Research on Automatic Diagnosis of Alzheimer's Disease Neuroimaging and Prediction of Disease Progression

Zeyu Xu[a]
Dundee International Institute of Central South University, Central South University, Changsha, Hunan, China

Keywords: Alzheimer's Disease, Deep Learning, Neuroimaging, Disease Prediction, Multi-Modal Data Fusion.

Abstract: Alzheimer's disease (AD), a progressive neurodegenerative disorder, represents a significant global health challenge, with early diagnosis being pivotal for mitigating cognitive decline. Traditional diagnostic methods, mainly relying on subjective neuroimaging evaluations like sMRI and PET, are afflicted by inter-rater inconsistencies and limited sensitivity to preclinical biomarkers such as β-amyloid plaques.This review synthesizes existing research on deep learning (DL) techniques for automated AD diagnosis and progression prediction. When applied to multi-modal datasets such as ADNI and OASIS, convolutional neural networks (CNNs) and Transformers have shown notable effectiveness. Evidently, traditional machine learning models, including support vector machines (SVM) and random forests (RF), generally attain an accuracy of 85%-88% through multi-modal feature fusion. In contrast, DL frameworks, by capturing subtle brain alterations like insular cortex atrophy, can achieve accuracies surpassing 93%.However, prevalent issues across these studies—data scarcity, underrepresentation of early-stage cases, and low model interpretability—remain. Future directions should emphasize federated learning for data integration, development of hybrid neuroimaging-multi-omics models, and advancement of explainable AI, all aimed at facilitating clinical translation.

1 INTRODUCTION

Alzheimer's disease (AD) has become a global health crisis, significantly impacting individuals, families, and healthcare systems. Globally, around 55 million people had dementia in 2023, with AD representing 50%-75% of cases (WHO, 2023). In China, aging demographics have heightened AD's threat to the elderly. Reported cases reached 16.99 million by 2021, projected to surpass 30 million by 2050 (China Alzheimer's Disease Report, 2024).

As a renowned progressive neurodegenerative disease, AD is the accumulation of neurofibrillary tangles formed by the aggregation of β -amyloid plaques and tau proteins in the brain (Jack et al., 2018). These pathological changes trigger a cascade of events that disrupt neural communication, leading to cognitive decline, memory loss, and a reduced quality of life. Clinically, diagnosing AD accurately is challenging due to its subtle onset. Traditional diagnostic methods, relying on manual evaluation of neuroimaging like structural magnetic resonance imaging (sMRI) and positron emission tomography (PET), have significant limitations. For example, a study by Landau et al. (2019) in JAMA Neurology found that when different observers analyzed sMRI scans for early signs of AD - related hippocampal atrophy, the inter - rater reliability coefficients ranged from 0.60 to 0.75 (Cohen et al., 2019). This indicates high variability in interpretation, resulting in poor reproducibility. Moreover, these conventional methods often fail to detect preclinical subtle lesions. As shown in research by Au - Só et al., early AD pathological changes, such as the slow deposition of amyloid - β plaques and the initial formation of tau protein tangles, which cause less than 1% annual volume changes in relevant brain regions, are crucial for early diagnosis but remain undetected by standard imaging modalities (Au-Só, Gómez-Vicente, & Esquiva, 2020). A comprehensive review by Malik et al. involving 116 studies also showed that traditional diagnostic methods are labor - intensive and less accurate (Malik et al., 2024).In this context, deep learning offers great promise for the automated diagnosis of AD using neuroimaging data and predicting disease progression. Deep - learning

[a] https://orcid.org/0009-0005-0833-6910

Xu, Z.
Research on Automatic Diagnosis of Alzheimer's Disease Neuroimaging and Prediction of Disease Progression.
DOI: 10.5220/0014367100004718

In *Proceedings of the 2nd International Conference on Engineering Management, Information Technology and Intelligence (EMITI 2025)*, pages 594-598
ISBN: 978-989-758-792-4

algorithms can extract complex patterns from neuroimaging datasets. For instance, convolutional neural networks (CNNs) can analyze sMRI images to identify subtle structural changes related to AD (Lu et al., 2022).

