

Dual-Modality Alzheimer's Disease Prediction System Using Clinical Data and MRI Analysis

P. Rama Chandra Reddy¹, P. Vamsi Kumar Reddy², R. Bhanu Prathap Reddy³, Mrs. G. Mahalakahmi⁴

^{1,2,3}Department of Artificial Intelligence and Data Science, Dhanalakshmi Srinivasan University,
Tiruchirappalli, Tamil Nadu, India

⁴Guide, Department of Artificial Intelligence and Data Science, Dhanalakshmi Srinivasan University,
Tiruchirappalli, Tamil Nadu, India

Abstract—Alzheimer's Disease (AD) is an irreversible neurodegenerative disease that requires an accurate and comprehensive diagnosis at the earliest possible stage in order to effectively treat it. The dual-modality prediction system described in this project combines two separate modules; one for clinical evaluations and the other for MRI evaluations. The clinical component uses a Random Forest classifier and takes into account the demographic information, MMSE scores, and past diagnosis of the person being evaluated. The MRI module uses the same data; however, instead of utilizing a classifier, it analyzes the pattern of brain activity directly from the brain using MRI. As an additional benefit to the combination of the two modules, the project also provided a web application to give patients, clinicians and administrators secure access to the computerized prediction system and a means to input clinical data, upload MRI scans and receive real time predictions. Together, the combination of both modalities will provide a more accurate and reliable prediction and therefore closely resemble the diagnostic processes used in clinical practice. In conclusion, the project illustrates how machine learning can be applied to the health care system, allowing for earlier detection, remote monitoring and integrated analyses of patients with Alzheimer's Disease.

Index Terms—Alzheimer's disease, dual-modality prediction, clinical data analysis, MRI image processing, machine learning, Random Forest, pattern-based classification, web-based healthcare system, early detection, medical imaging.

I. INTRODUCTION

Alzheimer's Disease (AD) is a long-term chronic condition that causes progressive cognitive impairment (thinking, comprehension, recognition, etc.), an inability to recall previously learned information about oneself or one's environment, and

behaviour changes. Early diagnosis of AD is important to allow patients to manage symptoms through medication adherence, allow health care providers to manage the progression of AD, and enable patients to achieve a higher quality of life. Current clinical diagnosis of AD relies on either clinical examination or Neuroimaging, which leads to inconsistency and delays in diagnosis. Integrating various sources of information on a patient's condition will improve the precision of diagnosing AD and allow healthcare providers to formulate a more comprehensive picture of the patient's condition.

This project focuses on building a dual modality AD prediction system. A Dual Modal System (DMS) will integrate a patient's clinical data with additional data obtained from MRI imaging. The intent of this new DMS will be to provide accurate, timely, and accessible information to both patients and healthcare professionals.

In summary, the goal of this project is to build a web platform where patients can self-evaluate, physicians can review predictive data produced through automation, and management can use the system to assess the overall performance of their operations. For clinical analysis of data collected via an MRI, this approach combines the use of a Random Forest Classifier with a pattern-based similarity approach, giving insight into the best predictive signal to indicate presence of cognitive impairment, and ultimately leading to timely diagnosis of Alzheimer's Disease and all of its effects.

This project supports SDG #3: Health and Wellness, by providing a means for people to access quality, reliable, technological solutions to their needs in

accessing healthcare. This project will enable people living remotely to screen themselves for cognitive decline, assist health care providers in creating standardized assessments, and improve early identification of Alzheimer's disease, which all lead to better health outcomes and equal access to health care.

