# Advancing Early Alzheimer's Disease Detection in Underdeveloped Areas with Fair Explainable AI Methods

**Quoc-Toan Nguyen**
University of Technology Sydney
Sydney, NSW, Australia
quoctoan.nguyen@student.uts.edu.au

## Abstract

Artificial intelligence (AI)-based telemedicine systems for early Alzheimer's detection using low-cost modalities are vital for rural or underdeveloped areas where travelling distance and high-cost devices like MRI are drawbacks. These systems require eXplainable AI (XAI) for reliable outcomes and intuitive explanations. Current XAI evaluations lack input from medical professionals and overlook stakeholder diversity, leading to potential biases. This project aims to develop a cost-effective AI telemedicine system, enhance early AD detection in underdeveloped areas, reduce healthcare disparities, and assess XAI methods with quality and fairness to mitigate biases for high-quality and fair explained outcomes.

## Introduction

Alzheimer's Disease (AD), the most common form of dementia, accounts for 60-80% of cases. It progresses through two main stages: Mild Cognitive Impairment (MCI) and dementia AD. This neurodegenerative disorder leads to cognitive decline, such as memory loss and impaired reasoning, as well as Behavioral and Psychological Symptoms of Dementia (BPSD), including delusions, aggression, wandering, and disinhibition (Hampel et al. 2011). Early detection is paramount to help medical professionals intervene to slow disease progression, which is vital for individuals' Quality of Life (QoL), as no curative treatments currently exist (van der Flier et al. 2023).

Utilising magnetic resonance imaging (MRI) with artificial intelligence (AI) models has proven effective in distinguishing normal controls (NC), MCI, and AD. However, the high-cost devices hinder the access of this technique to people in underdeveloped areas (Murali et al. 2023), and the distance to access to the medical premises poses a disadvantage as well for them. Hence, AI-based telemedicine comes into play, which uses digital communication and technologies to provide healthcare services remotely at home and rural clinics (Sekhon et al. 2021).

On top of that, in the healthcare sector, eXplainable AI (XAI) plays a vital role in supporting both medical professionals and individuals in understanding the why of models' outcomes, leading to a reliable system. Nevertheless, the current evaluations of XAI methods in AD detection have vital limitations (Viswan et al. 2024). Firstly, XAI researchers often rely solely on their judgement rather than consulting medical experts, leading to explanations that lack specialist input. Secondly, interpretation varies with expertise, causing uncertainty and disagreement when explanations conflict with expectations. Perceptions can vary among different demographic groups, and explainability is a multifaceted characteristic (Nauta et al. 2023). However, the evaluation is human-grounded but typically overlooks the demographic characteristics of the stakeholder team, which can lead to a biased evaluation of XAI methods, leading to inaccurate and biased decisions and worsening healthcare disparities. In summary, the abovementioned problems motivated the development of this project, tackling a paramount need for underdeveloped areas.



## Research Questions

Building on the identified gaps in the previous section, this project seeks to address the following research questions (RQ):

- **RQ1**: How can we develop a versatile AI-based telemedicine system for early AD detection supported underdeveloped areas that utilises multiple modalities, with unimodal and multimodal model(s) applying both home-based and clinic-based settings?
- **RQ2**: How can we develop XAI methods for both unimodal and multimodal AI models so that the quality and fairness of explanation outputs are assessed effectively?

Figure 1 in Appendix A illustrates the proposed system with models' outputs as NC, MCI and AD. Five key models are expected to be developed: four unimodal (model 1,2,3,4) and one multimodal (model 5). The unimodal models ensure the system is applicable even in underdeveloped areas with varying access to modalities and devices. Besides, the bottom of the figure highlights the developed **Fair-AD framework**, which is crucial for addressing RQ2, considering **quality** and **fairness** in to-be-developed XAI methods' evaluation.

## Methodologies

This section explains the to-be-developed AI models, XAI methods and developed Fair-AD framework for XAI evaluation as depicted in Figure 1.

### Primer of To-be-developed AI Models

- **Model 1**: Using speech with Boston Diagnostic Aphasia Examination (BDAE), as it is commonly seen in AD (Loukas and Dimitris 2023).
- **Model 2**: Drawing and handwriting will be conducted, given the strong correlation between motor control and AD (Azzali et al. 2024).
- **Model 3**: Cognitive tests with questions are used because accessing devices of the required quality for speech, drawing, and handwriting assessments can also be inaccessible.
- **Model 4**: Using Electroencephalography (EEG), preliminary results have been achieved with a proposed model named EEG-SSM (Tran et al. 2024), which demonstrated an accuracy of 91%. The next step for EEG involves developing a model using the microstate approach, which has shown high effectiveness (Lassi et al. 2023). This modality can only be utilised on premises since setting up an EEG scalp at home is limited.
- **Model 5**: Combining all modalities used in models 1-4 for a multimodal model, representing a comprehensive and ideal scenario when all modalities are available.

