

Prototypical System to Detect Anxiety Manifestations by Acoustic Patterns in Patients with Dementia

Netzahualcoyotl Hernandez^{1,*}, Matias Garcia-Constantino¹, Jessica Beltran², Pascal Hecker^{3,4}, Jesus Favela⁵, Joseph Rafferty¹, Ian Cleland¹, Hussein Lopez⁵, Bert Arnrich³, and Ian McChesney¹

¹Ulster University, Belfast, Northern Ireland, United Kingdom.

²CONACYT-IPN (Instituto Politecnico Nacional), Tijuana, Mexico.

³Hasso Plattner Institute, University of Potsdam, Potsdam, Germany.

⁴audEERING GmbH, Gilching, Germany.

⁵CONACYT-CICESE (Centro de Investigacion Cientifica y de Educacion Superior de Ensenada), Ensenada, Mexico.

Abstract

INTRODUCTION: Dementia is a syndrome characterised by a decline in memory, language, and problem-solving that affects the ability of patients to perform everyday activities. Patients with dementia tend to experience episodes of anxiety and remain for extended periods, which affects their quality of life.

OBJECTIVES: To design AnxiDetector, a system capable of detecting patterns of sounds associated before and during the manifestation of anxiety in patients with dementia.

METHODS: We conducted a non-participatory observation of 70 diagnosed patients in-situ, and conducted semi-structured interviews with four caregivers at a residential centre. Using the findings from our observation and caregiver interviews, we developed the AnxiDetector prototype and tested this in an experimental setting where we defined nine classes of audio to represent two groups of sounds: (i) Disturbance which includes audio files that characterise sounds that trigger anxiety in patients with dementia, and (ii) Expression which includes audio files that characterise sounds expressed by the patients during episodes of anxiety. We conducted two experimental classifications of sounds using (i) a Neural Network model trained and (ii) a Support Vector Machine model. The first evaluation consists of a binary discriminating between the two groups of sounds; the second evaluation discriminates the nine classes of audio. The audio resources were retrieved from publicly available datasets.

RESULTS: The qualitative results present the views of the caregivers on the adoption of AnxiDetector. The quantitative results from our binary discrimination show a classification accuracy of 98.1% and 99.2% for the Deep Neural Network and Support Vector Machine models, respectively. When classifying the nine classes of sound, our model shows a classification accuracy of 92.2%. Whereas, the Support Vector Machine model yielded an overall classification accuracy of 93.0%.

CONCLUSION: In this paper, we presented the outcomes from an observational study *in-situ* at a residential care centre, qualitative findings from interviews with caregivers, the design of AnxiDetector, and preliminary qualitative results of a methodology devised to detect relevant acoustic events associated with anxiety in patients with dementia. We conclude by signalling future plans to conduct *in-situ* validation of the effectiveness of AnxiDetector for anxiety detection.

Keywords: smart microphones, anxiety, dementia, sound recognition.

Received on 03 August 2019, accepted on 21 August 2019, published on 23 August 2019

Copyright © 2019 Netzahualcoyotl Hernandez *et al.*, licensed to EAI. This is an open access article distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/3.0/>), which permits unlimited use, distribution and reproduction in any medium so long as the original work is properly cited.

doi: 10.4108/eai.10-2-2020.163097

*Corresponding author. Email: hernandez_cruz-n@ulster.ac.uk

1. Introduction

It is estimated that 50 million people across the world live with dementia [1]. Predictions suggest that the incidence of dementia will double within 20 years, reaching 75 million diagnosed cases in 2030 and 131.5 million in 2050 [1]. Dementia's early signs may consist of a set of symptoms including memory loss, difficulty with planning or solving problems, misplacing things, having problems understanding visual information, speaking problems, and physical dysfunction, which result in limitations carrying out daily living activities. Symptoms of dementia also include stress, progressive anxiety, wandering and aggression, which can lead to situations of self-harm or danger [2]. Studies have shown a positive correlation between stress and anxiety [3]; these factors have been shown to recur people with mental health disorders. In this context, stress acts as a physical defence mechanism in threatening situations, and it is gradually reduced after the situation is over. Whereas, anxiety, is a sustained mental disorder that can be triggered by stress and remains for an extended period [2]. Anxiety is defined as a feeling of unease, such as worry or fear that can be rated from moderate to severe stage.

In patients with dementia (PwD), anxiety is normally associated with a lack of engagement with health care and a low quality of life [2]. The relationship between anxiety and dementia has been already studied in the literature [4]–[7]. Currently, quantification of anxiety leverages self-reporting as a standard clinical approach. These tools are adequate to measure anxiety at specific points in time, providing a general assessment. However, they do not provide an opportunity to prevent nor detect anxiety episodes on a constant, daily, basis [8]–[10]. This is due to this approach requiring a degree of cognitive effort which tends to interrupt people's routine activities. In this context, our research interest lies in the automatization of detecting anxiety episodes, so caregivers can provide the appropriate assistance by being remotely informed more accurately.

This work builds upon our previous study [11]. In this current manuscript, we present the design of AnxiDetector as a piece of Ubiquitous Health (UbiHealth) technology to assist caregivers by identifying acoustic patterns derived from episodes of anxiety in PwD. We describe a non-participatory observational study [12] and qualitative results in which we introduce a prototype of AnxiDetector. In collaboration with the “Digital Health – Connected Healthcare” research group at Hasso Plattner Institut, we also extend our previous research work by introducing the methodology, evaluation, and quantitative results for detecting acoustic patterns derived from anxiety episodes.

While it is acknowledged that there are studies in the literature which are designed to address the automatic detection of anxiety, the novelty of our study and related approach is comprised mainly by the following three aspects: (i) unobtrusiveness, (ii) automated, and (iii) uses robust technology. AnxiDetector is unobtrusive because it is placed

on the environment where PwD live. There is no requirement for PwD to wear nor carry any device that could cause them distressed or uncomfortable. AnxiDetector is automated in the way in which it collects and analyses data, and in sending notifications to caregivers, requiring minimum technical configurations during its usage. Finally, AnxiDetector uses robust technology in the form of smart microphones enabled to compute the directionality of acoustic signals out of an array of microphones, as described in further sections.

