Spontaneous Speech and Emotional Response modeling based on One-class classifier oriented to Alzheimer Disease diagnosis

K. Lopez-de-Ipiña[1], J.B. Alonso[2], N. Barroso[1], J. Solé-Casals[3], M. Ecay-Torres[4], P. Martinez-Lage[4]

[1] Department of System Engineering and Automation, University of the Basque Country, Spain, karmele.ipina@ehu.es
[2] Universidad de Las Palmas de Gran Canaria, IDeTIC
[3] Digital Technologies Group, University of Vic
[4] Neurology Department CITA-Alzheimer Foundation

Abstract— **The purpose of our project is to contribute to earlier diagnosis of AD and better estimates of its severity by using automatic analysis performed through new biomarkers extracted from non-invasive intelligent methods. The methods selected in this case are speech biomarkers oriented to Spontaneous Speech and Emotional Response Analysis. Thus the main goal of the present work is feature search in Spontaneous Speech oriented to pre-clinical evaluation for the definition of test for AD diagnosis by One-class classifier. One-class classification problem differs from multi-class classifier in one essential aspect. In one-class classification it is assumed that only information of one of the classes, the target class, is available. In this work we explore the problem of imbalanced datasets that is particularly crucial in applications where the goal is to maximize recognition of the minority class as in medical diagnosis. The use of information about outlier and Fractal Dimension features improves the system performance.**

.

Keywords— **One-class classifier, Nonlinear Speech Processing, Alzheimer disease diagnosis, Spontaneous Speech, Fractal Dimensions.**

I. Introduction

Alzheimer's Disease (AD) is the most common type of dementia among the elderly. It is characterized by progressive and irreversible cognitive deterioration with memory loss and impairments in judgment and language, together with other cognitive deficits and behavioral symptoms. The cognitive deficits and behavioral symptoms are severe enough to limit the ability of an individual to perform everyday professional, social or family activities. As the disease progresses, patients develop severe disability and full dependence. An early and accurate diagnosis of AD helps patients and their families to plan for the future and offers the best opportunity to treat the symptoms of the disease. According to current criteria, the diagnosis is expressed with different degrees of certainty as possible or probable AD when dementia is present and other possible causes have been ruled out. The diagnosis of definite AD requires the demonstration of the typical AD pathological changes at autopsy [1,2,3].In addition to the loss of memory, one of the

major problems caused by AD is the loss of language skills. We can meet different communication deficits in the area of language, including aphasia (difficulty in speaking and understanding) and anomia (difficulty in recognizing and naming things). The specific communication problems the patient encounters depend on the stage of the disease [3,4,5]. The main goal of the present work is feature search in Spontaneous Speech Analysis (ASSA) an Emotional Response Analysis (ERA) oriented to pre-clinical evaluation for the definition of test for AD diagnosis. These features will define control group (CR) and AD disease. Non-invasive Intelligent Techniques of diagnosis may become valuable tools for early detection of dementia. Moreover, these techniques are very low-cost and do not require extensive infrastructure or the availability of medical equipment. They are thus capable of yielding information easily, quickly, and inexpensively [6,7]. This study is focuses on early AD detection and its objective is the identification of AD in the pre-clinical (before first symptoms) and prodromic (some very early symptoms but no dementia) stages. The research presented here is a complementary preliminary experiment to define thresholds for a number of biomarkers related to spontaneous speech. Feature search in this work is oriented to pre-clinical evaluation for the definition of test for AD diagnosis Obtained data will complement the biomarkers of each person [8].

The rest of this paper is organized this way: In Section II, the materials are presented. Section III explains the methodology of the experiments, Section IV shows the experimental results, and finally, conclusions are presented in Section V.

II. Materials

Trying to develop a new methodology applicable to a wide range of individuals of different sex, age, language and cultural and social background, we have built up a multicultural and multilingual (English, French, Spanish, Catalan, Basque, Chinese, Arabian and Portuguese) database with video recordings of 50 healthy and 20 AD patients (with a prior diagnosis of Alzheimer) recorded for 12 hours and 8 hours respectively. The age span of the individuals in