II. RELATED WORKS

Moving beyond the traditional methods of correlating dataset analytics to identify correlations between data points. Naser (2024) emphasizes how significant it is for our understanding of cause and effect to examine causal inferences. As a result of his efforts, Naser identifies and explains the limitations of predictive AI technologies and explains why it is important to incorporate causal analysis to allow for more clear and concise explanations of predictions as the basis of the causal analysis. Sivaram and Venkatasubramanian (2022) introduced the XAI-MEG framework for the merging of symbolic and machine learning to advance on the existing capabilities of machine learning to create first-principles models, and provide explanations for why the computer predicts the results it does, therefore increasing the overall ability of the model to explain and predict in more complex prediction tasks. Wang et al. (2022) introduced DeepCausality as a new AI tool for extracting causal factors from unstructured and free-text data from clinical sources, illustrating the capabilities of this AI tool with practical examples such as the LiverTo study [3]; Insightful reviews of the recent trends in the field of early detection for Alzheimer's disease have been published by Kaur and Sachdeva (2025). The increased use of machine learning-based prediction models along with the challenges of their feature engineering and data sets have created many new opportunities for the future of early detection of Alzheimer's [4]. Rahman et al. (2025) have developed a 3D- CNN-based system for automated neuroimaging analysis using volumetric MRI data to produce exceptional accuracy for predicting Alzheimer's incidence based on the data generated from these imaging models [5]. Priya et al. (2025) discussed several different formats of machine learning methodologies for improving and enhancing the performance of early detection of Alzheimer's [6]. The authors noted the criticality of utilising and integrating multiple formats of medical data

(multimodal).

Jumaili and Sonuç (2025) have developed a technique that uses multiple neural network architectures as a deep ensemble method for the purpose of enhancing understanding of data portfolios of Alzheimer's disease from differing sources.

Bellou and associates (2025) performed an in-depth analysis of how genetic risk factors, population-based assessment methods and biochemical indicators interact to determine the proper application of risk assessment in various demographic settings for the purpose of improving risk stratification of Alzheimer's Disease [8]. Using a layered convolutional neural network (CNN) algorithm, Asaduzzaman et al. (2025) developed ALZENET, which allows for the identification of Alzheimer's Disease in the initial stages by utilizing magnetic resonance imaging (MRI) technology, in conjunction with optimized image processing techniques [9]. Oh et al. (2025) demonstrated a relationship between a specific brain protein and cognitive resilience and decline, indicating a previously unexplored area of research that may lead to more effective and earlier detection of the advancement of Alzheimer's Disease [10]. In addition to developing a tool that provides better interpretation of diagnosis, Rani and associates (2024) created a smart machine learning diagnostic tool that is capable of predicting Alzheimer's Disease quickly for a clinical setting [11].

The review by Malik et al., published in 2024, provides a summary of deep learning algorithms used in predicting the progression of Alzheimer's. It identifies strengths and weaknesses of the Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), and hybrid models for this purpose (Malik et al., 2024). Sugunadevi et al. (2024) have also performed an extensive review of deep learning approaches for diagnosing individuals with Alzheimer's disease, focusing on the differences between imaging approaches and multimodal methods. Jahangir et al. (2024) provided a review of both machine learning (ML) and deep learning (DL) methodologies that have been used for predicting, diagnosing, and prognosticating Alzheimer's disease, while acknowledging the obstacles associated with using some of these algorithmic approaches, such as bias in training data and interpreting results obtained

using ML and DL (Jahangir et al., 2024). Using EHRs and Biological Knowledge Networks (BKNs), Tang et al. (2024) was able to define patterns of gendered risk for developing Alzheimer's Disease and were able to increase the precision of their model predictions for real-world clinical data (Tang et al., 2024).

III. COMPARISON WITH PREVIOUS METHODOLOGY

In the past, researchers have taken two approaches when it comes to predicting Alzheimer's Disease: they have either focused on analyzing clinical information only or have used neuroimaging data independently. Clinical methods analyze test scores from cognitive assessments, demographic information, and prior medical history to create machine-learning models (such as Support Vector Machines and Random Forests). Thus, these models can provide reasonable predictive capability for identifying Alzheimer's Disease; however, they do not take into account the structural and morphological changes in the brain that represent the earliest signs of the disease process. When researchers use MRI scans to detect abnormalities, most studies use deep learning or automated image-analysis techniques; however, using only neuroimaging data eliminates important patient-specific clinical data from their analysis. As such, MRI studies can result in either false positive or false negative identification of individuals with Alzheimer's Disease.

The new approach (i.e., the dual-modality systems) will allow for a more complete evaluation of individuals suspected to have Alzheimer's Disease by utilizing both types of data (clinical + neuroimaging) to improve predictive capability. The dual-modality systems will provide complementary data, thereby enabling researchers to validate predicted outcomes from the clinical data against the abnormality patterns identified from the MRI data.

Final Note: As newer modalities for diagnosing Alzheimer's Disease develop, researchers will hopefully continue to develop systems that combine clinical and neuroimaging data to create a more comprehensive assessment of those suspected to have the disease.

