A primary data-related hurdle in AD research is the pervasive issue of class imbalance. In clinical datasets such as OASIS, the distribution of subjects is naturally skewed, with healthy controls significantly outnumbering individuals in the early or transitional stages of the disease (e.g., "Converted" subjects). This imbalance poses a substantial challenge for standard ML algorithms, which are optimized for balanced datasets and consequently develop a predictive bias toward the majority class. This bias diminishes their sensitivity to the minority classes, which are often of the highest clinical interest for early detection and intervention. To mitigate this, researchers have adopted various strategies that can be broadly categorized into data-level and algorithm-level solutions. The most prevalent data-level approach is oversampling, particularly the Synthetic Minority Over-sampling Technique (SMOTE), which generates synthetic instances of the minority class to create a more balanced training set [25]. However, SMOTE is not a panacea ; as noted by [26], it can introduce artificial noise and create overlapping data points that may compromise data integrity and blur decision boundaries. Algorithm-level solutions offer an alternative by modifying the learning process itself. These include cost-sensitive learning, which assigns a higher misclassification cost to minority class instances, and the use of weighted ensemble techniques that give greater influence to classifiers performing well on underrepresented classes [27]. Compounding the issue of class imbalance is the challenge of data quality. Clinical datasets are often susceptible to noise, artifacts, and outliers, which can further degrade model performance. The application of robust anomaly detection algorithms, such as DBSCAN [28], as a preprocessing step is therefore a critical but frequently overlooked measure to ensure data integrity before model training.

Beyond data-centric issues, two significant model-related challenges hinder the clinical adoption of ML diagnostics: interpretability and generalizability. The first, interpretability, addresses the "black box" nature of many complex algorithms, particularly in deep learning. The inability to articulate the reasoning behind a model's prediction is a major barrier to clinical trust and adoption. In response, the field of Explainable AI (XAI) has emerged, developing techniques such as SHAP (SHapley Additive exPlanations) to provide transparent, feature-based justifications for individual predictions [29]. The second challenge, generalizability, refers to a model's ability to maintain its predictive performance on new, unseen data from different populations or clinical settings. A model that performs exceptionally well on its training dataset may fail to generalize due to "dataset shift," where the statistical properties of the new data differ from the training data. The work by [18] provides a crucial, citable example of this issue, demonstrating that model performance is highly contingent on the specific characteristics of the input data. Their findings underscore the necessity of rigorous, external, and cross-dataset validation to ensure that a model is robust and reliable enough for real-world clinical application.

A final, and perhaps most formidable, challenge lies in the biological heterogeneity of the disease itself. Clinical presentations of dementia are frequently not attributable to a single, "pure" pathology. Instead, a significant proportion of cases, particularly in older populations, involve mixed pathologies where the hallmark features of AD coexist with other neurodegenerative or vascular conditions, such as Lewy body dementia or cerebrovascular disease [7]. This neuropathological complexity poses a profound challenge for machine learning models. Most models are trained on well-characterized research datasets (e.g., ADNI) that often represent more homogenous, "pure" forms of AD. Consequently, these models learn to recognize feature patterns specific to AD pathology. When deployed in a real-world clinical setting, where mixed pathologies are common, the models may struggle to generalize. The presence of co-pathologies can alter the neuroimaging and clinical data in ways that deviate from the learned "pure AD" template, leading to decreased classification accuracy and reduced diagnostic confidence. This discrepancy between the idealized data used for training and the complex reality of clinical populations is a major barrier to the translation of ML models into routine practice. Future research must therefore focus on developing models that are robust to this heterogeneity, potentially through multi-label classification frameworks capable of identifying multiple co-existing pathologies or through the integration of multimodal data that can provide a more comprehensive picture of the underlying neurobiology.

4. CONCLUSION AND FUTURE DIRECTIONS

This systematic review has surveyed the application of machine learning techniques for the early detection of Alzheimer's disease, elucidating the rapid advancements and significant potential of these methods to enhance clinical diagnosis. The analysis confirms that a diverse array of models, from classical SVMs to advanced CNNs, can achieve high classification accuracy using both neuroimaging and clinical data. However, despite these promising results, the translation of these models into routine clinical practice is impeded by several persistent challenges. Key issues identified in this review include data-related problems such as class imbalance and quality control; model-related limitations concerning interpretability and generalizability; and the overarching biological complexity of pathological heterogeneity.