the database was 20-98 years and there were 20 males and 20 females. This database is called AZTIAHO. All the work was performed in strict accordance with the ethical guidelines of the organizations involved in the project. The recordings consisted of videos of Spontaneous Speech – people telling pleasant stories or recounting pleasant feelings as well as interacting with each other in friendly conversation. The recording atmosphere was relaxed and non-invasive. The shorter recording times for the AD group are due to the fact that AD patients find speech more of an effort than healthy individuals: they speak more slowly, with longer pauses, and with more time spent on efforts to find the correct word and uttering speech disfluencies or break messages. In the advanced stage of the disease, they find this effort tiring and often want to stop the recording. We complied with their requests. The video was processed and the audio extracted in wav format (16 bits and 16 Khz). The first step was removing non-analyzable events: laughter, coughing, short hard noises and segments where speakers overlapped. Next, background noise was removed using denoiser adaptive filtering. After the pre-processing, about 80% of the material from the control group and 50% of the material from the AD group remained suitable for further analysis. The complete speech database consists of about 60 minutes of material for the AD group and about 9 hours for the control. The speech was next divided into consecutive segments of 60 seconds in order to obtain appropriate segments for all speakers, resulting finally in a database of about 600 segments of Spontaneous Speech. Finally for experimentation from the original database, a subset of 20 AD patients was selected (68-96 years of age, 12 women, 8 men, with a distribution in the three stages of AD as follows: First Stage [ES=4], Secondary Stage [IS=10] and Tertiary stage [AS=6]). The control group (CR) was made up of 20 individuals (10 male and 10 female, aged 20-98 years) representing a wide range of speech responses. This subset of the database is called AZTIAHORE[7].

III. METHODS

In previous work [7] the goal of the experimentation was to examine the potential of the selected features to help in the automatic measurement of the degradation of Spontaneous Speech, Emotional Response and their manifestation in people with AD as compared to the control group. The experiments have analyzed the automatic measurement and definition of appropriate features in Spontaneous Speech, Emotional Response Analysis and Integral Response in the speech of people with AD. Four groups/classes (CR, ES, IS, AS) and four different feature sets were evaluated (section 3.2.1.4). The automatic classification of speech is based on several classifiers. We used five different classifiers: (1) a Support Vector Machine (SVM), (2) a Multi Layer Percep-

tron (MLP) (3) a KNN Algorithm, (4) Decision Trees (DT) and (5) a Naive Bayes net. The approach's performance was very satisfactory and promising results for early diagnosis and classification of AD patient groups but medical doctors propose new experimentation oriented to detect mainly early stage. The goal of this new experimentation is to detect changes with regard to CR group and outliers which point to presence of AD's symptoms. One class classification will be use for this propose.

A. Feature extraction

A.1 Automatic Spontaneous Speech Analysis (ASSA)

Spoken language is one of the most important elements defining an individual's intellect, social life, and personality; it allows us to communicate with each other, share knowledge, and express our cultural and personal identity. Spoken language is the most spontaneous, natural, intuitive, and efficient method of communication among people. Therefore, the analysis by automated methods of Spontaneous Speech (SS – free and natural spoken communication), possibly combined with other methodologies, could be a useful non-invasive method for early AD diagnosis. The analysis of Spontaneous Speech fluency is based on three families of features (SSF set), obtained by the Praat software package [9] and software that we ourselves developed in MATLAB. For that purpose, an automatic Voice Activity Detector (VAD) [10] has extracted voiced/unvoiced segments as parts of an acoustic signal. These four families of features (Spontaneous Speech Features, SSF) include: duration, time and frequency domain analysis (pauses, short time energy, centroid).

A.2 Emotional Response Analysis

In this study, we aim to accomplish the automatic selection of emotional speech by analyzing three families of features in speech (Emotional Features, EF): acoustic, voice quality: duration an Emotional Temperature (ET) [XXX] (pitch, energy, intensity, shimmer, jitter) .

A.3 Fractal Dimension

The Fractal Dimension (FD) is one of the most significant features which describe the complexity of a system and could help in the detection of subtle changes for early diagnosis. Moreover this feature has in the ability to capture the dynamics of the system and thus relevant variations in speech utterances. More precisely, an implementation of Higuchi's algorithm [11] in order to add this new feature to the set that feeds the training process of the model. Most of the fractal systems have a characteristic called self-similarity. An object is self-similar if a close-up examination of the object reveals that it is composed of smaller versions of itself.
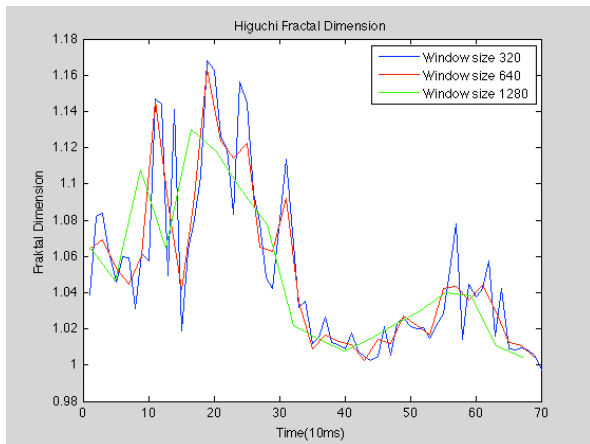
Figure 1. Higuchi Fractal Dimension (HFD) for an AD case for different window sizes.

