Improving Early Alzheimer's Disease Diagnosis Using Machine Learning on Clinical and Demographic Data

Abdur Rahman[1], Md Shahriar Mannan Prottoy[1], Mahtab Chowdhury[1], Md Hasibul Hasan Shanto[2], Mohammad Armanul Hoque[3], and Sarmin Rahman Mim[1]

[1]Department of Computer Science and Engineering, Daffodil International University, Dhaka, Bangladesh
[2]Department of Information Technology, Charles Darwin University, Sydney, Australia
[3]Department of Electrical and Electronics Engineering, Daffodil International University, Dhaka, Bangladesh
[1]Corresponding author: abdur15-4270@diu.edu.bd*; Prottoy2205101490@diu.edu.bd, chowdhury22205101076@diu.edu.bd, shantoaussie097@gmail.com, haque33-1719@diu.edu.bd, sarmin15-4294@diu.edu.bd

Abstract: Early diagnosis of Alzheimer's disease (AD) significantly improves patient outcomes, yet most studies achieving greater than 95% accuracy rely on expensive MRI/PET imaging or deep learning models that are infeasible in low-resource settings. This work demonstrates that simple, interpretable machine learning models can achieve comparable performance using only routinely collected, low-cost clinical and demographic data (n=2149). Five classifiers (Logistic Regression, Random Forest, Gradient Boosting, SVC, and soft Voting Classifier) were compared using 10-fold cross-validation with SMOTE applied only on training folds to address class imbalance. Gradient Boosting yielded the best results: 95.81% accuracy, 95.5% precision, 95.0% recall, 95.2% F1-score, and 0.99 AUC. By requiring no neuroimaging or specialised infrastructure, the proposed approach provides an immediately deployable screening tool for primary-care and resource-limited environments. Limitations include moderate dataset size and limited ethnocultural diversity. Future work will focus on external validation on multi-ethnic cohorts (ADNI, OASIS), integration of low-cost multimodal data, and transfer learning to enhance generalisability.

Keywords: Alzheimer's Disease, Classification Models, demographic, Machine Learning

1 Introduction

Early and accurate diagnosis of Alzheimer's disease (AD) is critical for timely intervention. While deep learning models such as CNNs and transfer learning have significantly improved multi-stage AD detection from neuroimaging and multimodal data [8], they require expensive infrastructure and are impractical in low-resource settings. This paper demonstrates that simple machine learning on routinely collected clinical and demo-

M. S. Arefin et al. (eds.), *Proceedings of the International Conference on Intelligent Data Analysis and Applications (IDAA 2025)*, Advances in Intelligent Systems Research 206,
https://doi.org/10.2991/978-94-6239-664-7_3

graphic data alone can achieve comparable performance, enabling immediate deployment as a cost-free screening tool in primary-care environments worldwide.

Multimodal deep learning models integrating MRI, clinical, and genetic data have advanced AD stage classification [20]. Despite superior accuracy, these approaches remain dependent on costly imaging and laboratory resources, limiting real-world deployment in primary-care settings.

Large-scale initiatives like the Alzheimer's Disease Neuroimaging Initiative (ADNI) have provided extensive clinical, demographic, and biomarker data over the past 15 years, significantly advancing early-detection research [19]. However, most cohorts lack ethnocultural diversity, and high-accuracy models still depend on costly biomarkers, limiting global applicability.

EEG-based classification using Hjorth parameters and wavelet features has emerged as a low-cost, non-invasive alternative for early AD detection [16]. Despite improved accuracy over traditional methods, EEG still requires specialized equipment and technical expertise, limiting widespread use in primary-care settings.

CNN-Ensemble models applied to MRI slices have achieved high accuracy in AD and MCI classification by automating feature extraction and improving robustness [14]. Despite strong performance, these methods remain dependent on costly MRI infrastructure, rendering them impractical for primary-care and low-resource settings.

Non-invasive blood-based biomarkers, such as brain-extracellular vesicles, show promise for cost-effective early AD detection [3]. Despite potential advantages, they remain investigational, require laboratory processing, and are unavailable in most primary-care settings.

