Detection and Management of Depression in Cancer Patients using Augmented Reality Technologies, Multimodal Signal Processing and Persuasive Interfaces

Alexandros Roniotis[1], Haridimos Kondylakis[2], Manolis Tsiknakis[1,2]
[1]Department of Informatics Engineering, TEI Crete, Heraklion, Greece
[2]Institute of Computer Science (ICS), FORTH, Heraklion, Greece
roniotis@staff.teicrete.gr

***Abstract*—This visual paper aims at proposing a framework for detecting depression in cancer patients using prosodic and statistical features extracted by speech, while chatting with a virtual coach.**

Keywords- virtual coach; cancer; detecting depression; machine learning; MFCCs

I. INTRODUCTION

Cancer is a major health problem for developed countries and accounts for 14.6% of all deaths worldwide [1]. In addition to physical weakness, knowledge about the disease and the upcoming death has a direct impact on the patient's psychological condition, resulting in increased frequency of depression [2]. Clinical depression, one of the most common mental disorders in cancer patients can only be diagnosed by a clinician and a clinical psychologist and treated through psychotherapy and antidepressants [3].

The occurrence of depression in patients with cancer accelerates disease progression, but the timely assessment and recognition of mental disorders requires intensive training and longtime experience [4]. The development of an automatic machine learning system for the automatic detection of depressive symptoms could significantly help clinicians in the early detection and psychotherapeutic intervention [5].

The development of health applications for modern smart devices such as mobile phones or tablets, is a rapidly growing sector. Automated health monitoring using programs that interact with the patient, called virtual coaches, can contribute to the self-treatment of the patient, reducing clinical monitoring costs and helping with the timely notification of the supervising doctor [6]. Not only is the financial cost of using such a system low, but also the familiarity of patients with modern devices is increasing over the years.

By integrating a virtual coach in a mobile device of a cancer patient, the patient could monitor his mental state at anytime, anywhere, without the use of additional equipment or the need to get monitored within any clinical setting. Moreover, the technical requirements of modern mobile phones now allow the processing of complex data, such as voice signals, as they have powerful processors, large storage space and memory.

II. BACKGROUND

Clinical depression affects thinking, mood, behavior and physical condition. Specifically, the voice and the articulation of a person is directly affected by his mental state [7]. This fact favors the use voice features as biomarkers of depression [8]. Voice characteristics correlate with the presence and grade of depression and are often used to develop automatic classifiers. These features are usually classified as normal, mild, moderate, severe or very severe mental disorder according to the Hamilton Rating Scale for Depression (HAM-D), the 9-item Patient Health Questionnaire (PHQ-9) or [9] Beck Depression Index (BDI) [10].

Depression causes physiological changes that affect speech production process by mainly differentiated stimulation of muscles and vocal cords and the change in respiratory rate [11]. As a result, the acoustic quality of the sound produced is affected and to such an extent that it is objectively measurable [12].

Some features that could be used for the classification of mental status of the patient are prosodic or spectral. Some prosodic features are the rate of speech, the fundamental voice frequency (F0), the intensity and the energy of the voice and glottal features [8]. The most frequent spectral features include the formants (the eigenfrequencies of the vocal organ), the power spectral density (PSD) and Mel Frequency Cepstral coefficients (MFCCs) [13].

III. OBJECTIVES

The high costs and the modest effectiveness of health system are often attributed to lack of patient's engagement at home. The effective engagement of patients is considered the "trillion opportunity" in health [14]. In previous years, dozens of companies have invested in developing applications for smart devices in health (mHealth apps) to involve patients in monitoring their health themselves, follow the medication therapy, reporting their symptoms, etc. But it seems that the devotion of patients was moderate to low, mainly because they were not prompted by a third party to keep using the application. Instead, when using the application under the constant presence of a doctor or nurse, the results were encouraging [15]. Thus, it appears that the existence of a coach during the usage of the application could improve its effectiveness [15].

It is interesting that after the boom of the first years, applications for smartphones have lost their initial momentum [16]. Now the new generation of applications is considered that of chatbots, i.e. automated communication applications in which the user converses with the device itself [14]. Chatbots are considered the next big station to healthcare applications where a virtual coach will discuss

with the patient, urge him to pursue his action, raise questions and guide the discussion according to the answers received and processed [14]. Such an application could be developed for exporting depressive biomarkers, such as those mentioned in the preceding paragraphs.

IV. METHODS

The proposed scientific research is divided into two main stages. At first a training set made of cancer patients is developed. This first stage includes the extraction of audio data from the patients, but also filling a depression questionnaire. Then, data is processed and feature vectors are extracted to compose the training set.

In the second stage, a user's speech signal is recorded and is then processed to generate a feature vector. The vector is then classified into a depression scale, using the training set of the previous stage. Then, depending on the results of the classification, the virtual coach urges the patient to do some more psychotherapeutic activities, follow his therapy, and notify the supervising doctor when depression scale is classified as severe.