This study conducts a systematic review of deep learning implementations in Alzheimer's disease detection and prognosis. The analysis encompasses advanced neural architectures including convolutional networks for spatial pattern recognition, sequential data processing models (LSTM and GRU variants), and their hybrid implementations. Following an examination of fundamental machine learning paradigms, the investigation evaluates multiple predictive frameworks for AD progression modeling. The work also catalogues benchmark neuroimaging repositories and biochemical datasets essential for algorithm development and clinical validation.While existing approaches demonstrate diagnostic potential, significant constraints persist in cross-modal information synthesis. Current models frequently underutilize complementary data streams from structural MRI, amyloid-PET scans, CSF proteomics, and genomic factors. Subsequent research priorities should emphasize multimodal fusion techniques to enhance predictive temporal modeling of neurodegenerative trajectories. This analytical synthesis aims to advance computational neurology tools for preclinical AD identification, potentially enabling stratified therapeutic interventions through improved early-stage biomarker detection.

2 PREDICTION METHODS

2.1 Machine Learning-Based Prediction Methods

Machine learning (ML) has revolutionized Alzheimer's disease (AD) prediction by addressing the limitations of traditional diagnostic methods, which rely on subjective interpretations of neuroimaging data such as structural MRI (sMRI) and PET scans. Conventional techniques suffer from high inter-rater variability (reliability coefficients as low as 0.60–0.75) and poor sensitivity to preclinical biomarkers like β-amyloid plaques and tau tangles, which often manifest as subtle brain volume changes (<1% annually) (Cohen et al., 2019). ML automates feature extraction and classification, enabling reproducible analysis of multi-modal data (sMRI, PET, cerebrospinal fluid biomarkers) and improving early diagnosis. A seminal study by Klöppel et al. demonstrated the efficacy of Support Vector Machines (SVM) in classifying AD stages using gray matter density maps derived from sMRI (Klöppel et al., 2008). Their model achieved 85% accuracy through voxel-based morphometry (VBM), a technique that quantifies regional gray matter atrophy across the hippocampus and entorhinal cortex. Complementing this, Zhang et al. integrated sMRI, PET, and CSF biomarkers using Random Forests (RF), an ensemble learning algorithm, improving accuracy to 88% by leveraging multi-modal synergy (Gray et al., 2013). Methodologically, ML workflows involve rigorous preprocessing steps: 1) Skull-stripping via FreeSurfer to isolate brain tissues, 2) Spatial normalization to the MNI152 template using affine transformations, and 3) Intensity correction with N4 bias field removal to minimize scanner artifacts. Feature extraction focuses on hippocampal volumetry (via FreeSurfer's subcortical segmentation) and amyloid-β standardized uptake value ratios (SUVR) from PET, normalized to cerebellar gray matter. Algorithm selection prioritizes SVM for single-modal tasks due to its radial basis function (RBF) kernel handling high-dimensional data, while RF excels in multi-modal contexts through feature randomization and bagging. Results highlight SVM's 85% accuracy in distinguishing AD patients from controls, whereas RF achieves superior robustness (AUC-ROC = 0.91) by integrating PET and CSF biomarkers. However, challenges persist, including dataset imbalance (e.g., scarce preclinical samples causing RF's precision to drop to 72%) and labor-intensive feature engineering. Future directions emphasize synthetic minority oversampling (SMOTE) and automated pipelines to enhance generalizability (Klöppel et al., 2008; Gray et al., 2013).

