Advances in artificial intelligence for diagnosing Alzheimer's disease through speech

Mishal Abid, MBBS[a], Maham Asif, MD[b], Zoya Khemane, MBBS[a], Afia Jawaid, MBBS[a], Aimen Waqar Khan, MBBS[c], Hufsa Naveed, MBBS[d], Tooba Naveed, MBBS[d], Asma Ahmed Farah, MBBS[e,*], Mohammad Arham Siddiq, MD[b]

Alzheimer's disease (AD) is a brain disorder that has been ranked as the seventh leading cause of death in the United States mainly affecting older adults. It is estimated that approximately more than 6 million Americans have dementia caused by Alzheimer's[1]. In the early stages of AD, individuals may experience subtle changes in language abilities, such as difficulty finding words or repeating phrases. As the severity advances, these language impairments become more pronounced, affecting both expressive and receptive language skills. Furthermore, deficits in cognition can contribute to the language processing difficulties observed in AD patients. Executive dysfunction, for example impairs the ability to plan and organize thoughts, leading to disorganized speech and difficulty staying on topic during conversations. Semantic memory deficits result in difficulties understanding and producing meaningful language, while attention deficits contribute to distractibility and poor concentration during language tasks.

These language and cognitive impairments not only impact communication and social interactions but also pose significant challenges for accurate diagnosis and monitoring of disease progression. As advancements are made in artificial intelligence (AI)—particularly natural language processing (NLP), there have been more ways being explored to diagnose Alzheimer's dementia at an earlier stage. For diagnoses through speech, data from Alzheimer's Dementia Recognition through Spontaneous Speech (ADReSS) is usually used and for knowledge of the severity of AD, the Mini-Mental State Examination (MMSE) cognitive test results are taken into account as well. The speech recordings used are of patients with an AD diagnosis and healthy controls describing a Cookie Theft picture from Boston Diagnostic Aphasia Examination[2].

While traditional diagnostic approaches, such as cognitive testing and neuroimaging, have demonstrated utility, they may not capture subtle changes in linguistic patterns that can manifest in the early stages of the disease, highlighting the need for more sensitive and objective measures. Our manuscript stresses on the importance of leveraging advancements in AI-based systems to develop more accurate and efficient tools for diagnosing AD at earlier stages.

Recently, GPT-3—a language model produced by OpenAI was experimented with to discover if it could predict Alzheimer's dementia[3]. Spontaneous speech was converted to a GPT-3 based embedding and was compared with previously used methodologies including acoustic feature-based approach and fine-tuned model. The data were confined to the dataset of ADReSS challenge and for AD and non-AD classification, the embeddings were rendered to a machine-learning model such as support vector classifier (SVC) or random forest (RF) for a 10-fold cross-validation (CV). The results of the study concluded that GPT-3 embeddings proved to be more advanced for predicting AD dementia than the conventional methods with an accuracy of 0.8028 by SVC. Even after these results, there is a concern of its public usage as it has only been experimented with a small ADReSS data hence more enhancements need to be made for more reliability.

A study was conducted to diagnose dementia using speech, comparing the performance of two AI-based systems. The first system utilized the pre-trained data2vec-audio-base-960h model, while the second system used the pre-trained wav2vec2-base-960h model and an acoustic feature approach as a benchmark for comparison[4]. For improved testing, an internal dataset (ADReSS) along with an external dataset (DementiaBank Pitt) were used in the 10-fold CV. The assessment after going through a thorough technical procedure reflected that the internal dataset achieved an area under the receiver operating characteristic (ROC) curve (AUC) of 0.846, whereas the external dataset gained an AUC of 0.835 complying that data2vec can be used for Alzheimer's diagnosis. The study also showed that data2vec worked better than wav2vec2. Although multiple datasets were used, they are still considered small and require more in-depth assessment before it is considered an official diagnosing option.

Different studies explore ways of Alzheimer's dementia through different angles. Fine-tuning text converted from speech, through Bidirectional Encoder Representations from Transformers (BERT) and Enhanced Representation through Knowledge Integration (ERNIE) language models was experimented where atypical pauses between sentences were observed. Speech recordings and transcripts were used from ADReSS challenge and were annotated using CHAT coding system—an integrated tool of Child Language Data Exchange System (CHILDES) project[5]. The transcripts were force-aligned, fine-

[a]Department of Medicine, Dow University of Health Sciences, Departments of [b]Medicine, [c]Internal Medicine, Jinnah Sindh Medical University, [d]Department of Medicine, Ziauddin Medical College, Karachi, Pakistan and [e]Department of Medicine, East Africa University, Boosaaso, Somalia

*Corresponding author. Address: 66R3 + CC4, East Africa University, Boosaaso, Somalia, Tel.: +252 907 814 006. E-mail: asmaaaxmad16@gmail.com (A.A. Farah).

tuned along with encoding pauses and classified for authentic results. The leave-one-out (LOO) cross-validation showed 89.6% accuracy with ERNIE fine-tuning including the pauses that were encoded for AD classification. It was also noticed that 'um' was comparatively repeated less often in AD, indicating that it may have a lexical status[6]. Due to low research on the importance of pauses for AD classification, it is still a question whether it should be paid attention to or not.

