Speech-Based Detection and Severity Assessment of Alzheimer's Disease

Speech-Based Alzheimer's Detection

Chidubem Valentine Ezikeoha

Computer Science, Howard University, valentineezikeoha@gmail.com

Saurav K. Aryal

Computer Science, Howard University, saurav.aryal@howard.edu

Howard Prioleau

Computer Science, Howard University, howard.prioleau@bison.howard.edu

Early detection in Alzheimer's remains challenging, as current diagnostic methods rely on clinical expertise, and cognitive testing, which can be costly and inaccessible. Because cognitive decline is reflected in speech through changes in vocabulary, readability, pauses, articulation, and prosody, spoken language offers a scalable, low-cost, and non-invasive signal for screening and monitoring. Prior work demonstrated good performance in predicting Mini-Mental State Examination (MMSE) scores using large acoustic-linguistic feature sets, with optimized LightGBM and ensemble models significantly reducing RMSE and achieving strong Alzheimer's detection accuracy. In this work, we extend beyond text-focused modeling toward a multilingual, audio-driven model for dementia detection and severity assessment. We train Audio Spectrogram Transformer (AST) models across ADReSS-style datasets to establish robust audio-only baselines, apply multilingual automatic speech recognition to obtain transcripts, and develop multimodal fusion models that integrate acoustic embeddings with linguistic features. This work lays the foundation for scalable and multilingual speech-based tools that can support early dementia detection. By using both audio and linguistic signals, it moves towards a more accessible approach to identifying cognitive decline.

CCS CONCEPTS • Machine Learning • Speech Recognition • Health Informatics

Additional Keywords and Phrases: Dementia analysis, Multimodal learning, Multilingual Learning, computational paralinguistics

REFERENCES

[1] Saturnino Luz, Fasih Haider, Sofia de la Fuente, Davida Fromm, and Brian MacWhinney. 2020. Alzheimer's Dementia Recognition through Spontaneous Speech: The ADReSS Challenge.https://talkbank.org/dementia/ADReSS-2020/index.html.

[2] [1] Saturnino Luz, Fasih Haider, Sofia de la Fuente, Davida Fromm, and Brian MacWhinney. 2021. Detecting cognitive decline using speech only: The ADReSSo Challenge. In Proceedings of the Annual Conference of the International Speech Communication Association (Interspeech 2021). ISCA. https://talkbank.org/dementia/ADReSSo-2021/index.html

[3] Luz S, Garcia SdLF, Haider F, Fromm D, MacWhinney B, Lanzi, A, Chang, YN, Chou CJ and Liu YC. Connected Speech-Based

Cognitive Assessment in Chinese and English . Proceedings of Interspeech, pp 947--951, Kos, Greece, 2024. https://taukadial-luzs-69e3bf4b9878b99a6f03aea43776344580b77b9fe54725f4.gitlab.io/

[4] [2] Saturnino Luz, Fasih Haider, Davida Fromm, Ioulietta Lazarou, Ioannis Kompatsiaris, and Brian MacWhinney. 2022. An overview of the ADReSS-M signal processing grand challenge on multilingual Alzheimer's dementia recognition through spontaneous speech. In Proceedings of the IEEE Signal Processing Grand Challenge 2022. https://talkbank.org/dementia/ADReSS-M/index.html