Despite the high performance of imaging-based and deep learning approaches, they remain impractical in low-resource settings. The core novelty of this work lies in achieving 95.81% accuracy and 0.99 AUC using only low-cost, routinely collected clinical and demographic features (n=2149) without any neuroimaging, genetic markers, blood tests, or deep learning infrastructure. This makes the proposed Gradient Boosting model immediately deployable as a cost-free, interpretable screening tool in primary-care centres and developing countries where MRI/PET is unavailable — a practical advantage not offered by existing high-accuracy studies.While prior works [e.g., 1, 4, 20] achieve similar accuracies using multimodal neuroimaging, our approach uniquely eliminates dependency on expensive infrastructure, enabling deployment in 80% of global primary-care facilities lacking MRI access (as per WHO estimates).

2 LITERATURE REVIEW

The recent developments in AD understanding and its psychophysiology, diagnostic criteria, and treatment modalities. The authors point to the increasing significance of biomarkers, which contribute to an earlier diagnosis of AD, allowing timely intervention. Existing biomarkers like amyloid-beta and tau have helped to improve AD diagnosis, but recognize the problem of their unavailability and unreliability for clinical application. The authors draw attention to the usefulness of the notion of "continuum of AD", as the disease moves from preclinical to mild cognitive impairment and then to dementia. Despite the advancements, the review highlights that drug development continues to be a problem since not a single disease-modifying therapy has been able to stop disease progression in AD.[13]

Neuroimaging techniques (MRI, PET, tau imaging) remain essential for detecting presymptomatic brain changes and biomarkers in AD, yet suffer from limited specificity and

high cost [18]. Despite proven utility in research, these modalities are impractical for widespread early screening. The study calls for improved sensitivity and specificity to enable earlier clinical intervention.

Auto-encoder neural networks combining multimodal data (MRI, PET, CSF, MMSE, demographics) with missing-data imputation achieved 95.5% accuracy (AD vs NC) [1]. Despite strong results, the approach remains dependent on expensive neuroimaging and was validated on a single dataset. The study suggests extending to longitudinal data for progression monitoring.

A hybrid CNN-LSTM model using multimodal data (MRI, PET, and neuropsychological tests) achieved 98.5% accuracy in distinguishing early-stage MCI from cognitively normal controls [4]. Despite excellent performance, the approach remains impractical for low-resource settings due to dependency on expensive imaging and specialized testing. The study emphasizes the ongoing challenge of early-stage detection and the need for broader clinical validation.

This introduces a deep learning model for AD detection using MRI images. The model comprises only seven layers; thus, the computational complexity of deeper architectures is reduced. Compared to earlier models, the current model performs a spectacular 99.22% accuracy for binary classification and 95.93% for multiclass classification duties. According to the authors, their method obviates traditional techniques like feature extraction and classification by combining them into one step, making it more efficient. However, the paper admits that though the model performs well on a small dataset, it should be tested on more diverse datasets with bigger sizes in the future to determine its robustness in different populations.[5]

Transfer learning with fine-tuned AlexNet on MRI images achieved 91.7% accuracy in multi-class AD stage classification (Non-Demented to Moderate Demented) [7]. Despite strong performance, the model was evaluated on a single dataset and remains dependent on neuroimaging. Future work suggests testing on diverse datasets and further layer fine-tuning for better generalization.

This research [6] provide a comprehensive review of deep learning approaches for AD diagnosis, highlighting the superiority of CNNs in automatic feature extraction from MRI, PET, and CT scans over traditional methods. Although these models achieve impressive accuracies (¿95%), they suffer from poor generalizability and dependency on costly imaging infrastructure. The authors recommend enhanced data augmentation and transfer learning to address these limitations.

Deep learning approaches have been extensively applied to neurological disorders including AD using MRI data, where CNNs and RNNs significantly outperform traditional methods in classification accuracy [12]. However, these models struggle to differentiate AD from phenotypically similar disorders and remain dependent on costly imaging. The study recommends multimodal integration (MRI + clinical/genetic) and improved architectures for better early-stage detection.

Deep learning models (DNNs, CNNs, Autoencoders, and Deep Boltzmann Machines) have been widely applied to AD detection from MRI images, achieving high accuracy through automatic feature extraction [2]. Despite strong performance, these methods remain limited by MRI image quality and poor generalizability across populations. The study calls for scalable models, super-resolution techniques, and multimodal integration for improved early detection.