In Section 2, we present the related work and limitations of current solutions. Section 3 introduces our research methodology and elaborates on scenarios that illustrate the application of our approach. Section 4 describes our AnxiDetector prototype. We then elaborate on its technical details in Section 5. In Section 6, we present the results from a preliminary evaluation. Section 7 introduces some of the challenges and limitations of our approach. Finally, we discuss some recommendations and conclude with future work and discussion in Sections 8 and 9, respectively.

2. Related work on anxiety detection

A review of studies concerned with the automatic detection of anxiety was performed. Some existing work is aligned to an analysis of facial cues from videos, such as eye and mouth activity [13], through the analysis of speech features, such as vocal pitch [14] and by a combination of sensing technology [15].

W. Simm *et al.* reported the “Clasp” system as a technology to detect anxiety [16]. This system consists of some anxiety coping devices, such as stress balls, that send usage data to a smartphone, enable support network communication, and a web-based interface. The web-based interface presents a timeline for community feedback [16].

A similar solution is Sensing Whether Affect Requires Mediation (SWARM). SWARM uses a scarf that incorporates conductive fabric circuitry to control heat, vibration, and audio actuators to help people with autism to relieve anxiety [17].

A wearable and mobile intervention solution, introduced by L. Cruz *et al.*, has the objective of reducing symptoms of panic disorder by guiding the PwD on performing breathing and relaxation exercises; in this case, anxiety was one of the three more severe symptoms identified [18].

The automated system for anxiety detection presented in a study conducted by H. Haritha *et al.* [19] analysed respiratory signals collected from an inductive band which was placed across the pectoral. Signals from such bands were analysed to detect breathe fluctuations. The output of this analysis was then used to inform of anxiety.

The approach introduced by N. Chaitanya *et al.* [20] is in the form of a massage headband based on ElectroEncephaloGraphy (EEG) to detect brainwaves within a subject. These brainwaves were processed to detect anxiety.

Zheng *et al.* [21] conducted a study to quantify anxiety based on features derived from EEG and PhotoPlethysmoGram (PPG) data collected from a wearable

headset and glasses. PPG data indicates blood volume changes within a subject. Data from these data sources were fused and processed to determine anxiety levels.

The applicability of leveraging electrocardiogram (ECG) features for detecting anxiety disorders using wearable devices was analysed in a recent review of state of the art [22]. Specifically, applicability across panic, post-traumatic stress, generalized anxiety, social, mixed, and obsessive-compulsive anxiety disorders was investigated.

Portable Autonomous Multisensory Intervention Device (PAMID) [23] is a proprietary device designed to monitor anxiety as well as other negative behavioural symptoms wirelessly. PAMID used automatic audio analysis to detect disruptive behaviours from PwD to prevent anxiety episodes that could escalate [23]. To support that approach, the audio was used to provide information about different levels of contexts, such as speech, activities [18], and environmental sound events [12].

The review concluded that, although much of the previous works reported in the literature has obtained satisfactory results; in many cases, the data collection method is intrusive and raises important questions regarding privacy.

It has been shown that wearable solutions are intrusive to observed users by requiring devices to be physically affixed to individuals. This physical attachment can cause issues with discomfort, unfamiliarity and notably, can cause distress within PwD. Wearable devices typically require maintenance such as charging, which must be managed. Additionally, they require a user to be remembered to wear these devices each day or employ caregivers to do this. A requirement for a PwD to maintain or remember to wear such devices cannot be assured due to the nature of the illness [24]. As such, the use of ambient sensings, such as audio is a more appropriate model when observing PwD.

The recent popularity in the use of smart microphones; which provide high-quality audio and include speech recognition algorithms, is very relevant as these are potential tools for detecting anxiety due to their ease of placement in the living environment and their intuitive interaction/operation. An important advantage of using microphones over other types of technologies is that they are an unobtrusive technology that allows capturing data from people without using wearable devices. Moreover, analysing the audio within the smart microphone addresses privacy concerns as it is not necessary to transmit recordings for further investigation but only to send notifications regarding detected anxiety episodes.

Given their potential to provide useful information and their ease of use, smart microphones can be utilised to non-intrusively monitor the wellbeing of PwD through the identification of acoustic events. Smart microphones can be installed in living spaces (e.g., bathrooms, living rooms) and start collecting data from the PwD without interrupting or interfering in their daily living activities but rather non-intrusively detecting relevant acoustic events. To explore these capabilities, we conducted a non-participatory study aiming to acquire a broader understanding of the sounds

occurring before and during the manifestation of an episode of anxiety. Respective outcomes provided us with design clues to develop a prototype of AnxiDetector. We then build a Neural Network model to detect acoustic patterns associated with the manifestation of anxiety in PwD.

3. Development of anxiety in patients with dementia in their daily living

Designing UbiHealth technology requires a deep understanding of the user's profile and the environmental conditions in which the technology will be deployed. Our research methodology addressed this by incorporating elements of qualitative data collection from PwD and caregivers and quantitative evaluation of machine learning models for recognising categories of sounds. In this section we describe a four-day study based on a non-participatory observation of PwD [12] and interviews with their caregivers, to better understand how anxiety is manifested within Activities of Daily Living (ADL).

This data collection informed an iterative and incremental approach to the evaluation of AnxiDetector, as described in sections 4-6, including quantitative evaluation of machine learning models for sound recognition and qualitative evaluation of prototype outcomes with caregivers.

3.1. Consultation with Domain Experts

In coordination with the AgeingLab Foundation (host entity of the European project: REMIND[†]), the observational study was conducted at the Residential Centre 'Ángeles Cobo López', located in the city of Alcaudete in Jaén, Spain. This is a residence that welcomes cognitively impaired patients who voluntary accept to be hosted by the care centre or have been clinically referred. Patients are hosted according to their level of dependency which is characterised by their impairment profile (e.g., dementia, Parkinson's, bipolar disease) and impairment severity (i.e., emerging, moderate, and severe). Some of the services that patients receive consist of physical treatment and psychological therapies. They are also supported with medication management and constant monitoring towards reaching the stage in which they are sufficiently healthy to temporarily leave the residence centre under the condition that they return to the facilities at night.