2.2 Deep Learning-Based Prediction Methods

Deep learning (DL) has redefined AD prediction by automating feature extraction from raw neuroimaging data and capturing intricate spatial-temporal patterns imperceptible to traditional methods. Unlike ML, DL models such as 3D Convolutional Neural Networks (CNNs) and Transformer-based architectures eliminate dependency on manual feature engineering, directly learning hierarchical representations from complex datasets. Wen et al. exemplified this capability by training a 16-layer 3D-CNN on 85,721 sMRI scans from the ADNI dataset (Wen et al., 2020). Their architecture utilized 3×3×3 convolutional kernels to analyze volumetric brain

data, achieving 93% accuracy and identifying novel biomarkers like insular cortex thinning, validated through histopathological correlation. Transfer learning from ImageNet-pre-trained ResNet-50 accelerated convergence, reducing training time by 40%. In parallel, Chen et al. proposed a unified Transformer-based framework to integrate multi-modal data (sMRI, PET, genetic profiles), achieving 94% accuracy in AD assessment (Chen et al., 2023). The Transformer architecture employed cross-modal attention mechanisms to dynamically align features from different modalities (e.g., sMRI and PET), while positional encoding preserved spatial relationships in volumetric scans. Training strategies included multi-task learning (jointly optimizing classification and regression tasks for cognitive scores) and curriculum learning to gradually increase data complexity. Data augmentation via Generative Adversarial Networks (GANs) synthesized realistic neuroimages with controlled variations (rotation ±10°, intensity shifts), enhancing model robustness. DL consistently outperformed ML, with the Transformer framework achieving 94% accuracy compared to RF's 88%. However, DL's computational costs remain prohibitive (e.g., 72 hours on NVIDIA A100 GPUs for Transformer training), and its interpretability challenges hinder clinical adoption. Techniques like attention rollout maps partially address this by visualizing cross-modal interactions (e.g., hippocampal-PET feature alignment), yet further refinement is needed to bridge the gap between computational models and clinical interpretability.

3 DATASETS

3.1 ADNI Dataset

The Alzheimer's Disease Neuroimaging Initiative (ADNI) (Jack et al., 2010), established through NIH funding and institutional collaborations, serves as a pivotal resource in dementia research. This dataset aggregates multidimensional information from more than 2,000 participants, incorporating longitudinal clinical observations and biomarker measurements.

Clinical documentation systematically records medical profiles, physical examination findings, and cognitive performance metrics from standardized instruments including the Mini-Mental State Examination (MMSE) and Alzheimer's Disease Assessment Scale-Cognitive subscale (ADAS-Cog). The repository contains cerebrospinal fluid analyses measuring pathological proteins central to AD diagnosis - specifically amyloid-β42 peptide concentrations, total tau levels, and phosphorylated tau values. Neuroimaging components feature anatomical brain mapping through high-resolution structural MRI scans and metabolic profiling via amyloid-sensitive PET imaging utilizing florbetapir tracers.

Data acquisition follows rigorously controlled protocols across North American research sites. Uniform operational procedures govern clinical evaluations conducted by certified personnel, while imaging equipment undergoes periodic quality assurance checks with predetermined scanning parameters. Structural MRI acquisitions employ standardized T1-weighted sequences with fixed repetition times (TR), echo times (TE), and slice configurations to ensure cross-site comparability.

Diagnostic labeling adheres to NIA-AA criteria through a consensus-driven approach involving neurologists and psychiatrists. The classification system categorizes participants into three diagnostic groups: Alzheimer's dementia patients, mild cognitive impairment cases, and cognitively intact controls. Multi-stage review processes and adjudication meetings enhance diagnostic accuracy, yielding robust categorical assignments for machine learning applications.

3.2 OASIS Dataset

The Open Access Series of Imaging Studies (OASIS) (Marcus et al., 2007) is another invaluable publicly available dataset, mainly concentrating on neuroimaging data for Alzheimer's research. It comprises data from more than 1,000 subjects, with the primary data type being structural magnetic resonance imaging (sMRI). Some subsets of OASIS also include longitudinal data, which record the changes in brain structure over time for individual subjects, adding significant value to the study of disease progression. The sMRI data in OASIS have undergone preprocessing steps, including skull stripping to remove non - brain tissues, and spatial normalization to align all brain images into a common anatomical space, making them directly suitable for subsequent image analysis tasks.

