

The joint effects of audio and linguistic markers for early identification of mild cognitive impairment

Fengyi Tang¹ | Jun Chen² | Hiroko H Dodge³ | Jiayu Zhou¹ | I-CONNECT

¹ Michigan State University, East Lansing, MI, USA

² University of Michigan, Ann Arbor, MI, USA

³ NIA-Layton Aging & Alzheimer's Disease Center, Oregon Health & Science University, Portland, OR, USA

Correspondence

Jiayu Zhou, Michigan State University, East Lansing, MI, USA.

Email: jiayuz@msu.edu

Abstract

Background: In recent years, speech-related behavioral markers such as spoken language and lexical preferences have been studied in the early detection of mild cognitive impairment (MCI). While the combination of linguistic and acoustic signals has been shown to be effective in detecting MCI, they have generally been restricted to structured conversations in which the interviewee responds to fixed prompts. In this study, we propose a stratification scheme that proves to be effective in extracting quality linguistic and acoustic MCI markers in semi-structured more natural conversational settings.

Method: We obtained transcripts and audio recordings from an ongoing single blind randomized controlled clinical trial NCT02871921. Briefly this clinical trial aims to improve psychological well-being and cognitive functions of the non-demented (normal or MCI) older adults aged 75 and older by enhancing their social interactions. The experimental group receives 30 minutes of semi-structured conversations (video chats) with interviewers four times per week, for up to one year. Recruitment started in July 2018 and we used the data from the first 39 participants enrolled in the experimental group with audio and transcribed data as of March 2020 in the current analysis. Our stratification scheme eliminates drastic train-test distribution mismatch, allowing for apples-to-apples comparison across participant conversations, despite drastic variations in word choices, sentence structure and length of conversations – obstacles that typically present challenges to the semi-structured conversational setting for automatic acoustic marker extraction.

Results: Figure 1 gives an overview of our MCI prediction pipeline. The combination of linguistic and audio features achieve a mean AUC of 82.7, significantly ($p < 0.01$) outperforming linguistic-only (AUC = 74.9) or acoustic-only (AUC = 65.0) detections on hold-out data. Our subtopic stratification method identifies several conserved clusters (Figure 2) in interview structure, allowing more balanced train-test split in terms of linguistic properties, despite the semi-structured nature of conversations.

Conclusion: A combination of linguistic and audio features improves the ability to detect MCI beyond that achieved by each feature alone in semi-structured conversation. Our method provides a potentially valuable approach for using speech for early detection of MCI and clinical trials enrichment.

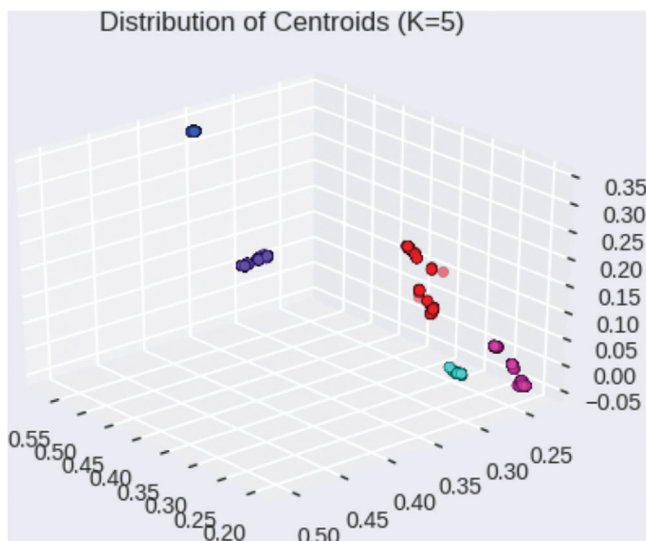


FIGURE 1

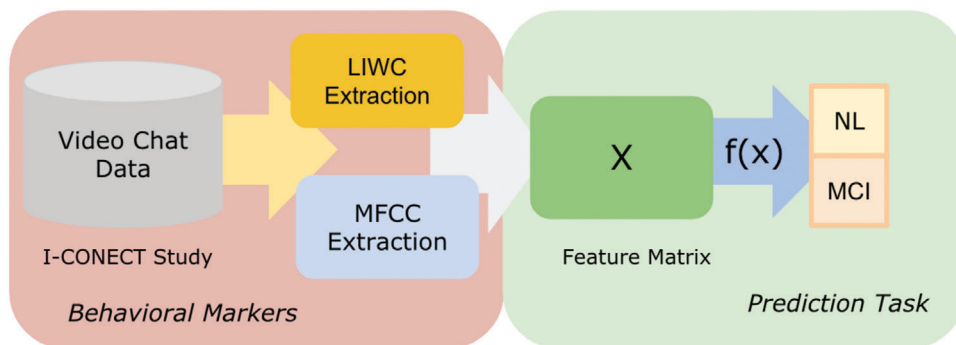


FIGURE 2