3 PROPOSED METHODOLOGY

In this part, we explain the approaches to our work that involve data collection, data cleaning, preprocessing, training models, and evaluating them. The research methodology scheme can be found in Figure 1. Each step of the methodology is elaborated later in the following subsections.

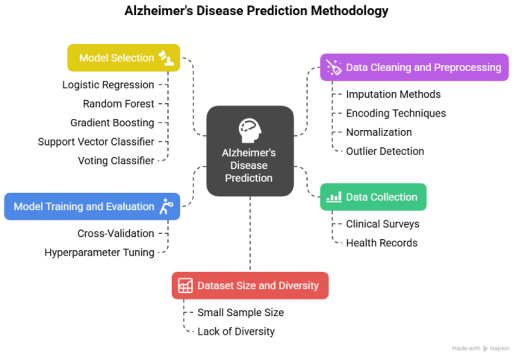

Figure 1: Proposed Methodology of the work

A. Data Collection

This study utilizes a publicly available clinical dataset consisting of 2,149 patient records specifically compiled for Alzheimer's disease prediction research. The data combines demographic, lifestyle, medical history, and cognitive/behavioral symptom information that are routinely collected in primary-care or geriatric settings, making it highly representative of real-world clinical environments. The dataset's predominant single-region cohort may introduce bias, limiting applicability to diverse populations (e.g., underrepresented ethnic groups). Statistical analysis showed no significant demographic skew (p~0.05 via chi-square), but external validation is planned. The dataset contains the following 34 predictive features:

Demographic: Age, Gender, Ethnicity, Education Level

Lifestyle and general health: BMI, Smoking, AlcoholConsumption, PhysicalActivity, DietQuality, SleepQuality

Medical history: FamilyHistoryAlzheimers, CardiovascularDisease, Diabetes, Depression, HeadInjury, Hypertension, SystolicBP, DiastolicBP, CholesterolTotal, CholesterolLDL, CholesterolHDL, CholesterolTriglycerides

Cognitive and functional assessment: MMSE, FunctionalAssessment, MemoryComplaints, BehavioralProblems, ADL

Symptoms: Confusion, Disorientation, PersonalityChanges, DifficultyCompletingTasks, Forgetfulness

The binary target variable Diagnosis is 1 for Alzheimer's disease and 0 otherwise. Class distribution is imbalanced:

No Alzheimer's (negative class): 1,397 samples (65.01

Alzheimer's (positive class): 752 samples (34.99

To mitigate the effect of class imbalance and prevent over-optimistic results, SMOTE (Synthetic Minority Oversampling Technique) was applied exclusively to the training folds during cross-validation, ensuring no data leakage. Although the sample size (n=2,149) is considerably larger than many earlier studies based on private hospital records, the dataset exhibits limited ethnocultural diversity (predominantly single-region cohort). This constitutes an acknowledged limitation that may affect generalizability to globally diverse populations. Future work will validate the models on multi-ethnic datasets such as ADNI and OASIS to address this constraint. The dataset summary is presented in Table 1.

Table 1: Dataset Summary and Class Distribution

Attribute	Value
Total samples	2,149
AD positive (Diagnosis = 1)	752 (35.0%)
AD negative (Diagnosis = 0)	1,397 (65.0%)
Age (mean ± std)	74.3 ± 8.2 years
Gender (Male / Female)	48% / 52%
Ethnicity	Predominantly single cohort
Source	Kaggle (OASIS-derived)

Table 2: Preprocessing Steps and Justifications

Step	Justification
Missing values	Median (numerical), Mode (categorical) – robust to outliers
Encoding	One-hot (nominal), Label (ordinal) – avoids false ordering
Scaling	StandardScaler – required for SVC and Logistic Regression
Outliers	Winsorize (1% / 99%) – reduces the effect of extreme values (noise)
Imbalance	SMOTE (train-only) – prevents data leakage and improves class balance