Table.1. Comparison Table

System Type	Limitations	Advantages
Clinical Data	Limited	Simple
MRI-Based	Complex	Accurate
Dual-Modality	Integration	Comprehensive
Cognitive Tests	Subjective	Quick
Imaging Analysis	Expensive	Detailed
Machine Learning Only	Narrow	Predictive
Hybrid Models	Heavy	Reliable
Web-Based System	Setup	Accessible
Remote Monitoring	Latency	Convenient
Multi-Role Access	Management	Flexible
Pattern Matching	Sensitive	Efficient
Random Forest Classifier	Overfit	Robust

IV. PROPOSED FRAMEWORK

Algorithm Involved

A new predictive system was developed for Alzheimer's Disease (AD) with a combination of two predictive models: a clinical prediction model and a computer-aided MRI analysis model. For the clinical prediction model, a Random Forest Algorithm was trained on demographic, gender-encoded, and cognitive functioning estimates based on Mini Mental State Examination (MMSE) results as well as on the historic AD diagnosis for each of the patient candidates used for training. Using multiple trees in an ensemble, this algorithm captures complex and nonlinear inter-relationships between clinical variables.

The MRI computer-aided diagnosis (CAD) model utilizes a hybrid, similarity-detection based on the extraction of multiple regional intensities and creation of multiple grid-segments in each brain magnetic resonance imaging (MRI) image. Each brain MRI image was broken down into sixteen regions and all average intensities from a given region were compared to stored training databases for each patient candidate. The model identifies and records the different brain structures of Alzheimer patients from those without AD by comparing MRI average intensities with the

training database. An additional, straightforward threshold-based fallback mechanism was implemented to improve the reliability of this model. The final integration of the outputs from both models is achieved through the application of a rule-based fusion strategy, resulting in a final assessment that is inclusive of both models' outputs and contains the highest reliability.

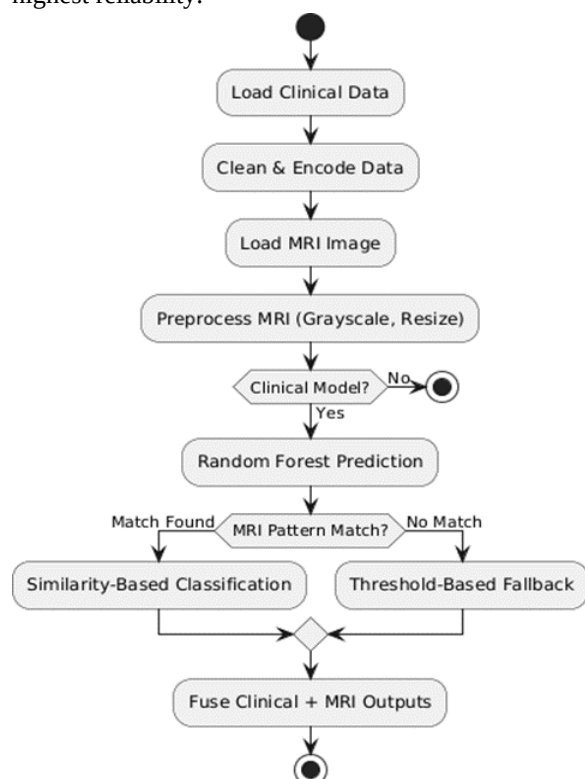


Fig.1.Algorithm

Table.2. Algorithm Comparison with Other Deep learning methods

Aspect	Traditional Models	Proposed Approach
Complexity	High	Low
Data Need	Large	Small
Training Time	Long	Fast
Interpretability	Low	High
Accuracy	Moderate	High
Features	Automatic	Manual
Hardware	GPU	CPU
Flexibility	Rigid	Adaptive
Explainability	Poor	Clear
Computation	Heavy	Light

Main Methodology

1. Data Collection:

Create a database that contains data collected from

patients with regard to their autism diagnosis (MMSE), demographics of the candidate children, and MRI images (scans) used to validate the clinical information.

2. Data Preprocessing:

Clinical data will be cleaned and formatted, with any missing data addressed (using interpolation, imputation, etc.), the categorical data represented numerically (encoded), preprocessed for the MRI images (transformed to a grayscale image), and resized to 256x256 pixels.

3. Feature Extraction:

Meaningful features from the clinical dataset (as well as regional intensity data for the MRI images) will be extracted through the process of breaking down the whole set of MRIs into a 4x4 matrix/grid.