The on-going research on the effect on spontaneous speech due to AD also caters to proving its authenticity through challenging it via multi-lingual machine learning. Evidence was collected from participants fluent in English and French language. They were extracted from different collections of datasets (English, 2020 ADReSS INTERSPEECH challenge and French, EIT-Digital ELEMENT project). Speech recordings of participants describing the Cookie theft picture were used as speech samples. Different language features that is semantics, syntax, task-specific and paralinguistic, were checked and measured using a program created in python programming language. The results contribute to the conclusion that language impairment in AD can be used to diagnose the disease since it can be measured through advanced NLP techniques combined with machine learning (ML) without undermining the performance[7].

With growing usage of AI in all fields around the globe, it is essential that it is used as a good beneficial resource in the medical field as well. Along with diagnosing Alzheimer's dementia at an early stage, it can also prove to play a vital role in diagnosing or curing other diseases as well. It is not only cost-effective but sooner with upgrades, different applications might come into action, which could allow self-diagnose and reduce patient's mental stress and expenses. There is a need for transparency and accountability in the development and deployment of AI-based systems, as well as regulation and oversight to ensure their responsible use. The development of AI-based systems for diagnosing and managing AD has the potential to revolutionize the field of dementia care. However, it is important to note that more research is needed to ensure the reliability and public usage of these AI-based systems. As these systems rely on large datasets for training, it is crucial that the datasets used are diverse, representative, and of high quality. Furthermore, ethical considerations such as data privacy and informed consent need to be considered when implementing these systems. Patient's confidentiality matters and it shall be ensured that no software or any AI by-product abuses patient's information since using unencrypted voice recordings of patients for dementia diagnosis can be considered a breach to privacy. There is a need for transparency and accountability in the development and deployment of AI-based systems, as well as regulation and oversight to ensure their responsible use.

With continued advancements and collaboration between researchers, clinicians, and policymakers, these systems have the potential to transform the landscape of dementia care and improve the lives of millions of individuals and their families affected by AD.

Ethical approval

Not applicable as this research is a short communication so an ethics approval statement wasn't required.

Consent

Not applicable as this research is a short communication so consent wasn't required.

Sources of funding

The authors received no extramural funding for the study.

Author contribution

The conceptualization was done by M.A. and M.A.S. The literature and drafting of the manuscript were conducted by M.A., Z.K., A.J., H.N., T.N., A.A.F. and A.W.K. The editing and supervision were performed by M.A.S. and M.A. All authors have read and agreed to the final version of the manuscript.

Conflicts of interest disclosure

The authors declare no potential conflicts of interest concerning the research, authorship, and/or publication of this article.

Research registration unique identifying number (UIN)

Not applicable.

Guarantor

All the authors take full responsibility.

Data availability statement

Not applicable.

Provenance and peer review

Not commissioned, externally peer-reviewed.

References

[1] Alzheimer's Disease Fact Sheet | National Institute on Aging. 2023. Accessed April 5, 2023. https://www.nia.nih.gov/health/alzheimers-and-dementia/alzheimers-disease-fact-sheet

[2] Borod JC, Goodglass H, Kaplan E. Normative data on the boston diagnostic aphasia examination, parietal lobe battery, and the boston naming Test. J Clin Neuropsychol 1980;2:209–15.

[3] Agbavor F, Liang H. Predicting dementia from spontaneous speech using large language models. PLOS Digital Health 2022;1:e0000168.

[4] Agbavor F, Liang H. Artificial intelligence-enabled end-to-end detection and assessment of Alzheimer's disease using voice. Brain Sci 2023;13:28.

[5] Pye C, MacWhinney B. The CHILDES Project: Tools for Analyzing Talk. Language 1994;70:156.

[6] Yuan J, Cai X, Bian Y, Ye Z, Church K. Pauses for detection of Alzheimer's disease. Front Computer Sci 2021;2:624488.

[7] Lindsay H, Tröger J, König A. Language impairment in Alzheimer's disease—robust and explainable evidence for AD-related deterioration of spontaneous speech through multilingual machine learning. Front Aging Neurosci 2021;13:642033.