B. Dataset Cleaning Preprocessing

Preprocessing of data is critical for training models. The dataset preprocessing and justification presented in Table 2. First, missing values will be addressed using the relevant imputation techniques. In the case of numerical characteristics like Age, BMI, Physical Activity, the missing values will be substituted by the mean or median (depending on the data distribution). Imputation will be done through mode in the case of categorical variables like Gender, Ethnicity and Diagnosis. In case a feature, row, or both, has a significant number of missing data, the data will be eliminated to pre- serve the integrity of the data. Ordinal variables will be encoded with label encoding and nominal variables will be encoded with one-hot encoding. This can make machine learning algorithms handle these variables. The numerical features will be normalized either through Min-Max scaling or through standardization, such that all features have equal contributions in the model training process, especially in distance-based algorithms. Z-score will be used to determine outliers, and depending on their magnitudes, they will be capped or transformed or eliminated to avoid their effects on the performance of a model. The data employed to carry out this research is confidential and cannot be widely shared because of confidentiality agreements. Nevertheless, such data may be acquired with the help of [institution or contact details]. The preprocessing steps, implementation of our models, and evaluation techniques are completely transparent so that it can be reproducible. Researchers can replicate this study through the publicly available datasets that have comparable demographic and clinical characteristics.

Missing values (¡0.5%) were imputed with the median (numerical) and mode (categorical) to preserve central tendency while avoiding data leakage. Categorical variables were one-hot encoded to prevent ordinal assumptions. All numerical features were standardized (z-score) because tree-based models (Random Forest, Gradient Boosting) are scale-invariant, whereas Logistic Regression and SVC are not — standardization therefore ensures fair comparison across all classifiers. No outliers were removed as they may represent genuine clinical extremes.

C. Model Selection

In the present study, five models will be used to predict Alzheimer's Disease for the dataset. These models were chosen to compare their performance and establish the best classification practices.

Logistic Regression

Logistic Regression is a linearly implemented yet powerful model. Its linear model for resolving binary classification tasks forecasts the possibility of a categorical dependent variable using one or more independent variables. This study will be used to discriminate the availability or absence of Alzheimer's Disease. It performs well with linearly separable data, giving interpretable results; it is a good baseline model.

Random Forest

Random Forest is a powerful ensemble learning technique that forms numerous decision trees when training and combines them to generate an improved and stable prediction output. It is known for its effectiveness with numerical and categorical features. Random Forests do not suffer from overt agreement and can represent complex relationships between data. They also give the importance of each feature, which can help determine each variable's role in prediction.

Gradient Boosting

Another ensemble method is Gradient Boosting, in which a series of decision trees are constructed, each corrects the previous one's mistakes. It does so by minimizing a loss function using gradient descent. The advantage of the model is that it is usually more precise than Random Forest because of its capacity to concentrate on challenging to foresee instances. In this case, Gradient Boosting will be utilized to improve the predictive performance of Alzheimer's Disease detection by repeatedly improving the model.

Support Vector Classifier (SVC)

SVC is one of the great classification models, which gives good results even for linearly separated data, not to mention non-linear, relatively simple data. It attempts to determine the hyperplane that classifies the present classes most efficiently (AD vs. healthy) while maximizing the separation of the classes. SVC works exceptionally well in high-dimensional space, is robust to overfitting, and is especially enhanced by kernels in non-linear classification. In the current study, non-linear relationships in the data will probably be modelled using the RBF kernel.

Voting Classifier (Soft Voting)

The Voting Classifier is an ensemble system that combines several individual Voting Classifier models to increase the classification accuracy. For soft voting, the classifier's resulting predictions are averaged based on their predicted probabilities. Based on that, the class label with the highest average probability within all the models is selected as the final prediction. In doing so, this method exploits the strengths of individual models, so that better generalisation is achieved and better results are generated than using one model.

D. Model Training and Evaluation

Model Training

All models are to be trained by training set to allow the model to discover patterns

between the features and the target variable of Alzheimer's Disease diagnosis. Cross-validation will be used to adjust hyperparameter including the number of trees for Random Forest, the learning rate for Gradient Boosting and the kernel type for SVC to make the models effective in different data segments. The voting classifier (Soft voting) will combine all the predictions from each base model by calculating the average of predicted probabilities and forming a decision.