### XAI Methods and Fair-AD Framework

Firstly, regarding the to-be-developed XAI methods (Viswan et al. 2024), natural language and visualisation will be employed. Combining these approaches can make explanations more intuitive and understandable for humans while ensuring efficiency for local samples.

Notably, to assess the quality and fairness of to-be-developed XAI methods, the Fair-AD framework has been developed as crucial part to the project's deployment. It is an evaluation technique for all the to-be-developed XAI methods used in all AI models. This framework includes a Likert questionnaire with twenty-four questions, leveraging twelve explanation quality properties across three dimensions presented by Nauta *et al.* (Nauta et al. 2023): **Content** (correctness, completeness, consistency, continuity, contrastivity, covariate complexity), **Presentation** (compactness, composition, confidence), and **User** (context, coherence, controllability).

Each property above has two questions; each question scores from 0 to 5, resulting in an overall score ($Overall_{sc}$) ranging from 0 to 120. $Overall_{sc}$ can be used to rank developed methods as well:

- **Poor (0-20)**: Highly inadequate and fails to meet basic requirements.
- **Subpar (21-40)**: Below standard and lacks important elements
- **Average (41-60)**: Fine but needs improvement.
- **Good (61-80)**: Solid and meets most expectations.
- **Very Good (81-100)**: Highly satisfactory and covers most aspects well.
- **Excellent (101-120)**: Outstanding and comprehensively addresses all criteria

For the fairness assessment of outputs from the XAI method, two fairness metrics are tailored for XAI evaluation: Disparate Impact ($DI$) and Demographic Parity ($DP$), which are widely used in fairness evaluations of algorithms (Pessach and Shmueli 2022). The fairness results from the developed Fair-AD are calculated by comparing $Overall_{sc}$ pairwise across different demographic groups of stakeholders based on sex, race, marital status, etc. An acceptable difference threshold for a fair XAI method between groups is set at a 20% or below difference based on the 80% rule; otherwise, it is unfair and needs improvements. The XAI method should meet both quality and fairness criteria for real-life deployment. The case study below demonstrates it.

**Case Study**: Assuming that we have already developed Model 1 and an XAI method for this model $\rightarrow$ We have outputs $x \in X$, which are explained outputs $\rightarrow$ The evaluation of the developed XAI method will be conducted by AI developing and medical teams from different demographic groups, including sex ($g_s \in G_s$) and race ($g_r \in G_r$) $\rightarrow$ Each $g_s$ and $g_r$ will be compared pairwise by their $Overall_{sc}$ in their $G_s$ and $G_r$, to have $DI$ or $DP$ $\rightarrow$ If the threshold is satisfied and the average $Overall_{sc}$ among all $g_s$ and $g_r$ are in the range of **Good** $\rightarrow$ This developed XAI method is a fair good one and can be considered to be deployed, otherwise not.

## Conclusion

This project addresses a paramount need in the field of telemedicine for early AD detection by developing high-performing AI models combined with robust XAI methods. Notably, the developed Fair-AD framework for evaluating XAI methods will play a key role in the project, providing comprehensive evaluations of the quality and fairness of XAI methods with users from different demographic groups to identify and mitigate potential biases for medical decision-making, ensuring equitable and accurate outcomes for all users.

## Acknowledgements

This work was funded by the Australian Research Council (ARC) via discovery grants DP210101093 and DP220100803, as well as the Australian National Health and Medical Research Council (NHMRC) Ideas Grant APP2021183. Further support was provided by the UTS Human-Centric AI Centre, sponsored by GrapheneX (2023-2031), Australia Defence Innovation Hub under Contract No. P18-650825, and Australian Cooperative Research Centres Projects (CRC-P) Round 11 CRCPXI000007.

## References

Azzali, I.; Cilia, N. D.; De Stefano, C.; Fontanella, F.; Giacobini, M.; and Vanneschi, L. 2024. Automatic feature extraction with Vectorial Genetic Programming for Alzheimer's Disease prediction through handwriting analysis. *Swarm and Evolutionary Computation*, 87: 101571.