Self-similarity can be quantified as a relative measure of the number of basic building blocks that form a pattern, and this measure is defined as the Fractal Dimension. Higuchi was choice because it has been reported to be more accurate in previous works with under-resourced conditions [12]. In this feature set Higuchi Fractal Dimension (HFD), and its max, min, mean and standard deviation have been used in HFD set.

A.3 Automatic classification by One-class classifier

One-class classification problem [13] differs from multi-class classifier in one essential aspect. In one-class classification it is assumed that only information of one of the classes, the target class, is available. This means that just example objects of the target class can be used and that no information about the other class of outlier objects is present. The different terms such as fault detection, anomaly detection, novelty detection and outlier detection originate from the different applications to which one class classification can be applied. The boundary between the two classes has to be estimated from data of only control class. The task is to define a boundary around the target class, such that it accepts as much of the target/control objects as possible, while it minimizes the chance of accepting outlier objects. In the literature a large number of different terms have been used for this problem. The term one-class classification originates from [14]. The application is as follows: The first application for one class classification (also called data description as it forms the boundary around the whole available data) is outlier detection, to detect uncharacteristic objects from a dataset. Secondly, data description can be used for a classification problem where one of the classes is sampled very well, while the other class is severely under sampled. The measurements on the under sampled class might be very expensive or difficult to obtain. Finally, the last possible use of the outlier detection is the comparison of two data sets. Assume that a classifier has been trained (in a long and difficult optimization process) on some (possibly expensive) data [15]. As explained above, the second application of outlier detection is in the classification problem where one of the classes is sampled very well but it is very hard and expensive, if not impossible, to obtain the data of the second class. One of the major difficulties inherent in the data (as in many medical diagnostic applications) is this highly skewed class distribution. The problem of imbalanced datasets is particularly crucial in applications where the goal is to maximize recognition of the minority class. WEKA [16] software was used in carrying out the experiments. The base classifiers to be used in the experimentation were Bagging and MLP, with k-fold cross-validation (k=10).

IV. EXPERIMENTAL RESULTS

The task was Automatic Classification, with the classification targets being: healthy speakers without neurological pathologies and speakers diagnosed with AD. The experimentation was carried out with AZTIAHORE. The results have been analyzed with regard to the feature set described in III. Experimentation has been divided to test One-class classifier only with speech samples from CR group and with information about outliers (patients with AD). The results are shown in Figure 1.

A. One-class Classifier only with information about CR group:

The results are satisfactory for this study in the case of MLP classifier. Bagging paradigm presents lower computational cost but very poor results. The new fractal features improve the system but they can improve Bagging performance. Window size: For HFD algorithm three different window size have been used: 1280, 640 and 320 samples. The best results have been obtained for a window size of 320 samples. The best results are obtained for integral feature set, which mixes features relative to Spontaneous Speeech and Emotional Response (SSF+HFD+EF+ET).

B. One-class Classifier with information about outliers:

The results are satisfactory for this study because they obtained very good results not only for MLP classifier but also for Bagging, which presents lower computational cost. The new fractal features improve the system for both paradigms. The best results have been obtained for a window size of 320 samples in HFD. In this case also the best results are obtained for SSF+HFD+EF+ET set.
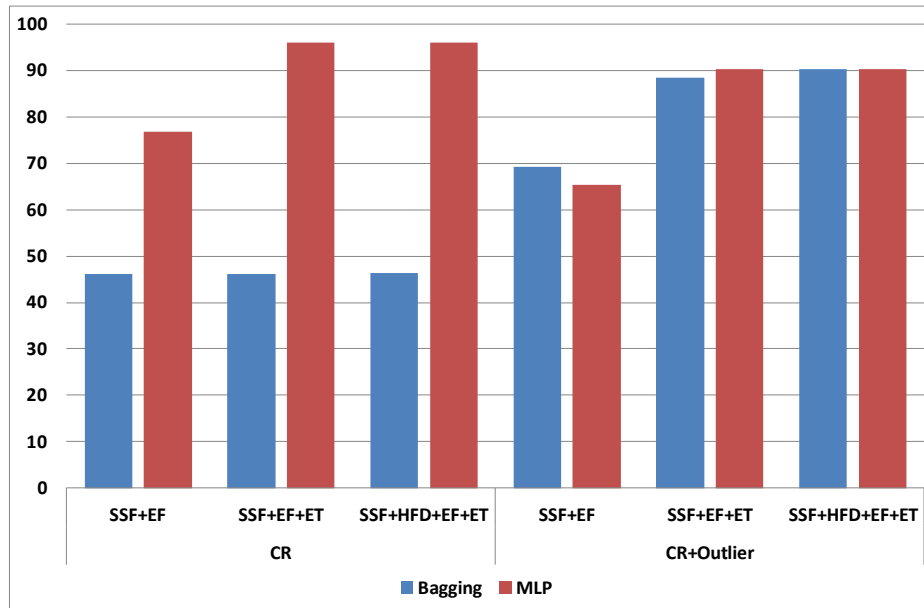
Figure 2. Higuchi Fractal Dimension (HFD) for an AD signal for different window sizes.