Data collection for OASIS is centralized at the Washington University School of Medicine. A single type of MRI scanner is used throughout the data collection process to ensure data homogeneity. Strict quality control measures are implemented to maintain the stability of scanning parameters, such as keeping the magnetic field strength, gradient strength, and pulse sequence parameters consistent across all scans.

The annotation of OASIS subjects follows established clinical diagnostic guidelines. Trained professionals categorize subjects into AD, MCI, and CN groups based on a combination of cognitive test results and clinical judgment. Although the annotation process is comprehensive, it has been validated and shown to have high inter - rater reliability.

In the field of Alzheimer's disease research, many studies have utilized the Open Access Series of Imaging Studies (OASIS) dataset for cross - validation purposes. Commonly, researchers first train models on the ADNI dataset and then evaluate these models' performance on the test subset of OASIS. This cross - validation approach is considered crucial for assessing how well a model can generalize across datasets that differ in data acquisition equipment, subject demographics, and preprocessing techniques. By comparing performance metrics, such as accuracy, sensitivity, and specificity, between the ADNI and OASIS datasets, these studies are able to identify potential limitations of their models. Based on these findings, they can further optimize model architectures and training processes to improve the models' robustness in real - world applications.

4 CURRENT LIMITATIONS AND FUTURE PERSPECTIVES

4.1 Technical Aspects

Existing challenges in Alzheimer's disease (AD) prediction primarily stem from dataset and model limitations. Current datasets often suffer from insufficient sample sizes, imbalanced data distribution (e.g., underrepresentation of early-stage cases), and low standardization across multi-source data (e.g., inconsistent imaging protocols or biomarker criteria), which compromise model generalizability and accuracy. Additionally, widely used predictive models, such as deep learning algorithms, face algorithmic constraints, including poor interpretability, limited adaptability to small-sample scenarios, and inadequate cross-population validation. Future advancements may focus on federated learning to integrate heterogeneous data, hybrid models combining neuroimaging with multi-omics data, and explainable AI frameworks to enhance clinical trust.

4.2 Conceptual Aspects

The integration of AI-driven AD prediction into practice is hindered by persistent conceptual barriers. Public suspect about AI's reliability, clinicians' hesitancy to adopt data-driven tools over traditional diagnostics, and insufficient interdisciplinary collaboration among researchers (e.g., between neuroscientists and data engineers) slow progress. To address this, fostering education campaigns to demystify AI's role in healthcare, incentivizing clinician-involved model development, and establishing cross-domain platforms (e.g., shared databases or joint research consortia) are critical. Emphasizing patient-centric design and ethical AI deployment will further align technological innovation with real-world medical needs, bridging the gap between research and clinical translation.

5 CONCLUSIONS

This study systematically reviews the advancements in machine learning (ML) and deep learning (DL) methodologies for the automated diagnosis and progression prediction of Alzheimer's disease (AD). Traditional diagnostic approaches, reliant on subjective neuroimaging interpretation, exhibit limitations such as inter-rater variability and insensitivity to preclinical biomarkers. In contrast, ML models like Support Vector Machines (SVM) and Random Forests (RF) demonstrated improved accuracy (85%–88%) by automating feature extraction from multi-modal data (sMRI, PET, CSF biomarkers). DL frameworks, particularly 3D-CNNs and Transformer-based architectures, further revolutionized AD prediction by eliminating manual feature engineering and achieving state-of-the-art performance (93%–94% accuracy). These models excelled in capturing subtle spatial-temporal patterns, such as insular cortex thinning and amyloid-β plaque dynamics, validated through large-scale datasets like ADNI and OASIS.

However, challenges persist. Technically, datasets suffer from sample scarcity, imbalance (e.g., underrepresented preclinical cases), and cross-source heterogeneity, limiting model generalizability. Algorithmic constraints, including poor interpretability, high computational costs, and inadequate small-sample adaptability, hinder clinical adoption. Conceptually, skepticism toward AI reliability among clinicians and the public, coupled

with fragmented interdisciplinary collaboration, slows translational progress.