To bridge the translational gap between research and clinical application, future work must prioritize the development of integrated, robust, and interpretable frameworks. A promising avenue for future research lies in the creation of comprehensive systems that combine sophisticated data preprocessing—including robust anomaly detection and advanced techniques for managing class imbalance—with algorithms designed for real-world clinical data. Furthermore, there is a critical need to advance multimodal data fusion. The integration of structural and functional neuroimaging (e.g., sMRI, fMRI, PET) with non-imaging data such as genetic biomarkers and clinical assessments is essential for developing a more holistic and accurate representation of the disease process [30]. Another key direction is the development of models capable of patient stratification based on the probable mixture of underlying pathologies, moving beyond simple binary or multiclass detection toward a more personalized diagnostic paradigm. Finally, to fully capture the temporal dynamics of disease progression, continued investigation into longitudinal modeling with architectures such as Recurrent Neural Networks (RNNs) is essential. Addressing these multifaceted challenges through sustained, interdisciplinary collaboration between computer scientists, clinicians, and neuroscientists will be crucial for realizing the full potential of machine learning in the fight against Alzheimer's disease.

REFERENCES

[1] S. Vinze, J. Chodosh, M. Lee, J. Wright, and S. Borson, "The national public health response to Alzheimer's disease and related dementias: Origins, evolution, and recommendations to improve early detection," (in eng), Alzheimers Dement, vol. 19, no. 9, pp. 4276-4286, Sep 2023, doi: https://doi.org/10.1002/alz.13376.

[2] "2025 Alzheimer's disease facts and figures," Alzheimer's & Dementia, vol. 21, no. 4, p. e70235, 2025, doi: https://doi.org/10.1002/alz.70235.

[3] S. Safiri et al., "Alzheimer's disease: a comprehensive review of epidemiology, risk factors, symptoms diagnosis, management, caregiving, advanced treatments and associated challenges," (in eng), Front Med (Lausanne), vol. 11, p. 1474043, 2024, doi: https://doi.org/10.3389/fmed.2024.1474043.

[4] A. Takashima, Y. Soeda, R. Koike, and S. Maeda, "Tau Pathology and Neurodegenerative Disorders," in Aging Mechanisms II : Longevity, Metabolism, and Brain Aging, N. Mori Ed. Singapore: Springer Nature Singapore, 2022, pp. 355-366.

[5] J. A. Trejo-Lopez, A. T. Yachnis, and S. Prokop, "Neuropathology of Alzheimer's Disease," (in eng), Neurotherapeutics, vol. 19, no. 1, pp. 173-185, Jan 2022, doi: https://doi.org/10.1007/s13311-021-01146-y.

[6] A. Benussi et al., "Conceptual framework for the definition of preclinical and prodromal frontotemporal dementia," (in eng), Alzheimers Dement, vol. 18, no. 7, pp. 1408-1423, Jul 2022, doi: https://doi.org/10.1002/alz.12485.

[7] C. A. Derby, "Trends in the public health significance, definitions of disease, and implications for prevention of Alzheimer's disease," Current Epidemiology Reports, vol. 7, no. 2, pp. 68-76, 2020/06/01 2020, doi: https://doi.org/10.1007/s40471-020-00231-8.

[8] P. S. Aisen, G. A. Jimenez-Maggiora, M. S. Rafii, S. Walter, and R. Raman, "Early-stage Alzheimer disease: getting trial-ready," (in eng), Nat Rev Neurol, vol. 18, no. 7, pp. 389-399, Jul 2022, doi: https://doi.org/10.1038/s41582-022-00645-6.

[9] S. S. Bhujbal, M. M. Kad, and V. C. Patole, "Recent diagnostic techniques for the detection of Alzheimer's disease: a short review," (in eng), Ir J Med Sci, vol. 192, no. 5, pp. 2417-2426, Oct 2023, doi: https://doi.org/10.1007/s11845-022-03244-y.

[10] S. Carlson, H. Kim, D. P. Devanand, and T. E. Goldberg, "Novel approaches to measuring neurocognitive functions in Alzheimer's disease clinical trials," (in eng), Curr Opin Neurol, vol. 35, no. 2, pp. 240-248, Apr 1 2022, doi: https://doi.org/10.1097/wco.0000000000001041.

[11] A. Shukla, R. Tiwari, and S. Tiwari, "Review on alzheimer disease detection methods: Automatic pipelines and machine learning techniques," Sci, vol. 5, no. 1, p. 13, 2023, doi: https://doi.org/10.3390/sci5010013.

[12] A. Joseph and C. Jayaraman, "Machine Learning Approaches for Efficient Analysis of Neuroimaging Techniques," SHS Web of Conferences, vol. 139, p. 03027, 01/01 2022, doi: https://doi.org/10.1051/shsconf/202213903027

[13] B. A. Mohammed et al., "Multi-method analysis of medical records and MRI images for early diagnosis of dementia and Alzheimer's disease based on deep learning and hybrid methods," Electronics, vol. 10, no. 22, p. 2860, 2021, doi: https://doi.org/10.3390/electronics10222860.