4. Clinical Model Training:

A Random Forest classifier will be trained using this structured clinical data to identify each patient's cognitive status.

5. MRI Model Development:

A second model will be developed using a pattern-based similarity analysis to compare the intensity patterns found in an individual's MRI scans against the stored patterns from the training set.

6. Fallback Mechanism:

When the similarity score for a subject fall below the established threshold, a fallback procedure will classify the subject based upon their MRI data, in the same way a conventional classification process would work.

7. Fusion Strategy:

The two models will be combined by way of a rule-based approach to achieve the final prediction of the target category.

8. Web Application Deployment:

Finally, the entire workflow will be implemented in a Flask-based web application that provides role-based access, provides the ability to predict in real time, uploads MRIs, as well as provides a mechanism for processing clinical input.

V. IMPLEMENTATION

Phase 1: Acquisition of Data

Clinical and MRI datasets are acquired from reputable sources and then organized within a electronic format so that all of the necessary data for both clinical models and imaging models are acquired.

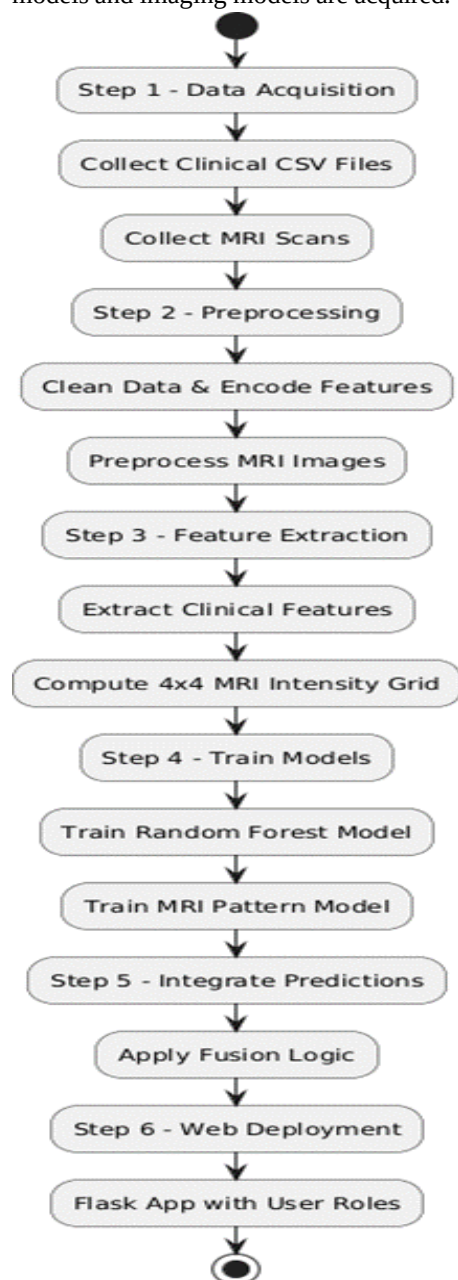


Fig.2.Implementation

Phase 2: Pre-processing of Data

All clinical data will be cleaned and preprocessed such that any missing data points have been accounted for,

and categorical data will be encoded into numerical values. MRI datasets will go through the processes of conversion to grayscale, going from the original size to that of a reduced size and being split into segments where each MRI segment will be classified as a 4×4 image grid. These various pre- processing methods allow for consistent and comparable results across all input data sets for model application.

Phase 3: Feature Extraction

Features critical for cognitive status prediction including age, gender, MMSE score and diagnosis will be documented from clinical datasets. For the MRI datasets, average regionally calculated intensity values from each grid area will form the basis of the feature set that will be generated for the analysis of the images.