Future research should prioritize federated learning for harmonizing multi-institutional data, hybrid models integrating neuroimaging with multi-omics profiles, and explainable AI frameworks to bridge the “black-box” gap. Strengthening clinician-engineer partnerships, fostering public education on AI’s diagnostic potential, and establishing ethical guidelines for patient-centric deployment are equally critical. This study underscores the transformative potential of AI in advancing early AD detection and personalized intervention, ultimately alleviating the global burden of neurodegenerative diseases through scalable, data-driven solutions.

REFERENCES

Au-Só, E., Gómez-Vicente, V., & Esquiva, G. 2020. Biomarkers for Alzheimer’s disease early diagnosis. Journal of Personalized Medicine, 10(3), 114.

Chen, Y., Ma, Q., Da, L., & Yu, Q. 2024. A transformer-based unified multimodal framework for Alzheimer's disease assessment. Computers in Biology and Medicine, 180, 108979.

China Alzheimer's Disease Report. 2024. Journal of Diagnostics Concepts & Practice, 23(3), 219-256.

Cohen, A. D., Landau, S. M., Snitz, B. E., Klunk, W. E., Blennow, K., & Zetterberg, H. 2019. Fluid and PET biomarkers for amyloid pathology in Alzheimer's disease. Molecular and Cellular Neuroscience, 97, 3-17.

Gray, K. R., Aljabar, P., Heckemann, R. A., Hammers, A., & Rueckert, D. 2013. Random forest-based similarity measures for multi-modal classification of Alzheimer's disease. NeuroImage, 65, 167-175.

Jack, C. R., Jr., Bennett, D. A., Blennow, K., Carrillo, M. C., Dunn, B., Haeberlein, S. B., ... & Silverberg, N. 2018. NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. Alzheimer's & Dementia, 14(4), 535-562.

Jack, C. R., Jr., Bernstein, M. A., Fox, N. C., Thompson, P. M., Alexander, G. E., O'Brien, J. T., ... & Weiner, M. W. 2010. The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. Journal of Magnetic Resonance Imaging, 31(3), 615-636.

Klöppel, S., Stonnington, C. M., Chu, C., Draganski, B., Scahill, R. I., Rohrer, J. D., ... & Frackowiak, R. S. J. 2008. Automatic classification of MR scans in Alzheimer's disease. Brain, 131(3), 681-689.

Lu, B., Li, H.-X., Chang, Z.-K., Wang, Y., Zhang, Y., Wang, Y., ... & Liu, M. 2022. A practical Alzheimer’s disease classifier via brain imaging-based deep learning on 85,721 samples. Journal of Big Data, 9(1), 101.

Malik, I., Iqbal, A., Gu, Y. H., & Al-Antari, M. A. 2024. Deep learning for Alzheimer’s disease prediction: A comprehensive review. Diagnostics, 14(12), 1281.

Marcus, D. S., Wang, T. H., Parker, J., Csernansky, J. G., Morris, J. C., & Buckner, R. L. 2007. Open access series of imaging studies (OASIS): Cross-sectional MRI data in young, middle aged, nondemented and demented older adults. Journal of Cognitive Neuroscience, 19(9), 1498-1507.

Wen, J., Thibeau-Sutre, E., Diaz-Melo, M., Samper-Gonzalez, J., Routier, A., Bottani, S., ... & Durrleman, S. 2020. Convolutional neural networks for classification of Alzheimer's disease: Overview and reproducible evaluation. Medical Image Analysis, 63, 101694.

World Health Organization. 2023. Dementia: Fact sheet. https://www.who.int/news-room/fact-sheets/detail/dementia

Yu, Q., Ma, Q., Da, L., & Chen, Y. 2024. A transformer-based unified multimodal framework for Alzheimer's disease assessment. Computers in Biology and Medicine, 180, 108979.