A. Training Set

Initially, an application for mobile devices will be developed in order to generate some data from at least 50 cancer patients at the Psychiatric Clinic of Heraklion University Hospital, Greece, with which all relevant legal agreements should have been signed. The application will be installed on a mobile device such as a mobile phone or tablet. The first time it runs, the program will show the patient a BDI-II questionnaire [10], instead of the PHQ. The answers will be stored for processing at a later stage.

The virtual coach then appears, which has been developed with the open source BotLibre (botlibre.com), which, based on Positive psychology theory [17]. It starts an interactive chat with the user, posing specific questions. The patient's speech is recorded and stored on the device for processing at a later stage.

BDI answers will be then used as the ground truth to extract the depression scale and develop the training set. Audio data from the user's conversation with the chatbot will be used to extract feature vectors. The spectral features to be extracted are the Mel-frequency cepstral coefficients (MFCCs) [13], that is the most common sound feature that is used internationally. More statistical features that are considered biomarkers of depression will be computed, such as speech duration, the duration of pauses, speech rate and the response delay after the question of the virtual coach. All these features, along with the gender of the user, will form the training set.

B. Classification of Depression Scale

In the second stage, a user can be classified in one of the depression scales. The user uses the application, which will repeat the same inquiry procedure with the first stage. More specifically, the virtual coach will ask the same questions and the signal of the user's speech will be recorded for producing the feature vector to be classified using the training set.

Another key step will be to assess the effectiveness of implementing the automatic classification of the patient's depression scale. Towards evaluation one could implement the leave-one-out method, in which each vector of the training set is sorted after removing the same vector. Finally, the resulting Confusion Matrix is used for estimating accuracy.

ACKNOWLEDGEMENT

This research is supported by IKY scholarships programme and co-financed by the European Union (European Social Fund - ESF) and Greek national funds through the action "Reinforcement of Postdoctoral Researchers" in the framework of the Operational Programme "Human Resources Development Program, Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) 2014 − 2020

REFERENCES

[1] World Health Organization, "World Cancer Report 2014," 2014.

[2] Greek National Research Institute, "Mental Health - Contemporary Approaches and Reflections", Athens, 2011.

[3] C. Mathers, T. Boerma, and D. Ma Fat, "The global burden of disease: 2004 update," WHO, Geneva, Switzerland, 2008.

[4] American Cancer Society, "Depression Increases Cancer Patients' Risk Of Dying," 2009.

[5] J. Cesar, and F. Chavoushi, "Depression, WHO - Priority Medicines for Europe and the World (2013 Update), 2013.

[6] A. Wissen, C. Vinkers, and A. Halteren, "Developing a Virtual Coach for Chronic Patients: A User Study on the Impact of Similarity, Familiarity and Realism," Proc 11th Int Conf Per Techn, 2016.

[7] S. Sahu, and C. Espy-Wilson, "Effect of depression on syllabic rate of speech," J Acoustical Society of America, vol. 138, p. 1781, 2015.

[8] H. Hamilton, "A rating scale for depression," Journal of Neurology, Neurosurgery and Psychiatry, vol. 23, pp. 56-62, 1960.

[9] A. Beck, R. Steer, and G. Brown, "Beck Depression Inventory-II," The Psychological Corporation, 1961-1996.

[10] S. Kreibig, "Autonomic nervous system activity in emotion: a review," Biol. Psych., vol. 84, no. 3, pp. 394-421, July 2010.

[11] N. Cummins, S. Scherer, J. Krajewski, S. Schnieder, J. Epps, and T. Quatieri, "A review of depression and suicide risk assessment using speech analysis," Speech Communications, vol. 71, pp. 10-49, 2015.

[12] A. Pampouchidou, O. Simantiraki, A. Fazlollahi, D. Manousos, Pediaditis M, A. Roniotis, G. Giannakakis, F. Meriaudeau, P. Simos, K. Marias, F. Yang, and M. Tsiknakis, "Depression Assessment by Fusing High and Low Level Features from Audio, Video, and Text,» AVEC 2016, Amsterdam, Netherlands, pp. 27-34, 2016.

[13] T. Laosaphan, and T. Yingthawornsuk, "Classification of Depressed Speakers based on MFCC in Speech Samples," ICAEEE 2012, Pattaya, Thailand.

[14] M. Yuan, "Chatbots: Building Intelligent Bots," Addison-Wesley, 2016.

[15] D. Freeney, "Usability versus Persuasion in an Application Interface Design," Inst Innov Design & Eng, Mälardalen Uni, Sweden, 2014.

[16] The Nielsen Company, "So Many Apps, So Much More Time for Entertainment," 2016.

[17] R. Proyer, F. Gander, S. Wellenzohn, and R. Willibald, "Positive psychology interventions in people aged 50–79 years: long-term effects of placebocontrolled online interventions on well-being and depression," Ag Ment Health, vol. 18, no. 8, pp. 997-1005, 2014.