Hampel, H.; Prvulovic, D.; Teipel, S.; Jessen, F.; Luckhaus, C.; Frölich, L.; Riepe, M. W.; Dodel, R.; Leyhe, T.; Bertram,

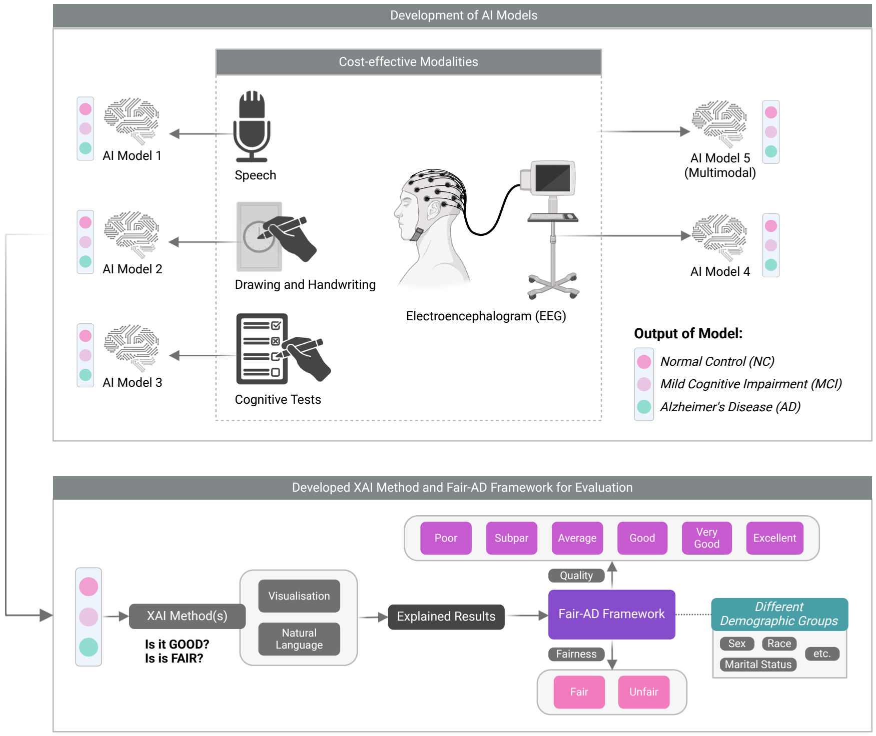


Figure 1: Comprehensive Overview of Proposed System for Early AD Detection using Cost-effective Modalities with Models, XAI Methods and Fair-AD Framework for XAI Evaluation.

L.; et al. 2011. The future of Alzheimer's disease: the next 10 years. *Progress in neurobiology*, 95(4): 718–728.

Lassi, M.; Fabbiani, C.; Mazzeo, S.; Burali, R.; Vergani, A. A.; Giacomucci, G.; Moschini, V.; Morinelli, C.; Emiliani, F.; Scarpino, M.; et al. 2023. Degradation of EEG microstates patterns in subjective cognitive decline and mild cognitive impairment: Early biomarkers along the Alzheimer's Disease continuum? *NeuroImage: Clinical*, 38: 103407.

Loukas, I.; and Dimitris, A. 2023. Context-aware attention layers coupled with optimal transport domain adaptation and multimodal fusion methods for recognizing dementia from spontaneous speech. *Knowledge-Based Systems*, 277: 110834.

Murali, S.; Ding, H.; Adedeji, F.; Qin, C.; Obungoloch, J.; Asllani, I.; Anazodo, U.; Ntusi, N. A.; Mammen, R.; Niendorf, T.; et al. 2023. Bringing MRI to low-and middle-income countries: directions, challenges and potential solutions. *NMR in Biomedicine*, e4992.

Nauta, M.; Trienes, J.; Pathak, S.; Nguyen, E.; Peters, M.; Schmitt, Y.; Schlötterer, J.; van Keulen, M.; and Seifert, C. 2023. From anecdotal evidence to quantitative evaluation methods: A systematic review on evaluating explainable ai. *ACM Computing Surveys*, 55(13s): 1–42.

Pessach, D.; and Shmueli, E. 2022. A review on fairness in machine learning. *ACM Computing Surveys (CSUR)*, 55(3): 1–44.

Sekhon, H.; Sekhon, K.; Launay, C.; Afililo, M.; Innocente, N.; Vahia, I.; Rej, S.; and Beauchet, O. 2021. Telemedicine and the rural dementia population: A systematic review. *Maturitas*, 143: 105–114.

Tran, X.-T.; Le, L.; Nguyen, Q. T.; Do, T.; and Lin, C.-T. 2024. EEG-SSM: Leveraging State-Space Model for Dementia Detection. *arXiv preprint arXiv:2407.17801*.

van der Flier, W. M.; de Vugt, M. E.; Smets, E. M.; Blom, M.; and Teunissen, C. E. 2023. Towards a future where Alzheimer's disease pathology is stopped before the onset of dementia. *Nature aging*, 3(5): 494–505.

Viswan, V.; Shaffi, N.; Mahmud, M.; Subramanian, K.; and Hajamohideen, F. 2024. Explainable artificial intelligence in Alzheimer's disease classification: A systematic review. *Cognitive Computation*, 16(1): 1–44.