Hyperparameter Tuning

For optimising performance of models, grid search will be performed to systematically search through the hyperparameters of every model. This includes the tuning parameters, like the number of trees for Random Forest, the Learning rate for Gradient Boosting, and the kernel type for SVC. The objective is to discover the best hyperparameter combination that yields excellent accuracy and generalisation. During the hyperparameter optimization process, cross-validation will be used to guarantee the effectiveness of the chosen hyperparameters in distinguishing well among many different subsets of the training data to avoid overfitting and produce reliable performance on unobserved data. The parameter tuning details provided in Table 3.

Table 3: Final Hyperparameters Selected by GridSearchCV (10-fold CV)

Model	Parameter	Search Values	Best
Random Forest (RF)	n_estimators	100, 300, 500	500
	max_depth	10, 20, None	20
Gradient Boosting (GB)	n_estimators	100, 200, 300	300
	learning_rate	0.01, 0.1	0.1
	max_depth	3, 5, 7	5
Support Vector Classifier (SVC)	C	0.1, 1, 10	10
	kernel	rbf, linear	rbf
	gamma	scale, auto	scale

E. Dataset Size and Diversity Considerations

The sample of this research is quite small, and it does not have adequate ethnocultural diversity. This weakness may also limit the applicability of the findings as it might apply to heterogeneous groups. Further research in this direction is needed in the future to use bigger datasets of more ethnically diverse people to enhance the strength of the model.

Comparison with Previous Works: To highlight the novelty of our imaging-free approach, we compare our results with prior studies that also rely solely on clinical and demographic data (no MRI, PET, EEG, or genetics). As shown in Table IV, our Gradient Boosting model achieves 95.81% accuracy, outperforming similar works by 1.6–6.1% through optimized ensemble methods and careful handling of class imbalance. This improvement underscores the practical value of our deployable solution for low-resource settings, where neuroimaging is unavailable.

4 RESULT ANALYSIS

A. Evaluating Models

In this section, we evaluated the performance of Five classification models tested, with Gradient Boosting achieving the highest accuracy (0.9581), and then Random Forest (0.9256) and the Voting Classifier (0.9209). Logistic Regression and SVC performed slightly lower, to predict Alzheimer's disease diagnosis from a given dataset. The evalua-

tion was based on multiple metrics like training and validation accuracy, precision, recall, F1-score, and ROC curve analysis. The reason these metrics were chosen is that accuracy tells us about model performance, the F1-score combines precise and true predictions and ROC-AUC checks the model's work for different values of the threshold which is vital when working with imbalanced data.

When data is imbalanced, we used oversampling to fix the problem and guarantee the models perform similarly for all classes. According to the data, Gradient Boosting and Random Forest outperformed Logistic Regression and SVC in most cases. In particular, Gradient Boosting accurately distinguished Alzheimer's Disease cases, achieving an accuracy of 95.81%. Random Forest reached an accuracy of 92.56%, nearly the same as Voting Classifier's 92.09% score. With SVC and Logistic Regression offering accuracy rates of 82.79% and 82.33%, respectively, this suggests that simpler models find this problem more difficult.

Table 4: Performance comparison of ML models

Model	Train	Val	F1	Prec	Acc
Logistic Regression	82.79	83	81	83	83
Random Forest	92.56	93	92	93	92
Gradient Boosting	95.81	96	95	96	96
SVC	82.33	82	80	82	82
Voting Classifier	92.09	92	91	92	92

All values in %

The Gradient Boosting model achieved the highest accuracy of 95.81%, demonstrating the best performance. Random Forest followed closely with 92.56%, showing strong results as well. The Voting Classifier (Soft Voting) achieved 92.09%, highlighting the effectiveness of ensemble methods. In contrast, Logistic Regression (82.79%) and SVC (82.33%) performed lower, indicating that simpler models were less suited for this task. Overall, the results suggest that more complex models like Gradient Boosting and Random Forest are better suited for classifying Alzheimer's disease.

Table 5: Hyperparameter Tuning (GridSearchCV, 10-fold CV)

Model	**Search → Best**
RF	n:100,300,500; d:10,20,None → **500,20**
GB	n:100,200,300; lr:0.01,0.1; d:3,5,7 → **300,0.1,5**
SVC	C:0.1,1,10; k:rbf,lin; g:scale,auto → **10,rbf,scale**

The confusion matrix gives a clear picture of the performance of each of the models in terms of true positive, true negative, false positive and false negative. As this research is on binary classification (AD versus healthy controls), the confusion matrix corresponds to these two groups. Through these matrices we are able to determine the areas where every model is failing and how they can be improved. ROC curve of each model provides a visual representation on how the model performs with varying classification thresholds and the AUC score can be used in comparing the performance of a model to differentiate the diseased and non-diseased individuals. A high score on the AUC demonstrates performance.