V. CONCLUSIONS

The main goal of the present project is feature search in through new biomarkers extracted from non-invasive intelligent methods as Spontaneous Speech and Emotional Response Analysis oriented to pre-clinical evaluation for the definition of test for AD diagnosis. These features are of great relevance for health specialists to define health people and AD disease symptoms. One-class classification has been used in the experimentation. In one-class classification it is assumed that only information of one of the classes, the target class, is available. In this work we explore the problem of imbalanced datasets by using information about outlier and Fractal Dimension features improves the system performance. The approach of this work complements the previous multi-class modelling and is robust in terms of capturing the dynamics of the whole waveform, and it offers many advantages in terms of computability, and it also makes easier to compare the power of the new features against the previous ones. In future works we will introduce new features relatives to speech modelling oriented to standard medical tests for AD diagnosis and to emotion response analysis.

REFERENCES

1. Mc Kahn G, et al.Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Workgroup on AD. 1984; 24:939-944.
2. McKhann GM et al. The diagnosis of dementia due to Alzheimer's disease: Recommendations from the NIAA's Association workgroups on diagnostic guidelines for AD. Alzheimers Dement. 2011 May;7(3):263-9.
3. Van de Pole, L.A., et al., The effects of age and Alzheimer's disease on hippocampal vol-umes, a MRI study. Alzheimer's and Dementia, 2005. 1(1, Supplement 1): p. 51.
4. Morris JC, The Clinical Dementia Rating (CDR): current version and scoring rules. Neu-rology, 1993. 43: p. 2412b-2414b.
5. American Psychiatric Association, 2000. Diagnostic and Statistical Manual of Mental dis-orders, 4th Edition Text Revision. Washington DC.
6. 7. M. Faundez-Zanuy et al. Biometric Applications Related to Human Beings: There Is Life beyond Security, Cognitive Computation, 2012, DOI 10.1007/s12559-012-9169-9
7. 8. K. López de Ipiña, J. B. Alonso, J. Solé-Casals, N. Barroso, M. Faundez, M. Ecay, C. Travieso, A. Ezeiza and A. Estanga Alzheimer Disease Diagnosis based on Automatic Spontaneous Speech Analysis, Proceedings of NCTA 2012. Barcelona, 2012
8. Alzheimer's Association.: http://www.alz.org/
9. Praat: doing Phonetics by Computer. Available online: www.fon.hum.uva.nl/praat
10. Solé J, Zaiats V., A Non-Linear VAD for Noisy Environment. Cognitive Computation. 2010; 2(3):191-198.
11. Higuchi T. Approach to an irregular time series on the basis of the fractal theory. Physica D 1988. 31277:283.
12. K. Lopez-de-Ipiña et al. Feature extraction approach based on Fractal Dimension for Spontaneous Speech modelling oriented to Alzheimer Disease diagnosis. LNAI, Proceedings of NOLISP, Mons Belgium, 2013.
13. David Martinus Johannes One-class classification, PhD Thesis,Tax. Technische Universiteit Delft, (2001)
14. Moya, M., Koch, M., and Hostetler, L. (1993). One-class classifier networks for target recognition applications. In Proceedings world congress on neural networks, pages 797–801, Portland, OR. International Neural Network Society, INNS.
15. Shehroz S. Khan, Michael G. Madden, A Survey of Recent Trends in One Class Classificatio Lecture Notes in Computer Science Volume 6206, 2010, pp 188-197
16. WEKA. Available online: http://www.cs.waikato.ac.nz/ml/weka/