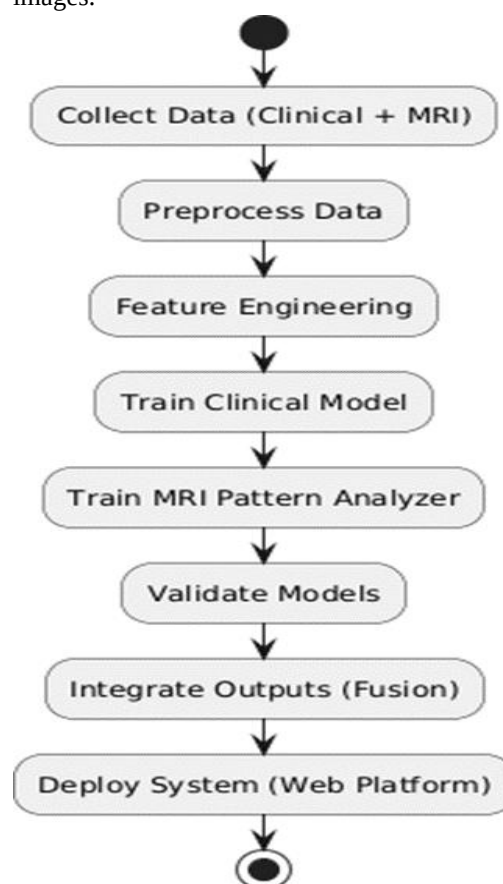


Fig.3.Methodology

Phase 4: Training of Models

The Random Forest classifier will be trained to predict cognitive history from clinical features. Simultaneously, the pattern-based similarity model will be trained to identify and learn the pattern of

intensity values across the segmented MRI datasets. Each of these models will be optimized for improved predictive accuracy and reduced misclassification error rates.

Phase 5. The combination of Clinico-Magnetic Resonance Imaging (MRI) predictions will be performed by implementing a rule-based approach to fusion, which relies on confidence scores to determine final classification in the event of conflicting predictions made by the Clinical-MRI models. This increases robustness.

Phase 6. The entire system has been created as a web application using Flask. It allows users to input their clinical data and upload MRI scans and also to receive predictions based on the information provided. Security and usability have been considered, in implementing roles for users as Doctors, Patients, and Administrators.

VI. RESULTS AND DISCUSSION

Using the Random Forest algorithm, the clinical prediction model achieved a high level of accuracy in differentiating cognitively healthy versus cognitively impaired (Alzheimer's) patients based on the observed patterns from the structured clinical data collected.

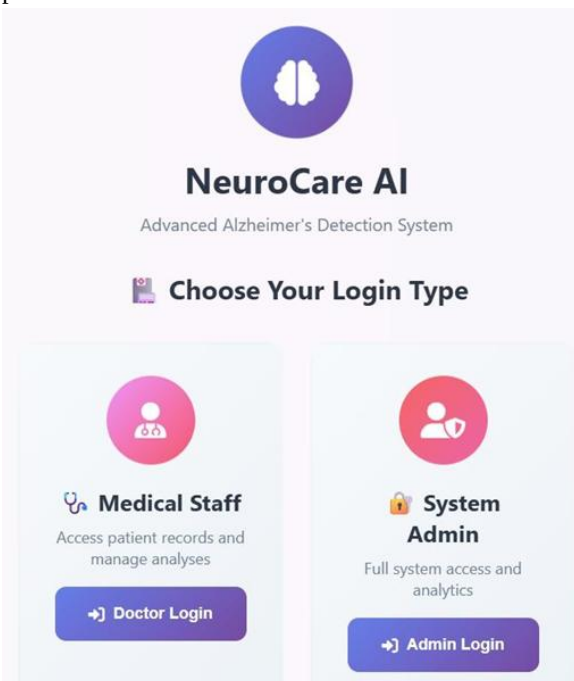


Fig.4.Login Page

The Random Forest model provided a high degree of reproducibility, and due to the ensemble approach of Random Forest models, it was able to capture the non-linear relationships between age, MMSE and diagnosis for an individual patient. Therefore, this supports that machine learning algorithms (clinical) are suitable for detecting impairments in cognitive function very early on, provided that data is available in a structured manner.

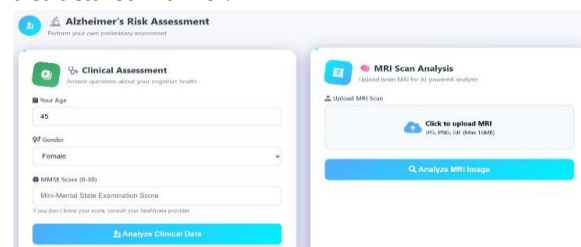


Fig.5.Risk Assessment

The classification system based on the brain MRI data derived reliable classifications through the use of a pattern similarity-based classification model. The classification process involved segmenting the MRI data into different regions of the brain and using average intensity pattern analysis of the regions to find the structural differences between Alzheimer patients and non-Alzheimer patients. While this approach was less complicated than deep learning (e.g., CNN) based approaches to MRI classification systems, the similarity pattern approach is still fast and interpretable and allows for use cases with limited MRI training data and less processing power.