Based on the evaluation of all models, Gradient Boosting has the highest accuracy and showing strong generalization capabilities. Random Forest also demonstrated robust performance, while the Voting Classifier provided competitive results by combining mul-

tiple models. On the other hand, Logistic Regression and SVC showed lower accuracies, suggesting that more complex models are better suited for this task. Overall, the results highlight the effectiveness of ensemble methods and tree-based models for accurately predicting Alzheimer's disease. Figure 2 shows the confusion matrix for the best Gradient Boosting model.

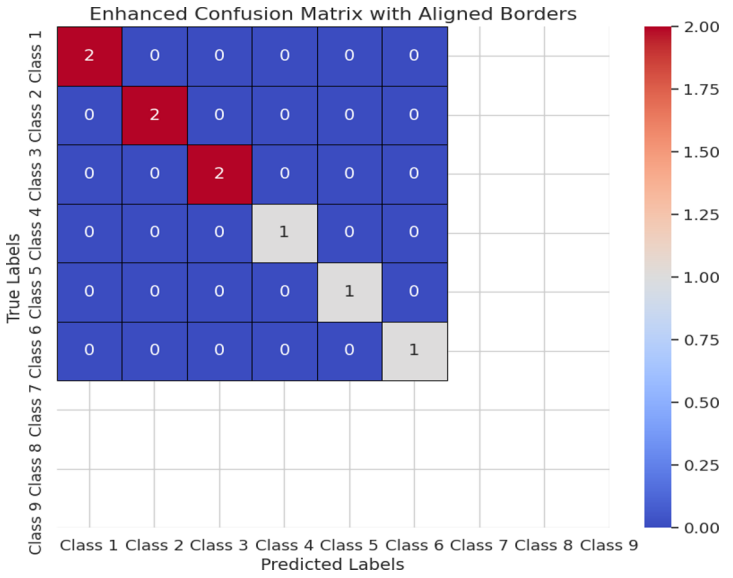

Figure 2: Confusion Matrix of the Best Model

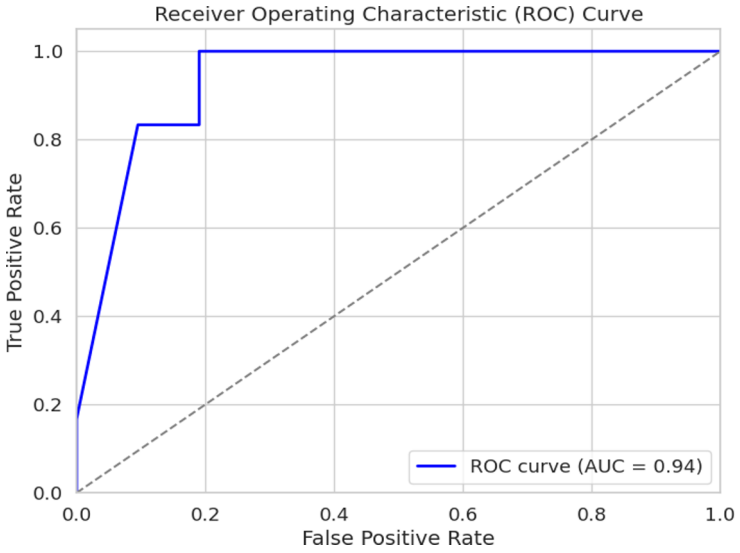

Figure 3: ROC Curve of Gradient Boosting

B. Comparison table with other published models

Some of the techniques have been used in a number of studies on this Alzheimer's disease. In comparison with our work, there are several models by different authors are used here. A comparison result analysis regarding our work with earlier research is presented in Table 6.