To conduct this study, the AgeingLab Foundation provided ethical approval under the auspices of Personal and Public Involvement, which includes the involvement of expert caregivers as specialist advisers in the evaluation. The non-participatory observation was conducted under the ethical and privacy countenance of the Residential Centre 'Ángeles Cobo López'. Note that the quotes provided in this report have been translated from Spanish by the researcher who is conducting the observational study *in-situ*.

[†] https://cordis.europa.eu/project/rcn/207045_en.html

3.2. Non-participatory observation procedure

The non-participatory observation aimed at identifying the characteristics of anxiety episodes while the PwD performs their daily living activities such as physical therapies, requesting assistance from the caregivers, and during socialisation. Hence, this study was conducted over three consecutive days (10 hours total) during different daily sessions: (i) morning (from the time the residents wake up until they have breakfast), (ii) afternoon (in which the residents frequently socialize, watch TV, and prepare to have lunch), and (iii) night (in which the residents receive dinner before resting).

Lunchtime was observed on two occasions. Firstly, 50 PwD with a low degree of impairment (who were located in the common lunchroom), and secondly, 20 PwD with a severe degree of impairment.

A psychomotor therapy session was observed during the morning session, in which 20 PwD were guided step by step by a physiotherapist through a pre-established set of activities. The activities had an occupational dimension, which focused on empowering the performance of activities of daily living, such as getting dressed, showering, and eating. The observation was complemented with a semi-structured interview addressed to two caregivers with eight- and two-years' experience, respectively, responsible for providing rehabilitation therapies. Questions included topics such as anxiety manifestation and strategies to address PwD experiencing anxiety.

Leisure time usually happens during the afternoon, when PwD have no scheduled activities; thus, they are allowed to independently watch TV, socialise with other residents, go for a walk within and outside the facilities, or attend the library or any other common area. Observation of 40 PwD in total was conducted (*i.e.*, 20 diagnosed with moderate dementia, and 20 diagnosed with severe levels of dementia). Caregivers answered questions about particular behaviours in a private office once observation tasks were concluded.

Due to privacy policies, access to personal rooms and night visiting was prohibited. However, we were informed about supervised activities such as waking up and showering PwD. This phase of the day is of particular interest due to a previously reported phenomenon called Sundown Syndrome [25], in which PwD tend to present confusion and behave irrationally after sunset.

4. Audio manifestation of anxiety in patients with dementia

The relationship between audio and anxiety in PwD was explored by conducting affinity diagramming to the notations collected during the non-participatory observation and interview sessions [26]. Below, we summarise the relevant findings in a narrative style.

4.1. Anxiety manifestation and canalization

In this subsection, we report some of the findings which have been anonymized due to the privacy agreement with the Residential Centre 'Ángeles Cobo López'.

Quote from a PwD with early impairment: [PwD-A] *"I don't like being in this room (i.e., leisure shared area) when PwD-B is here because he gets irritated if I want to change the TV's channel. He always turns it on the news."*

As reported above, it was observed that a PwD felt uncomfortable when another PwD with severe impairment behaved in a verbally aggressive manner and was loud when they were changing the TV channels in the common area.

Quote from a PwD with moderate impairment: [PwD-C] *"Oh, my legs, they hurt so much! Oh, my legs, they hurt so much!..."* [PwD-D] *"Oh, my God PwD-C, please be quiet!"*

As reported above, it was observed that a PwD with a higher impairment experienced anxiety more frequently. In the above quote, PwD-C loudly and repetitively complained about an experience she had the previous night. After a few minutes, PwD-D (who was sitting next to PwD-C) showed concern and started to threaten her to be quiet.

Quote from a Caregiver supervising a PwD with severe impairment: [Caregiver A] *"It is very common to see how PwD that are quietly lying in their couches suddenly start repeating what others do or want, like complaining about the noise/light or asking for water and support to go to the toilet, even when they don't need to use the toilet"*.

As reported above, it was observed that PwD that had fully developed a cognitive impairment presented anxiety most of the day. As an example, Caregiver A shared that due to a specific condition of PwD-E, she has been assigned a particular location during leisure time in which she is protected by a belt around her waist.

Additionally, it was observed, that during one of the non-participatory observational sessions PwD-E was overwhelmed and started to struggle with the chair and loudly complained about it, which then triggered a domino effect in which more than five PwD started to shout.

4.2. Sound patterns when anxiety is manifested

Through conducting this research, we found that during an episode of anxiety, PwD may have a particular behavioural habit. For instance, someone might produce a pattern noise with their hands, repeating specific sentences, whereas another will passively spend long-time daydreaming while sitting in a chair.

In two sessions, for example, PwD-F was repetitively hitting their chair with their hand for an extended period of

time. PwD-F was perceived as anxious; however, they stopped once a caregiver started conversing with them.

On another occasion, PwD-G was repeating a sentence (“My arm, my arm, my arm...”) many times before having dinner. PwD-G eventually stopped, had dinner and did not carry on with their previous repeating.

Findings of PwD with severe impairment are aligned with the production of unstructured sounds with their mouths when requiring caregivers’ attention. The intensity varied depending on what was understood as urgency; it was explained that this happens typically when assistance is required to use the toilet or when the PwD experiences pain or physical discomfort derived from sitting for an extended period of time.

[Caregiver A] “Normally, PwD that are unable to walk by themselves and can’t find the words to express themselves start hitting whatever object they have nearby to get our attention when they need to go to the toilet. If they don’t have any object, they will be beating the table, couch, or chair”.

Overall, it was observed a different pattern of expressed anxiety accordingly to the severity of mental impairment in patients. Those PwD with early impairment were able to wait for caregivers to bring their meal, while PwD with severe impairment were expecting immediate attention. In this regard, anxiety will progressively increase, and repetition of noise and sounds produced by the PwD will happen more frequently and with higher intensity. While this phenomenon varied across individuals, it was presented by most of the PwD observed.