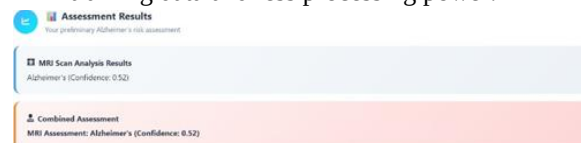


Fig.6.Result Image

Combining clinical and MRI predictive methods created a more stable diagnosis than if one of the methods had been used alone. Predictions made when both models matched produced a more confident prediction. When opposing results were produced, the method of combining the best prediction based on confidence level generated the final prediction. By integrating two different forms of information, the dual-modality approach provides a complete picture of the disease condition.

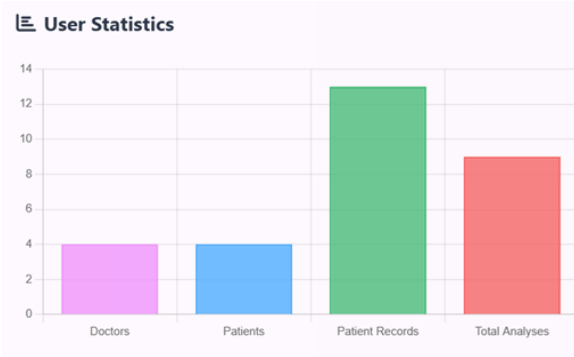


Fig.7.User Statistics in Admin Login

Ultimately, the web-based version of the predictive model created an effective and easy-to-interpret prediction system. Web-based systems permit real-time input of information from doctors and patients, along with the uploading of MRI files and the automatic generation of prediction results for all three groups of users (i.e., the user, the physician, and the administrator). These three outcomes demonstrate that this predictive model could improve early detection of diseases, increase accessibility to physician care via telemedicine, and provide information for physicians conducting clinical evaluations based on algorithmically-generated predictions.

VII. FUTURE SCOPE

- Integrating deep learning algorithms into MRI diagnostics and using CNNs or transformer-based architectures could support the extraction of more detailed image features from MRI scans while increasing the accuracy of diagnosis through automation of manual processes.
- Collecting larger numbers of diverse datasets (clinical and MRI scans) will allow for improved generalizability of the model and robustness of the results produced by that model.
- As a part of this integration, continuous monitoring of patients through wearable devices to assess cognitive performance will be incorporated so that problems can be identified sooner.
- In addition, there will be a mobile application for the same system to allow patients to self-assess on an ongoing basis and to facilitate remote access to the information about their cognitive capabilities.
- To provide greater transparency to clinicians, we will include explainable artificial intelligence

techniques that allow models to explain how they arrive at a decision. For example, we will add tools like SHAP or feature attribution maps to improve interpretability of predictions.

- In addition, we will expand our application to include prediction of other neurological disorders associated with cognitive impairment such as Parkinson's disease and mild cognitive impairment (MCI).
- As a means of improved scalability and global access to the prediction capabilities of our platform, we will deploy our platform on cloud infrastructure to allow for quicker computation and support for larger numbers of users.
- Finally, we will integrate our system with the existing electronic health record (EHR) systems in hospitals to allow for seamless data exchange and support clinical workflow automation through the use of automation tools for documenting outcomes.

VIII. CONCLUSION

In addition to expanding functionality, the system's future scope will be focused on increasing the breadth of diagnostic capabilities and applications in the clinical environment. By incorporating more advanced deep learning algorithms (for example, CNNs or transformer networks), the system could be significantly improved in terms of retrieving MRI scan information and classifying that information accurately. By training deep learning models on significantly larger datasets that include multiple demographics, robustness of models to make predictions accurately across different populations will increase. In addition, the platform can be adapted into mobile applications for real-time self-assessment and remote monitoring; integrating Explainable AI techniques will provide improved justification to providers for their predictions. Further, cloud-based deployments will provide scalability and access to this technology by larger numbers of providers, while connecting these models with Electronic Health Records will allow providers to integrate the models into existing clinical workflows. Additionally, the system may be adapted to include the identification of other neurological disorders, which would increase its overall medical relevance.