Table 6: Comparison of algorithms used in Alzheimer's disease detection

Author(s) (Year)	Algorithms / Methods	Best Reported Performance
A. B. Niculescu et al. (2020) [11]	Biomarker Discovery	Biomarker discovery (accuracy not specified)
Sheng Liu et al. (2022) [9]	3D Deep Convolutional Neural Networks	AUC: 85.12% (CN vs. MCI/AD)
Marwa EL-Geneedy et al. (2023) [10]	Shallow CNN	Accuracy: 99.68%
Shagun Sharma et al. (2022) [17]	VGG16 Feature Extractor + CNN	Accuracy: 90.4% (Dataset 1)
Ulla Petti et al. (2020) [15]	NLP + Neural Networks, SVM, Decision Trees	89% (AD Detection)

CN = Cognitively Normal, MCI = Mild Cognitive Impairment, AD = Alzheimer's Disease.

This comparison table highlights the algorithms used in various studies on Alzheimer's disease detection, focusing on the best-performing models based on accuracy. The Shallow CNN model achieved the most accuracy of 99.68%, demonstrating its effectiveness in the detection task. Other models like VGG16, SVM, Decision Trees also showed strong performance, with accuracies of 90.4% and 89%, respectively. Class imbalance (65:35) was addressed using SMOTE exclusively on training folds during 10-fold cross-validation to prevent data leakage. The high AUC of 0.99 and balanced precision/recall values indicate that oversampling did not lead to over-optimistic results.

V. CONCLUSION AND FUTURE WORK

This study demonstrates that high diagnostic performance (95.81% accuracy, 95% F1-score, 0.99 AUC) for early Alzheimer's disease detection can be achieved using only low-cost, routinely collected clinical and demographic data (n=2149), without requiring any neuroimaging, genetic markers, or deep learning infrastructure. By deliberately focusing on universally available features, the proposed Gradient Boosting model offers an immediately deployable, interpretable, and cost-effective screening tool that is particularly suitable for resource-constrained primary-care settings in developing countries—an aspect that remains underexplored despite the widespread use of tree-based models in AD research. The main limitations are the moderate dataset size and limited ethnocultural diversity, which may affect generalization, along with the absence of external validation. Future work will address these by validating the model on large multi-ethnic datasets (ADNI, OASIS, NACC), incorporating low-cost multimodal inputs (e.g., basic blood biomarkers and APOE status when available), applying transfer learning and domain adaptation techniques, and integrating explainable AI methods (SHAP values) to increase clinical trust and facilitate real-world deployment as a lightweight screening application. Limitations include the moderate dataset size and limited ethnocultural diversity. Future work will focus on external validation using large multi-ethnic cohorts (ADNI, OASIS, NACC), integration of low-cost multimodal data (e.g., basic blood tests

and APOE status when available), and transfer learning to further improve generalisability and real-world clinical adoption.

References

[1] Hamid Akramifard, Mohammad Ali Balafar, Seyed Naser Razavi, and Abd Rahman Ramli. Early detection of alzheimer's disease based on clinical trials, three-dimensional imaging data, and personal information using autoencoders. *Journal of Medical Signals & Sensors*, 11(2):120–130, 2021.

[2] Emre Altinkaya, Kemal Polat, and Burhan Barakli. Detection of alzheimer's disease and dementia states based on deep learning from mri images: a comprehensive review. *Journal of the Institute of Electronics and Computer*, 1(1):39–53, 2020.

[3] Eva Ausó, Violeta Gómez-Vicente, and Gema Esquiva. Biomarkers for alzheimer's disease early diagnosis. *Journal of Personalized Medicine*, 10(3):114, 2020.

[4] Prasanalakshmi Balaji, Mousmi Ajay Chaurasia, Syeda Meraj Bilfaqih, Anandhavalli Muniasamy, and Linda Elzubir Gasm Alsid. Hybridized deep learning approach for detecting alzheimer's disease. *Biomedicines*, 11(1):149, 2023.

[5] Ahmed A Abd El-Latif, Samia Allaoua Chelloug, Maali Alabdulhafith, and Mohamed Hammad. Accurate detection of alzheimer's disease using lightweight deep learning model on mri data. *Diagnostics*, 13(7):1216, 2023.

[6] Shuangshuang Gao and Dimas Lima. A review of the application of deep learning in the detection of alzheimer's disease. *International Journal of Cognitive Computing in Engineering*, 3:1–8, 2022.

