Tracking changes in daily routines of elderly users through acoustic sensing: An unsupervised learning approach

Pavlos Nicolaou and Christos Efstratiou
pn218@kent.ac.uk, c.efstratiou@kent.ac.uk
University of Kent, United Kingdom

1. Motivation

The current state-of-art in assisted living relies on a blend of wearable sensing technologies (wristbands, pendants) and location tracking technologies embedded in their living environment [1]. However, when considering people living with dementia, solutions that rely on the user to regularly wear, maintain and recharge their device have limited success in real-world deployments. Instead, solutions that rely on passive sensing through environmental sensors are considered a more promising approach.

In this work we investigate the use of acoustic sensing within the living environment of people living with dementia. The work is part of an Innovate UK funded project ADAPTIVE (https://www.miicare.co.uk/adaptive) investigating the use of acoustic sensing for elderly care. The project relies on the use of digital assistants designed to provide support to elderly users. These devices are capable of continuous acoustic sensing opening the opportunity for capturing rich acoustic datasets for passive monitoring of human activities at home. The project is currently in the process of deploying acoustic sensing technology in the living environments of 60 elderly participants (30 people living with dementia).

With the numbers of people living with dementia expected to reach more than 8% of the UK population by 2040 [2], there is an increasing need to provide technology that can observe their daily lives and the progress of their condition while reducing the need for unnecessary involvement of carers.

2. Approach

Developing systems for activity recognition requires significant effort to generate accurately labeled datasets. For system that target people living with dementia, such tasks become extremely difficult and in many cases impossible to achieve.

In this work we explore the feasibility of developing an acoustic sensing system that can track the changes of daily routines of elderly users, through a fully unsupervised learning approach, without the need for accurately curated datasets. Our approach is motivated by the observation that from a healthcare perspective, one of the main requirement is to be able to observe whether the elderly user tends to follow their typical daily routine, and notify carers when there are significant diversion from their routine.

Our work-in-process approach for developing such system is to rely on transfer learning, utilising the pretrained model VGGish [3] as a way of mapping sounds into a multi-dimensional hyperspace. VGGish is a pretrained Convolutional Neural Network from Google. The model can generate a 128-dimensional embedding vector for each second of sound, maintaining close similarity between embeddings of sounds of the same type (e.g. conversations, vacuum cleaner, etc), and increased dissimilarity between sounds of different types. The continuous stream of sound captured during the daily lives of people at home can be translated into a time series of 128-dimensional vectors that represent the changes of activities that are happening throughout the day. In order to allow us to track changes of high-level activities the one sec vectors are smoothed and regularised to help us capture changes of activities that may involve longer time windows. The smoothing approach includes the aggregation of multiple embeddings over a wider sliding window (5 mins - 50% overlap), where multiple embedding vectors are mapped into the centroid of the given embeddings. The transformed time series represents smooth transisions of activities over time, as detected through the sound source.

The main objective of our approach is to identify similarities (or diversions) between daily patterns. The problem is similar to that of detecting similar/dissimilar body gestures through wearable technologies, using 3-dimensional input to represent movement within 3-D space. Borrowing from techniques applied in gesture recognition, we explore the use of multidimensional dynamic time warping (DTW) as a technique that can help us identify similar activity patterns, and identify diversions from the expected patterns.

We currently explore the feasibility of this approach using labelled datasets collected though healthy volunteers using an acoustic sensing mobile application. The model will eventually be evaluated through real-world deployments that are part of the ADAPTIVE project.

References

[1] P. Rashidi and A. Mihailidis, "A survey on ambient-assisted living tools for older adults," *IEEE journal of biomedical and health informatics*, vol. 17, no. 3, pp. 579–590, 2012.

[2] R. Wittenberg, B. Hu, L. Barraza-Araiza, and A. Rehill, "Projections of older people living with dementia and costs of dementia care in the united kingdom, 2019-2040," *Care Policy and Evaluation Centre, London School of Economics and Political Science*, 2019.

[3] S. Hershey, S. Chaudhuri, D. P. Ellis, J. F. Gemmeke, A. Jansen, R. C. Moore, M. Plakal, D. Platt, R. A. Saurous, B. Seybold *et al.*, "Cnn architectures for large-scale audio classification," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2017, pp. 131–135.

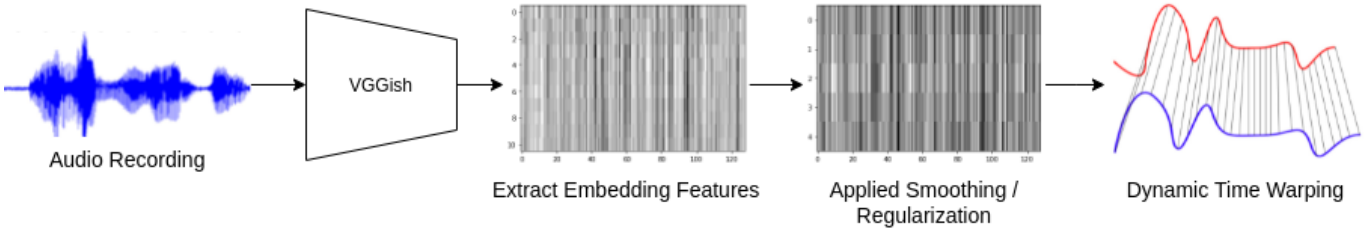

Figure 1. Unsupervised acoustic tracking of daily activities