4.3. Caregivers interviews outcomes

Arising from the interviews, we identified two key ways in which UbiHealth technology could have a practical benefit:

a) Anxiety detection in private room.

Description: Caregivers patrol residents’ areas on a daily basis; however, there is limited access to personal areas due to privacy policies. Patrolling is conducted along corridors from where caregivers can perceive loud noises.

Issue: There can be situations when a PwD will have anxiety episodes inside their private room that could not be spotted by caregivers due to the number of PwD to look after or quiet sounds a PwD might be producing.

[Caregiver A] “It is normal to hear PwD praying before sleep, laughing, or talking while sleeping. However, sometimes when we wake them up, we found they spent the night crying. They will have their eyes wet, so we work on cheer them up.”

b) Misbehaviour and violence detection.

Description: The residence is equipped with different shared areas, in which PwD conduct creative, or sports activities, as well as have a relaxing time watching TV. Loud noises and PwD misbehaviour are factors that increase the risk of aggressive discussion.

Issue: Arguing episodes might be overseen since common areas tend to be open to all PwD to socialize, thus individual assistance from caregivers might be limited due to the number of PwD.

[Caregiver B] “Normally we will struggle with PwD-A in the afternoons, when he wants to watch the TV, but PwD-B is already watching it. PwD-A is strong, so I am afraid he will throw a punch at any time; thus we tend to spot them and interfere as soon as we see they will be arguing”.

As shown by the non-participatory observation and the interviews, there are specific auditory manifestations that provide indications about anxiety episodes already happening, such as screams, or manifestations with the potential to evolve to anxiety episodes, such as growling words before crying. Although these manifestations can be specific to the PwD, its automatic detection can support caregivers to monitor prohibited areas or to avoid hazardous situations, such as aggression.

4.4. Scenario

As presented in Section 2, our approach relies on using environmental microphones as a means to detect anxiety in PwD. To show how our approach could be used, we re-visit the prototypical situation in Subsection 4.3a and build two scenarios of the usage of the system.

Scenario 1: *Maricela is an 80-years-old resident diagnosed with moderate dementia 5 years ago. She becomes distressed when she gets confused by not knowing where she is combined with her desire to return home. The caregivers know that Maricela suffers from the Sundown Syndrome, which is reflected in the confusion and restlessness in the PwD caused due to the deprivation of light when the sunsets.*

On a particular night (about 1 am), Maricela wakes up and starts crying because she does not recognise where she is, and she expects to see her husband by her side (who died a few years earlier).

AnxiDetector identifies the sounds of restlessness while Maricela sleeps (sending a notification to the caregiver to be aware of a potential assistance request) and then the sound of her awakening and getting out of her bed. A warning alert is sent to the caregiver who is on the night shift to go to Maricela’s room to provide individual support.

Scenario 2: *John is a 75-years-old resident diagnosed with moderated dementia 7 years ago. Since he was diagnosed, he has lived in a public residence who take care of him. He usually is calm and behaves friendly with other residents. However, John’s relatives have reported his sensibilities to disturbance such as low noises, when he tends to be aggressive to others.*

During the transition of a popular night TV show; John was calmly seated in the common room. Out of a sudden, two residents started to talk loudly to each other. AnxiDetector recognises that there is a disturbance event happening in the

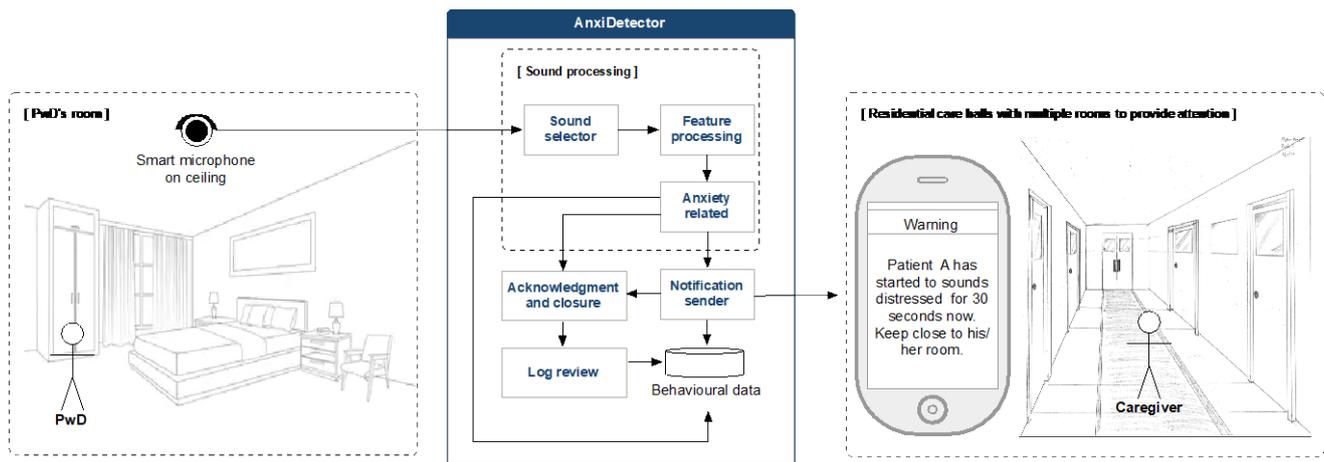


Figure 1. AnxiDetector architecture consisting of six components to collect, filter, and select audio features from the PwD’s daily life environment (left side of diagram). AnxiDetector automatically monitors anxiety episodes and triggers a notification to caregivers through their smartphone or smartwatch (right side of diagram).

shared room, hence, sends a warning notification to the caregivers. While, the caregivers, understand that this is not an emergency, but rather a potential anxiety treat, they decide to prioritise and finish their ongoing task before heading to the shared room.

After a few seconds, John loses control and start shouting and violently treating them. Given the early notification sent by AnxiDetector, the caregivers arrive just in time, before John behaviour escalates further. Given the early action of the caregivers, they quickly manage to calm John down.

