An avatar-based system for identifying individuals likely to develop dementia

Bahman Mirheidari[1], Daniel Blackburn[3], Kirsty Harkness[4], Traci Walker[5], Annalena Venneri[3,7], Markus Reuber[6], and Heidi Christensen[1,2]

[1]Department of Computer Science, [2]Centre for Assistive Technology and Connected Healthcare (CATCH), [3]Sheffield Institute for Translational Neuroscience (SITraN), [4]Department of Neurology, [5]Department of Human Communication Science, [6]Academic Neurology Unit, University of Sheffield, Sheffield, UK; [7]IRCCS Fondazione Ospedale San Camillo, Venice, Italy

Introduction

Dementia

- General term for a broad category of ***decline in mental ability***
- Starts with subtle ***word finding difficulties***, a decline in thinking or memorising; aggravates over time and ultimately leads to ***loss of communication***.
- Challenging to diagnose due to the lack of ***reliable bio-markers***, ***overlapping symptoms*** with normal ageing and ***low accuracy*** of existing cognitive (*"pen-and-paper"*) screening tools
- Effect on language ([1, 2]) includes:***loss of vocabulary***, impoverished and ***simplified syntax/semantics***, and overuse of semantically ***empty words***
- ***Conversation analysis*** (CA) [3] (an approach to study social interaction/communication ability) has been used for people with dementia (e.g. [4, 5]), but it requires audio/video recording, transcribing, and a qualitative analysis (carried out by an ***expert***); ***time-consuming*** and relatively ***expensive*** and not applicable for ***large scale use***

Research questions

- Is it ***feasible*** to develop an ***automatic tool*** to help doctors in diagnosing dementia? What kind of ***speech, text and ML technologies*** and tools can be used for designing such a system?
- Task is to classify between two types of memory diseases with very similar symptoms: ***neudegenerative dementia (ND)*** and ***functional memory disorder (FMD)***
- To what extent it is ***feasible*** to use ***an avatar*** front-end to elicit conversational diagnostic features?

Dementia detection system

- Diarisation (***who talks when***) (SHoUT toolkit)
- Automatic speech recognition (ASR) (Kaldi toolkit)
- Feature extraction (NLTK python + Praat toolkit)
- Machine learning classifier (SVM from Scikit-learn python)

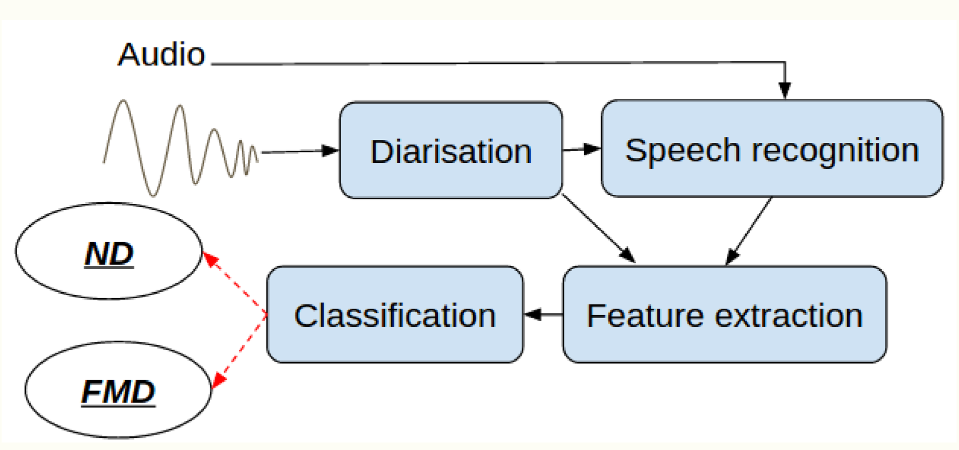

Figure 1: Block diagram of dementia detection system

Experiment

Data

- **A) Neurologist-patient conversations**:Audio files and transcriptions of interviews of 15 FMD and 15 ND participant-doctor consultations
- **B) Avatar-patient conversations**:Audio files and transcriptions of conversations of 6 FMD and 6 ND participant with the Avatar.

Features

- **Conversation Analysis inspired[3]:** 20 features, e.g. patient answered me for who's most concerned question, average number of empty words (CA is an approach to study social interaction/ communication ability of people which has been used recently for people with dementia (e.g. [4, 5])
- **Acoustic:** 12 features, e.g. silence, intonation, pitch, H1-H2
- **Lexical (Part of Speech):** 12 features, e.g. number of verbs, nouns adverbs, etc

Avatar system

- **Avatar head animation:** Botlibre (https://www.botlibre.com).
- **Avatar voice:** Pre-recorded human voice

Figure 2: Prototype avatar

Results

A.Speech recognition

Table 1: Speech recognition results.

System	Train	Test	WER
Baseline_HUM	HUM	HUM	55.7%
Baseline_AVA	AVA	AVA	77.0%
Cross domain	HUM	AVA	65.0%
MAP adaptation	Map on HUM	AVA	58.7%
Combining data	HUM+AVA	AVA	**46.2%**

B.Diarisation

DER: Diarisation Error Rate, a common metric to measure the performance of a diarisation tool , including the missing speaker error: E_{MISS}, false alarm:E_{FA}, and speaker error:E_{SPKR}. W-DER: Word diarisation error, extending the diarisation error to the words recognised by the ASR.