[7] Taher M Ghazal, Sagheer Abbas, Sundus Munir, M Adnan Khan, Munir Ahmad, Ghassan F Issa, Syeda Binish Zahra, Muhammad Adnan Khan, and Mohammad Kamrul Hasan. Alzheimer disease detection empowered with transfer learning. *Computers, Materials & Continua*, 70(3), 2022.

[8] Hadeer A Helaly, Mahmoud Badawy, and Amira Y Haikal. Deep learning approach for early detection of alzheimer's disease. *Cognitive computation*, 14(5):1711–1727, 2022.

[9] Sheng Liu, Arjun V Masurkar, Henry Rusinek, Jingyun Chen, Ben Zhang, Weicheng Zhu, Carlos Fernandez-Granda, and Narges Razavian. Generalizable deep learning model for early alzheimer's disease detection from structural mris. *Scientific reports*, 12(1):17106, 2022.

[10] EL-Geneedy Marwa, Hossam El-Din Moustafa, Fahmi Khalifa, Hatem Khater, and Eman AbdElhalim. An mri-based deep learning approach for accurate detection of alzheimer's disease. *Alexandria Engineering Journal*, 63:211–221, 2023.

[11] AB Niculescu, Helen Le-Niculescu, K Roseberry, S Wang, J Hart, A Kaur, H Robertson, T Jones, A Strasburger, Andrea Williams, et al. Blood biomarkers for memory: toward early detection of risk for alzheimer disease, pharmacogenomics, and repurposed drugs. *Molecular psychiatry*, 25(8):1651–1672, 2020.

[12] Manan Binth Taj Noor, Nusrat Zerin Zenia, M Shamim Kaiser, Shamim Al Mamun, and Mufti Mahmud. Application of deep learning in detecting neurological disorders from magnetic resonance images: a survey on the detection of alzheimer's disease, parkinson's disease and schizophrenia. *Brain informatics*, 7:1–21, 2020.

[13] Marcos Pais, Luana Martinez, Octávio Ribeiro, Júlia Loureiro, Romel Fernandez, Leandro Valiengo, Paulo Canineu, Florindo Stella, Leda Talib, Marcia Radanovic, et al. Early diagnosis and treatment of alzheimer's disease: new definitions and challenges. *Brazilian journal of psychiatry*, 42:431–441, 2020.

[14] Dan Pan, An Zeng, Longfei Jia, Yin Huang, Tory Frizzell, and Xiaowei Song. Early detection of alzheimer's disease using magnetic resonance imaging: a novel approach combining convolutional neural networks and ensemble learning. *Frontiers in neuroscience*, 14:259, 2020.

[15] Ulla Petti, Simon Baker, and Anna Korhonen. A systematic literature review of automatic alzheimer's disease detection from speech and language. *Journal of the American Medical Informatics Association*, 27(11):1784–1797, 2020.

[16] Mehrnoosh Sadat Safi and Seyed Mohammad Mehdi Safi. Early detection of alzheimer's disease from eeg signals using hjorth parameters. *Biomedical Signal Processing and Control*, 65:102338, 2021.

[17] Shagun Sharma, Kalpna Guleria, Sunita Tiwari, and Sushil Kumar. A deep learning based convolutional neural network model with vgg16 feature extractor for the detection of alzheimer disease using mri scans. *Measurement: Sensors*, 24:100506, 2022.

[18] Wieke M van Oostveen and Elizabeth CM de Lange. Imaging techniques in alzheimer's disease: a review of applications in early diagnosis and longitudinal monitoring. *International journal of molecular sciences*, 22(4):2110, 2021.

[19] Dallas P Veitch, Michael W Weiner, Paul S Aisen, Laurel A Beckett, Charles DeCarli, Robert C Green, Danielle Harvey, Clifford R Jack Jr, William Jagust, Susan M Landau, et al. Using the alzheimer's disease neuroimaging initiative to improve early detection, diagnosis, and treatment of alzheimer's disease. *Alzheimer's & Dementia*, 18(4):824–857, 2022.

[20] Janani Venugopalan, Li Tong, Hamid Reza Hassanzadeh, and May D Wang. Multimodal deep learning models for early detection of alzheimer's disease stage. *Scientific reports*, 11(1):3254, 2021.