Previous scenarios are summarized in a use case narrated by building upon Figure 1. Here, AnxiDetector represents the sensing actor which monitors the acoustic sounds from the environment; including those emitted by the PwD (Maricela or John). The entry condition for AnxiDetector consists of an anxiety manifestation represented by a pattern of sound (see Section 4), hence, as the AnxiDetector detects a critical sound, it exits condition is satisfied by sending a notification to the caregiver’s smartphone.

In these scenarios, the PwD would benefit from early care to prevent accidents, whereas caregivers would benefit from contextual information to assist PwD unobtrusively.

5. AnxiDetector

AnxiDetector is a prototypical solution that has been devised to investigate the feasibility of implementing smart microphones to anticipate anxiety episodes in PwD who live in a residential care environment.

AnxiDetector consists of a Matrix Voice ESP[‡] module, which is an array of 8 digital microphones which can sense and identify sound energy levels and location of where sounds were produced. The Matrix Voice ESP module includes a microcontroller, which enables processing of algorithms for

voice detection, de-reverberation and noise cancellation, amongst other capabilities.

In this initial prototype of AnxiDetector, the Matrix Voice ESP was connected to a Raspberry Pi Single Board Computer[§]. A software component which provides interaction with the Matrix Voice ESP and performs relevant processing was loaded on the Raspberry Pi.

AnxiDetector is connected to a power supply and to a wireless Internet connection nearby where PwD spend most of their time, such as private rooms or leisure areas. To reduce privacy concerns, data is processed locally on the microphone’s device. The system architecture is presented in Figure 1 and consists of the following components:

Sound processing. The smart microphone monitors sound from the environmental perimeter of the PwD.

- a) A sound selector filters noise focusing on sounds produced by the PwD (e.g., crying, screaming).
- b) A feature processor generates features retrieved from sound frequency, amplitude, speed of sound, direction, pitch, duration, loudness, timbre, sonic texture, spatial location, and pressure level.
- c) Anxiety-related sounds. It classifies the sound input to determine whether the sounds are related to anxiety. Hence, this component conceives a decision-maker that identifies the probability of an episode of anxiety occurring.

Notification sender. Two levels of notifications are available, warning and emergency. A warning notification it is sent when probabilities of an anxiety episode are low (e.g., < 70%). An Emergency notification it is sent when a behavioural pattern on the PwD has been clearly identified; therefore, certainty has been calculated by the Anxiety-related component. This setup is helpful to avoid entering private rooms when a false positive anxiety episode is detected.

Intervention and notification. The notification sent to the caregiver consists of a combination of sound and visual

[‡] <https://www.matrix.one/products/voice>

[§] <https://www.raspberrypi.org>

information triggered on their smartphone/smartwatch device. There is no need to acknowledge the Warning notification since its purpose is only to warn of potential assistance support. However, once the Emergency notification is sent, this component will expect a caregiver to attend the PwD's location and provide adequate support. A closure mechanism will be automatically triggered once anxiety has been treated by the caregiver.

Log review. To improve the detection of future episodes, this component generates a log of behaviours in which pattern behaviour pre-episodes are recorded towards a better understanding of anxiety events.

As more data is logged into the database stored in the server and more relevant patterns are found, it is anticipated there will be a greater insight into the individual behaviour of PwD concerning their anxiety (causes and interventions). This insight will inform the early detection algorithm located within the anxiety-related component.

Note that in this paper we introduce AnxiDetector as an embedded prototype system to be installed within a Raspberry Pi Single Board Computer, its deployment is envisioned to be compatible to Linux operative system platform (e.g., Java as a programming language and PostgreSQL as the database). The Sound processing is a machine learning component envisioned to be deployed utilising TensorFlow Lite** (for deploying a DNN model) or Firefly†† (for deploying an SVM model). Finally, we consider that the notifications should adopt standardised SMS or MMS content; which are compatible with popular mobile devices platforms such as Android and iPhone.

To evaluate the adoption of AnxiDetector, an iterative and incremental development model is used, in which (i) a prototype is co-designed including opinions of caregivers, (ii) the prototype is developed considering the architecture detailed previously, and (iii) validation of the acceptability of the prototype is conducted. Then, a new iteration is reworked when needed until a final prototype will be achieved, as presented in Figure 2.

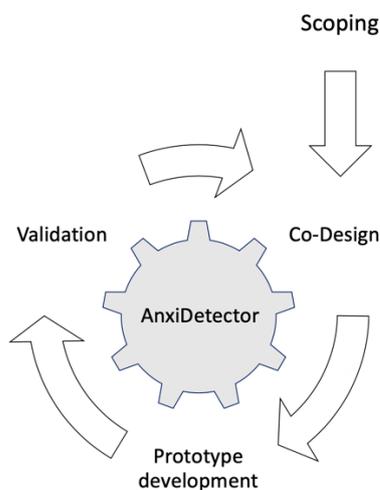


Figure 2. Prototype development model.

** <https://www.tensorflow.org/lite>

6. Preliminary validation

To validate the acceptability of AnxiDetector, we conducted two interviews with caregivers at the Residential Centre ‘Ángeles Cobo López’ in Spain. During this study, we observed that PwD expressed anxiety in different ways, given individual circumstance regarding their mental impairment and their personality.

In this work, we estimated the effectiveness of the detection by adopting nine classes of sounds as representative patterns manifested by PwD when experiencing anxiety. The variety of classes were determined as they were observed in our non-participatory study (Section 4) and as they are reported in the literature - that dementia can worsen the effects of sensory changes by altering how the person perceives external stimuli, such as noise and light [27].

Two groups of audio files were considered: (i) *disturbance* sounds, and (ii) *expression* sounds. The former are sounds that could disturb PwD and include: flushing toilet, knocking door, phone ringing, and bell ringing. The latter are sounds expressed by the PwD when experiencing anxiety, such as crying, verbal expressions (i.e., wow), typing, anger, and screaming.

Due to privacy restrictions, audio recording from the Residence Centre ‘Ángeles Cobo López’ was not possible. Therefore, we relied on publicly available datasets to validate our methodology, as presented below.