Table 2: DER and W-DER

Data	E_{MISS}	E_{FA}	E_{SPKR}	DER	W − DER
HUM_dia	2.7%	14.9%	12.8%	30.4%	**5.7%**
AVA_dia	11.6%	6.9%	11.1%	**29.6%**	16.8%

C.Classification and feature selection

Table 3: Classification accuracy;'_man':using gold-standard transcript instead of ASR-produced transcripts;'CA': CA-style features;'AC':acoustic features;'LX':lexical features;'T10':top 10 informative features.

Train/Test	CA	AC	LX	ALL	T10
A	**96.7%**	83.3%	66.7%	76.7%	**100%**
B	**76.7%**	60.0%	50.0%	76.7%	**90.0%**
C	58.3%	66.7%	**83.3%**	66.7%	75.0%
D	72.7%	63.6%	63.6%	**81.8%**	72.7%
E	63.6%	54.5%	63.6%	**90.9%**	72.7%

A: HUM_man/HUM_man
B:HUM/HUM
C:AVA_man+HUM_man/AVA_man
D:AVA_man+HUM_man/AVA
E:AVA+HUM_man/AVA

D.Differences between the two conversations

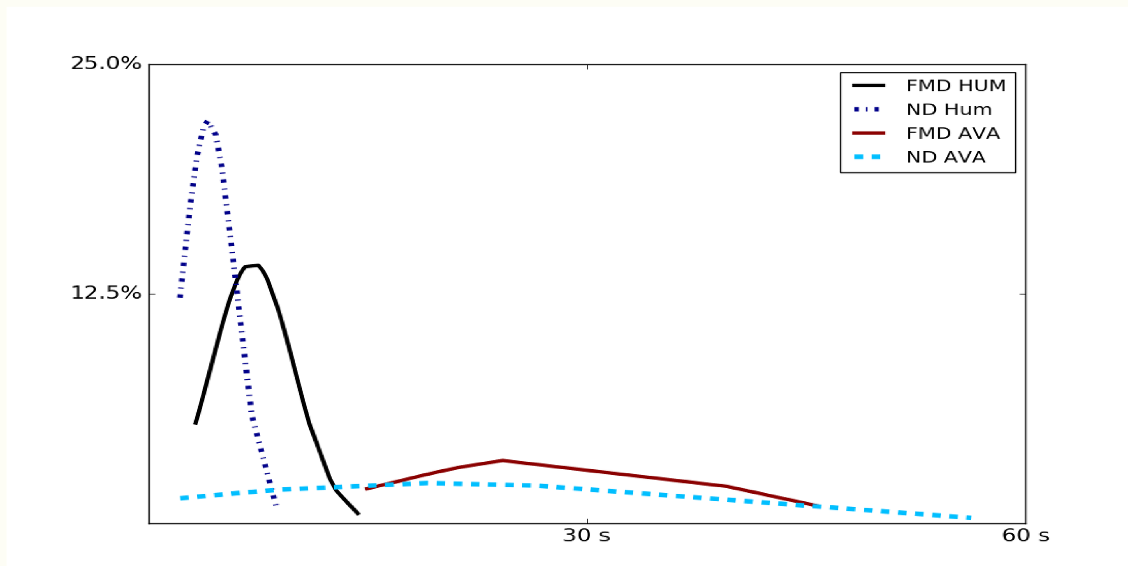

Figure 3: Histogram of the average turn length.

Conclusions and further work

Challenges

- ***Spontaneous speech*** resulting in ASR with high WER
- ***Background noise, mic far from patient***
- Challenging ***diarisation task***, high DER
- Large number of ***overlapping segments***
- Lack of ***feedback from the Avatar***, resulted in long turn responses

Conclusions

- We have proposed a fully ***automatic*** system for detecting dementia
- ***Feasible*** of replacing the neurologists with the Avatar
- Low cost of the potential tool to stratify patients with memory complaints

Future work

- Expanding to include ***more feature set***
- ***Improving*** the ASR, diarisation and feature extraction
- ***Improving the Avatar*** to make it more responsive

References

[1] Bayles, K.A. Kaszniak, A.W.(1987).Communication and cognition in normal aging and dementia. Taylor Francis Ltd London.

[2] Tang-Wai, D.F. Graham, N.L.(2008). Assessment of Language Function in Dementia.Geriatrics and Aging,11, 103-110.

[3] Lerner, G.H.(2004).Conversation Analysis: studies from the first generation. Amsterdam John Benjamins Pub.

[4] Kindell, J., Sage, K., Keady, J. Wilkinson, R.(2013). Adapting to conversation with semantic dementia: Using enactment as a compensatory strategy in everyday social interaction.International Journal of Language and Communication Disorders,48, 497-507.

[5] Elsey, C., Drew, P., Jones, D., Blackburn, D., Wakefield, S., Harkness, K., Venneri, A. Reuber, M.(2015). Towards diagnostic conversational proles of patients presenting with dementia or functional memory disorders to memory clinics.Patient Education and Counseling,98, 1071-1077.