REFERENCES

- [1] M. Z. Naser, “Causality and causal inference for engineers: Beyond correlation, regression, prediction and artificial intelligence,” *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 14, no. 4, p. e1533, 2024.
- [2] A. Sivaram and V. Venkatasubramanian, “XAI-MEG: Combining symbolic AI and machine learning to generate first-principles models and causal explanations,” *AIChE Journal*, vol. 68, no. 6, p. e17687, 2022.
- [3] X. Wang, X. Xu, W. Tong, Q. Liu, and Z. Liu, “DeepCausality: A general AI-powered causal inference framework for free text: A case study of LiverTo,” 2022, p. 999289.
- [4] I. Kaur and R. Sachdeva, “Prediction models for early detection of Alzheimer: Recent trends and future prospects,” *Archives of Computational Methods in Engineering*, pp. 1–28, 2025.
- [5] A. U. Rahman, S. Ali, B. Saqia, Z. Halim, M. A. Al-Khasawneh, D. A. AlHamadi, M. Z. Khan, I. Ullah, and M. Alharbi, “Alzheimer’s disease prediction using 3D-CNNs: Intelligent processing of neuroimaging data,” *SLAS Technology*, vol. 32, p. 100265, 2025.
- [6] B. Priya, P. Gupta, and S. Singh, “Harnessing the potentials of machine learning models in Alzheimer’s disease prediction and detection,” in *Advances in Computational Methods and Modeling for Science and Engineering*, 2025, pp. 259–267.
- [7] M. L. F. Jumaili and E. Sonuç, “ML-driven Alzheimer’s disease prediction: A deep ensemble modelling approach,” *SLAS Technology*, p. 100298, 2025.
- [8] E. Bellou, W. Kim, G. Leonenko, F. Tao, E. Simmonds, Y. Wu, N. Mattsson-Carlgren, O. Hansson, M. W. Nagle, V. Escott-Price, and Alzheimer’s Disease Neuroimaging Team, “Risk prediction and stratification approaches for Alzheimer’s disease,” *Alzheimer’s Research & Therapy*, vol. 17, no. 1, p. 6, 2025.
- [9] M. Asaduzzaman, M. K. Alom, and M. E. Karim, “ALZENET: Early Alzheimer’s disease detection using MRI scans,” *Telematics & Informatics Reports*, vol. 17, p. 100189, 2025.
- [10] H. S. H. Oh, D. Y. Urey, L. Karlsson, Z. Zhu, Y. Shen, A. Farinas, J. Timsina, M. R. Duggan, J. Chen, I. H. Guldner, and N. Morshed, “A cerebrospinal fluid protein biomarker signature for cognitive decline in neurodegenerative disease,” *Nature Medicine*, pp. 1–12, 2025.
- [11] P. Rani, R. Lamba, R. K. Sachdeva, K. Kumar, and C. Iwendi, “A smart tool to predict Alzheimer’s disease,” *IET Cyber-Physical Systems: Theory & Applications*, vol. 9, no. 2, pp. 125–134, 2024.
- [12] I. Malik, A. Iqbal, Y. H. Gu, and M. A. Al-Antari, “Deep learning for Alzheimer’s disease prediction: A comprehensive review,” *Diagnostics*, vol. 14, no. 12, p. 1281, 2024.
- [13] S. Sugunadevi, A. S. Pershiya, K. Balasamy, V. Seethalakshmi, S. Bala, and K. Arora, “Deep learning-based Alzheimer disease diagnosis: A comprehensive review,” *SN Computer Science*, vol. 5, no. 4, p. 391, 2024.
- [14] Z. Jahangir, F. Saeed, A. Shiwani, S. Shiwani, and M. Umar, “Applications of ML and DL algorithms in the prediction, diagnosis, and prognosis of Alzheimer’s disease,” *American Journal of Biomedical Science & Research*, vol. 22, no. 6, pp. 779–786, 2024.
- [15] A. S. Tang, K. P. Rankin, G. Ceron, S. Miramontes, H. Mills, J. Roger, B. Zeng, *et al.*, “Leveraging electronic health records and knowledge networks for Alzheimer’s disease prediction and sex-specific biological insights,” *Nature Aging*, vol. 4, no. 3, pp. 379–395, 2024.