Table 1. Relationship between the classes of sound and the open-access dataset from which they were retrieved. Description of the datasets is provided in Section 6.1.

Acoustic event	Group	Dataset
Flushing toilet	Disturbance	Mixed ES [28]
Knocking door	Disturbance	Mixed ES [28]
Phone ringing	Disturbance	Mixed ES [28]
Typing	Disturbance	Mixed ES [28]
Bell ringing	Disturbance	Mixed ES [28]
Verbal expression (i.e., Wow)	Expression	Speech Commands [29]
Screaming	Expression	Zapsplat [30]
Crying	Expression	Freesound [31], Soundsnap [32]
Anger	Expression	Soundsnap [32]

To validate the capabilities of our methodology, we implemented a Deep Neural Network model consisting of 16 hidden layers, in which eight features were conceived as input neurons (energy entropy, short-time energy, spectral roll-off, spectral centroid, spectral flux, relative spectral transform filtering, and Mel-frequency cepstral). Our model was built using a sigmoid as activation function and trained upon 500 epoch iterations.

We calculated the above features creating a vector by splitting each audio file into 23 milliseconds windows with 10 milliseconds overlap, as we have previously proof effective

†† <https://pypi.org/project/firefly-python>

[28]. Then the different length of the audio files was addressed by calculating the standard deviation from each vector.

To extend the capabilities to recognize different PwD’s anxiety events, we trained our model with a small sample of instances per subject (between-subject approach) as described in Section 6.1. We hypothesise that by doing so, we build a more general model capable of identifying a wider variety of acoustic events compared to (for example) training our model by utilising instances from a single subject.

Validation was furthermore done using a Support Vector Machine (SVM) with a linear kernel and the suggested default hyperparameters (version 0.21.3 of the scikit-learn toolkit).

Features for each audio file in the data set were extracted with the openSMILE framework for acoustic feature extraction [35]. We extracted single feature vectors (statistical functionals over whole utterances) of the eGeMAPS feature set [36]. eGeMAPS is a widely used knowledge-driven multi-purpose feature set, mostly used for voice sciences.

For scaling, the mean was removed, and data points were scaled to unit variance.

6.1. Datasets description

The audio files were standardised by converting them into 16-bit little-endian PCM-encoded, and formatting them into WAVE files with a sample rate of 44.1Khz which as it has been proven to be effective in different research studies [33] [34]. Audio files were listened one by one of the authors from this paper, in order to validate the audio file annotation. The datasets used are described below.

Mixed Environmental Sound (Mixed ES). This dataset consists of 40 three-seconds audio files captured from each of the next classes: flushing toilet, knocking door, phone ringing, typing, and bell ringing. The sounds were recorded in an apartment of an older adult using a low-end mobile phone. Four anonymous subjects participated. To conduct our study, the 40 samples of each class were utilized [28].

Speech Commands. This is a set of 64,727 one-second audio files which contains a single spoken English word. The audio files were collected in uncontrolled locations by different people around the world. The number and characteristics of the participants are not reported. To conduct our study, we randomly selected 40 audio files from the verbal expression “wow” [29].

Zapsplat. It consists of free sound effects and a music library offering up to 55,446 tracks for instant download. The length of the files varies. The subjects’ profile recordings are not reported nor the conditions under which they were recorded. To conduct our study, we randomly selected 40 two-seconds audio files from people screaming [30].

Freesound. This is an open-access online server that offers more than 400,000 sound and effects audio files. The length of the files varies. Neither technical information of the subjects nor recording equipment are provided. To conduct our study, we randomly selected 40 one-second audio files from adults crying [31].

Soundsnap^{‡‡}. It consists of a server with more than 278,428 audios files. The length of the files varies, and information of the subject’s producing the audio is not reported. To conduct our study, we randomly selected 40 audio files from emotion events such as adults crying and being angry. Length of the audio files is approximately 2 seconds [32].

6.2. Results

6.2.1 Caregivers interviews outcomes

The acceptability of AnxiDetector was validated by the outcomes of two scenarios presenting its functionality (Presented in Subsection 4.4). In the interviews, the caregivers expressed that the deployment of AnxiDetector could positively impact in two aspects: (i) by providing assurance that there may be no PwD in harm, and (ii) the opportunity to improve the quality of their services.

Quote from Caregiver A: *“This technology will totally make me feel more comfortable because normally I will be under stress thinking that PwD-A (whose room is located at the main entrance) could have an anxiety episode, while I am patrolling the other end of the corridor”.*

Quote from Caregiver B: *“I think this can help me to run my errands quickly so that I can invest more time with the PwD. Sometimes, in tween rooms (i.e., rooms with two PwD divided by a door), it takes me up to 10 minutes to make a PwD’s bed because I am (simultaneously) walking back and forth to the other room. Hence, with this technology, I can focus on a single one at the time, trusting the other PwD doesn’t require my assistance”.*

6.2.2 Quantitative results

The accuracy of the DNN model was estimated by validating the classification model under a 10-fold cross-validation. When classifying the two groups of classes (i.e., *expression* and *disturbance*), the classification accuracy reaches 98.1% with a precision of 98.0% and 98.1% respectively. The average accuracy when classifying the nine different classes reports a classification accuracy of 92.2%, with the “Verbal expression” and the “Typing” classes having the highest precision of 100% each. The lowest precision was obtained with the class “Scream” (i.e., 74.4%). Table 2 presents the confusion matrix for these results.

Table 2. Confusion matrix for the DNN on the 9 audio classes and 10-fold cross-validation.

a	b	c	d	e	f	g	h	i	Classified as:
32	4	0	0	0	2	0	0	2	a = Screaming
7	32	0	0	0	0	0	0	1	b = Anger
0	1	38	0	0	0	0	0	1	c = Verbal Expression

^{‡‡} Educational license was granted by the Soundsnap company.