[14] D. S. Marcus, A. F. Fotenos, J. G. Csernansky, J. C. Morris, and R. L. Buckner, "Open Access Series of Imaging Studies: Longitudinal MRI Data in Nondemented and Demented Older Adults," Journal of Cognitive Neuroscience, vol. 22, no. 12, pp. 2677-2684, 2010, doi: https://doi.org/10.1162/jocn.2009.21407.

[15] M. W. Weiner et al., "The Alzheimer's disease neuroimaging initiative: progress report and future plans," (in eng), Alzheimers Dement, vol. 6, no. 3, pp. 202-11.e7, May 2010, doi: https://doi.org/10.1016/j.jalz.2010.03.007.

[16] M. B. Antor et al., "A comparative analysis of machine learning algorithms to predict alzheimer's disease," Journal of Healthcare Engineering, vol. 2021, 2021, doi: https://doi.org/10.1155/2021/9917919.

[17] W. Putri, D. Hastari, K. U. Faizah, S. Rohimah, and D. Safira, "Implementation of Naïve Bayes Classifier for Classifying Alzheimer's Disease Using the K-Means Clustering Data Sharing Technique," Public Research Journal of Engineering, Data Technology and Computer Science, vol. 1, no. 1, pp. 47-54, 2023, doi: https://doi.org/10.57152/predatecs.v1i1.803.

[18] G. Waldo-Benítez, L. C. Padierna, P. Ceron, and M. A. Sosa, "Dementia classification from magnetic resonance images by machine learning," Neural Computing and Applications, vol. 36, no. 6, pp. 2653-2664, 2024, doi: https://doi.org/10.1007/s00521-023-09163-y.

[19] M. Rhman, F. Rahman, M. M. Hossain, U. H. Emu, K. Akter, and M. Mridha, "Predicting Alzheimer's Disease at Low Cost Using Machine Learning," 2021 International Conference on Science & Contemporary Technologies (ICSCT), pp. 1-5, 2021, doi: https://doi.org/10.1109/ICSCT53883.2021.9642536.

[20] L. Breiman, "Random forests," Machine learning, vol. 45, no. 1, pp. 5-32, 2001, doi: https://doi.org/10.1023/A:1010933404324.

[21] A. Balasundaram, S. Srinivasan, A. Prasad, J. Malik, and A. Kumar, "Hippocampus segmentation-based Alzheimer's disease diagnosis and classification of MRI images," Arabian Journal for Science and Engineering, vol. 48, no. 8, pp. 10249-10265, 2023, doi: https://doi.org/10.1007/s13369-022-07538-2.

[22] U. Singhania et al., "A predictive and preventive model for onset of Alzheimer's disease," Frontiers in Public Health, vol. 9, p. 751536, 2021, doi: https://doi.org/10.3389/fpubh.2021.751536.

[23] S. Basheer, S. Bhatia, and S. B. Sakri, "Computational modeling of dementia prediction using deep neural network: analysis on OASIS dataset," IEEE access, vol. 9, pp. 42449-42462, 2021, doi: https://doi.org/10.1109/ACCESS.2021.3066213.

[24] P. Gayathri et al., "Deep Learning Augmented with SMOTE for Timely Alzheimer's Disease Detection in MRI Images," International Journal of Advanced Computer Science & Applications, vol. 15, no. 2, 2024, doi: https://doi.org/10.14569/IJACSA.2024.0150253.

[25] N. Chawla, K. Bowyer, L. Hall, and W. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," J. Artif. Intell. Res. (JAIR), vol. 16, pp. 321-357, 06/01 2002, doi: https://doi.org/10.1613/jair.953.

[26] D. Elreedy and A. F. Atiya, "A comprehensive analysis of synthetic minority oversampling technique (SMOTE) for handling class imbalance," Information Sciences, vol. 505, pp. 32-64, 2019, doi: https://doi.org/10.1016/j.ins.2019.07.070.

[27] C. Chen, A. Liaw, and L. Breiman, "Using random forest to learn imbalanced data," University of California, Berkeley, vol. 110, no. 1-12, p. 24, 2004. [Online]. Available: https://api.semanticscholar.org/CorpusID:7308660.

[28] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," presented at the Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, Portland, Oregon, 1996.

[29] A. Abdellatif et al., "Computational detection and interpretation of heart disease based on conditional variational auto-encoder and stacked

ensemble-learning framework," Biomedical Signal Processing and Control, vol. 88, p. 105644, 2024, doi: https://doi.org/10.1016/j.bspc.2023.105644.

[30] S. Dwivedi, T. Goel, M. Tanveer, M. Raman, and R. Sharma, "Multimodal Fusion-Based Deep Learning Network for Effective Diagnosis of Alzheimer's Disease," IEEE MultiMedia, vol. 29, pp. 1-1, 04/01 2022, doi: https://doi.org/10.1109/MMUL.2022.3156471.