1	0	0	37	1	0	0	0	1	d = Flushing toilet
0	0	0	0	40	0	0	0	0	e = Knocking door
3	0	0	0	1	35	1	0	0	f = Phone ringing
0	0	0	1	0	0	39	0	0	g = Bell ringing
0	0	0	0	0	0	0	40	0	h = Typing
0	0	0	1	0	0	0	0	39	i = Crying

The classification accuracy of the SVM was evaluated with 10-fold cross-validation. The distinction of the two *expression* and *disturbance* categories was made with an accuracy of 99.2% each. For the multiclass classification problem, the model achieved an overall classification accuracy of 93.0%, with the classes “Knocking door, Phone ringing, Bell ringing, and Typing” having the highest precision of 100%, and the “Scream” class reporting the lowest precision of 77.8%. Table 3 presents the confusion matrix for these results.

Table 3. Confusion matrix for SVM on the classification of the 9 audio classes and 10-fold cross-validation.

a	b	c	d	e	f	g	h	i	Classified as:
28	9	1	0	0	1	0	0	1	a = Screaming
6	32	1	0	0	0	0	0	1	b = Anger
1	0	37	0	2	0	0	0	0	c = Verbal Expression
0	0	0	39	1	0	0	0	0	d = Flushing toilet
0	0	0	0	40	0	0	0	0	e = Knocking door
0	0	0	0	0	40	0	0	0	f = Phone ringing
0	0	0	0	0	0	40	0	0	g = Bell ringing
0	0	0	0	0	0	0	40	0	h = Typing
1	0	0	0	0	0	0	0	39	i = Crying

7. Limitations and Constraints

Overall, two significant challenges to consider during the implementation stage will be to differentiate between sounds produced by a PwD from another PwD from an artificial environmental source, such as on TV. This is of particular concern as this could be misinterpreted by the AnxiDetector and PwD within the environment being observed. To address this, future work will take advantage of the directionality feature of the Matrix Voice ESP smart microphones. Specifically, this will be leveraged to map the location of appliances such as TVs or radios and filtering as appropriate. Another challenge is the characterisation of parameters according to how PwD manifest anxiety episodes, since, for example, while one PwD could express it by shouting, another could do so by walking. For the next step of this project, we will implement a first hardware version of AnxiDetector; hence, a more detailed architecture will be provided.

8. Recommendations

In this section, we provide a brief list of recommendations acquired during the design and development of this paper:

- **Understanding the end-user and their environment.** Which includes the qualitative analysis to understand better the relevant sounds associated with the PwD’s behaviours and anxiety episodes. We also found it useful to identify the common background sounds exposed in the living environment, in order to properly discriminate them and reduce the amount of false positives classification events.
- **Training of the model with appropriate samples.** Although publicly available datasets provide useful audio to train a classification model, it is encouraging to collect audio samples from the real-living environment in order to build a more robust model.
- **Maintaining privacy.** Given the sensitive information that can be captured by audio sensors, edge processing is recommendable, so that data can be processed in-site rather than remotely.
- **Personalising.** Given that not all sounds manifested by PwD are related to anxiety episodes, systems like AnxiDetector should be personalised.
- **Feedback.** As a personalization technique, one can adopt the active learning approach, so the model builds up as the caregivers discriminate between the different manifestations of anxiety in PwD’s.

9. Conclusion and Future work

This paper has presented a feasibility study of using environmental smart microphones to detect early stages of anxiety in PwD based on the identified auditory manifestations of anxiety. We presented the outcomes from a non-participatory observation of PwD and interviews with the caregivers from the Residential Centre ‘Ángeles Cobo López’. We identified the existence of auditory manifestation before and during anxiety episodes. The design and intended use of a low-fidelity prototype called AnxiDetector have been presented. Initial results from an evaluation of our prototype were presented. These results indicate that the use of this type of smart microphone-based technology would be of great support for caregivers. Additionally, two sound classification models based on a DNN and SVM were implemented to showcase the technical feasibility of AnxiDetector. Using the DNN model, a classification accuracy of 98.1% was obtained, whereas accuracy of 99.2% was achieved utilising the SVM for classifying the two groups of classes (*i.e.*, *expression* and *disturbances*). An accuracy of 92.2% and 93.0%, for the DNN and SVM models; respectively, when classifying the nine different classes of sound selected as an acoustic representation of disturbances and manifestations associated to anxiety episode in PwD.

As part of our future work, we will rely on transfer learning and active learning techniques, so that only a small amount of data is required to initiate the anxiety identification task. In particular, active learning techniques are envisioned to improve the quality of anxiety detection by providing

annotation to describe the relevance of the notifications sent by AnxiDetector. Moreover, we are interested in carrying out trials in care homes in Northern Ireland, Mexico, Germany, and Spain to help generalise the acceptability of this approach across cultures.

Acknowledgements.

We especially thank AgeingLab for their support in logistics to make this study possible at the Residential Centre 'Ángeles Cobo López'. This project has received partial funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 734355 and the Global Challenges Research Fund (GCRF) under the GCRF Building Capacity and Networking Grant.

References

- [1] M. Prince, A. Wimo, M. Guerchet, A. Gemma-Claire, Y.-T. Wu, and M. Prina, "World Alzheimer Report 2015: The Global Impact of Dementia - An analysis of prevalence, incidence, cost and trends," *Alzheimer's Dis. Int.*, 2015.
- [2] S. Shaji, S. Bose, and S. Kuriakose, "Behavioral and psychological symptoms of dementia: A study of symptomatology," *Indian J. Psychiatry*, 2009.
- [3] A. M. Saks, "Moderating effects of self-efficacy for the relationship between training method and anxiety and stress reactions of newcomers," *J. Organ. Behav.*, 1994.
- [4] P. J. Seignourel, M. E. Kunik, L. Snow, N. Wilson, and M. Stanley, "Anxiety in dementia: A critical review," *Clinical Psychology Review*. 2008.
- [5] L. Rozzini et al., "Anxiety symptoms in mild cognitive impairment," *Int. J. Geriatr. Psychiatry*, 2009.
- [6] V. R. Badrakalimuthu and A. F. Tarbuck, "Anxiety: a hidden element in dementia," *Adv. Psychiatr. Treat.*, 2012.
- [7] H. Twelftree and A. Qazi, "Relationship between anxiety and agitation in dementia," *Aging Ment. Heal.*, 2006.
- [8] E. Labarge, "A preliminary scale to measure the degree of worry among mildly demented alzheimer disease patients," *Phys. Occup. Ther. Geriatr.*, 1993.
- [9] K. K. Shankar, M. Walker, D. Frost, and M. W. Orrell, "The development of a valid and reliable scale for rating anxiety in dementia (RAID)," *Aging Ment. Heal.*, 1999.
- [10] J. L. Cummings, et al., "The Neuropsychiatric Inventory: Comprehensive assessment of psychopathology in dementia," *Neurology*, 1994.
- [11] N. Hernandez-Cruz, M. Garcia-Constantino, J. Beltran-Marquez, D. Cruz-Sandoval, I. H. Lopez-Nava, I. Cleland, J. Favela, C. Nugent, A. Ennis, J. Rafferty, and J. Synnott. 2019. Study Design of an Environmental Smart Microphone System to Detect Anxiety in Patients with Dementia. In Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth'19). 2019.
- [12] J. Preece, Y. Rogers, and H. Sharp, "Interaction Design: Beyond Human-Computer Interaction," *Design*, 2002.
- [13] G. Giannakakis et al., "Stress and anxiety detection using facial cues from videos," *Biomed. Signal Process. Control*, 2017.
- [14] J. W. Weeks, A. Srivastav, A. N. Howell, and A. R. Menatti, "'Speaking more than words': Classifying men with social anxiety disorder via vocal acoustic analyses of diagnostic interviews," *J. Psychopathol. Behav. Assess.*, 2016.
- [15] J. Gu et al., "Wearable Social Sensing and Its Application in Anxiety Assesment," in Proceedings - 2017 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery, CyberC 2017, 2018.
- [16] W. Simm, M. A. Ferrario, A. Gradinar, and J. Whittle, "Prototyping 'Clasp': Implications for designing digital technology for and with adults with autism," in 2014 ACM SIGCHI Conference on Designing Interactive Systems, DIS 2014, 2014.
- [17] M. A. Williams, A. Roseway, C. O'Dowd, M. Czerwinski, and M. R. Morris, "SWARM: An Actuated Wearable for Mediating Affect," *Proc.* 9, 2015.
- [18] L. Cruz, J. Rubin, R. Abreu, S. Ahern, H. Eldardiry, and D. G. Bobrow, "A wearable and mobile intervention delivery system for individuals with panic disorder," in Proceedings of the 14th International Conference on Mobile and Ubiquitous Multimedia - MUM '15, 2015.
- [19] H. Haritha et al. 2017. Automating Anxiety Detection using Respiratory Signal Analysis. *2017 IEEE Region 10 Symposium*.
- [20] N. Chaitanya M., S. Jayakkumar, E. Chong, and C. H. Yeow. 2017. A wearable, EEG-based massage headband for anxiety alleviation. *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*.
- [21] Y. Zheng, T. C. Wong, B. H. Leung, and C. C. Poon, "Unobtrusive and multimodal wearable sensing to quantify anxiety," *IEEE Sensors Journal*, 2016.
- [22] M. Elgendi, and C. Menon, "Assessing Anxiety Disorders Using Wearable Devices: Challenges and Future Directions," *Brain sciences*, 2019.
- [23] S. Rajasekaran, C. Luteran, H. Qu, and C. Riley-Doucet. 2011. A Portable Autonomous Multisensory Intervention Device (PAMID) for Early Detection of Anxiety and Agitation in Patients with Cognitive Impairment. *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*.
- [24] Mahoney, Edward L and Mahoney, Diane F, "Acceptance of wearable technology by people with Alzheimer's disease: Issues and accommodations", *American Journal of Alzheimer's Disease and Other Dementias*. SAGE Publications Sage CA: Los Angeles, CA, 2010.
- [25] G. Cipriani, C. Lucetti, C. Carlesi, S. Danti, and A. Nuti, "Sundown syndrome and dementia," *Eur. Geriatr. Med.*, 2015.
- [26] D. Collector and F. G. Module, "Qualitative Research Methods Overview," *Qual. Res. Methods A Data Collect. F. Guid.*, 2011.
- [27] J. Dewing, "Caring for people with dementia: noise and light.," *Nursing older people*. 2009.
- [28] J. Beltrán, E. Chávez, and J. Favela, "Scalable identification of mixed environmental sounds, recorded from heterogeneous sources," *Pattern Recognit. Lett.*, 2015.
- [29] Warden P. Speech Commands: A public dataset for single-word speech recognition, 2017. Available from http://download.tensorflow.org/data/speech_commands_v0.01.tar.gz
- [30] Zapsplat. Online server for free sound effects. Available from <https://www.zapsplat.com/sound-effect-category/carts-and-trolleys/>
- [31] Freesounds. Online server for free sound effects. Available from <https://freesound.org/>
- [32] Soundsnap. Online server with sound effects. Available from <https://www.soundsnap.com/>
- [33] L. Hertel, H. Phan, and A. Mertins, "Comparing time and frequency domain for audio event recognition using deep learning," in Proceedings of the International Joint Conference on Neural Networks, 2016.
- [34] A. Manzo-Martinez, A. C. Ramos-Rascon, G. Ramirez-Alonso, F. Gaxiola, R. Cornejo, and J. A. Camarena-Ibarrola, "Classification of acoustic events in a kitchen environment using multiband spectral entropy," in 2018 IEEE International Autumn Meeting on Power, Electronics and Computing, ROPEC 2018, 2019.
- [35] Eyben, F., Weninger, F., Gross, F., Schuller, B., "Recent

developments in openSMILE, the munich open-source multimedia feature extractor”, In Proceedings of the 21st ACM international conference on Multimedia (MM '13). 2013.

- [36] Eyben, F., Scherer, K.R., Schuller, B.W., Sundberg, J., Andre, E., Busso, C., Devillers, L.Y., Epps, J., Laukka, P., Narayanan, S.S., Truong, K.P., “The Geneva Minimalistic Acoustic Parameter Set (GeMAPS) for Voice Research and Affective Computing”. IEEE Transactions on Affective Computing 7, 2016.