Computers and Electrical Engineering 68 (2018) 310–321

Contents lists available at ScienceDirect

Computers and Electrical Engineering

journal homepage: www.elsevier.com/locate/compeleceng

# A new database of healthy and pathological voices☆

Ugo Cesari[a], Giuseppe De Pietro[b], Elio Marciano[c], Ciro Niri[d], Giovanna Sannino*[,b], Laura Verde[e]

[a] Department of Otorhinolaryngology, University Hospital (Policlinico) Federico II of Naples, Via S.Pansini, 5 Naples, Italy
[b] Institute of High Performance Computing and Networking (ICAR-CNR), Via Pietro Castellino, 111, Naples, Italy
[c] Area of Audiology, Department of Neurosciences, Reproductive and Odontostomatological Sciences, University of Naples Federico II, Via S.Pansini, 5, Naples, Italy
[d] Independent Doctor Surgeon Specialized in Audiology and Phoniatrics, Naples, Italy
[e] Department of Engineering, University of Naples Parthenope, Centro Direzionale di Napoli, Isola C4, Naples, Italy

ARTICLE INFO

*Keywords:*
Database
Voice disorders
Clinical study
Smart healthcare systems

ABSTRACT

In the era of Edge-of-Things computing for the accomplishment of smart healthcare systems, the availability of accurate and reliable databases is important to provide the right tools for researchers and business companies to design, develop and test new techniques, methodologies and/or algorithms to monitor or detect the patient's healthcare status. In this paper, the study and building of the VOice ICar fEDerico II (VOICED) database are presented, useful for anybody who needs voice signals in her/his research activities. It consists of 208 healthy and pathological voices collected during a clinical study performed following the guidelines of the medical SIFEL (Società Italiana di Foniatria e Logopedia) protocol and the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) 2013 Statement. For each subject, the database contains a recording of the vowel /a/ of five seconds in length, lifestyle information, the medical diagnosis, and the results of two specific medical questionnaires.

## 1. Introduction

Research activities, especially related to the accomplishment of smart healthcare systems, may require databases. In order to produce high-quality research results, three critical features must be examined. Firstly, the *quality of the database*, which helps guarantee that the results are accurate and generalizable. Researchers need data that is correctly labeled and similar to the real world or originates from the real world. Secondly, the *quantity of the database*, which ensures that there is sufficient data to train and validate approaches/tools, a fact which is especially important when utilizing artificial intelligence techniques. Finally, the *availability of data*, which is critical as it allows the research to commence and ensures reproducible results, helping to improve the state of the art. In fact, the reproduction and replication of determined scientific results provides the potential to improve the efficiency of smart healthcare systems and guide the development of new techniques [1].

In the scientific community, a comparison and improvement of results is only possible if identical input data sources are used. In the last few years, the importance of available databases has also been addressed by granting organizations, governments and other agencies. For example, *"The Obama Administration is committed to the proposition that citizens deserve easy access to the results of research their tax dollars have paid for"* [2]. Additionally, as also remarked by Penrose et al. [3] *"in the scientific method it is important that results*

☆ Reviews processed and recommended for publication to the Editor-in-Chief by Guest Editor Dr. M. M. Hassan.

* Corresponding author.
*E-mail addresses:* ugo.cesari@unina.it (U. Cesari), giuseppe.depietro@icar.cnr.it (G. De Pietro), marciano@unina.it (E. Marciano), ciro.niri@gmail.com (C. Niri), giovanna.sannino@icar.cnr.it (G. Sannino), laura.verde@uniparthenope.it (L. Verde).

https://doi.org/10.1016/j.compeleceng.2018.04.008
Received 21 December 2017; Received in revised form 13 April 2018; Accepted 13 April 2018

*be reproducible. An independent researcher should be able to repeat the experiment and achieve the same results. Most research has been done with private or irreproducible corpora generated by random searches on the WWW".*

In this work, we present a new reliable database of several voice recordings useful for the design, development, testing and evaluation of the performance of new algorithms and systems for different purposes, such as for example the detection of voice disorders through an analysis of the voice signal. Additionally, the proposed database could be embedded at the Edge level in a BodyCloud system [4,5] to support smart healthcare applications. The database has been realized thanks to a clinical study that we have performed during 2016 and 2017, in which all participants have been examined by medical experts. We have collected recordings of patients, some with vocal fold disorders and some without, conditions verified after the appropriate medical examination.

This paper describes in detail the design of the study and the construction of the **VO**ice **IC**ar f**ED**erico II (VOICED) database, realised by the "Institute of High Performance Computing and Networking of the National Research Council of Italy (ICAR-CNR)" and the Hospital University of Naples "Federico II".

The rest of the paper is organized as follows. In Section 2 we report some background information useful for the reader in the introduction of the problem of dysphonia and the state of the art related to the currently available databases specific to this particular disease. Section 3 presents the methods and procedures adopted to perform the study, collect the data, acquire the voice signals and save all information of interest. A discussion of the collected data is presented in Section 4, while some conclusions are provided in Section 5.

## 2. Background and related work

Nowadays, the use of technology in the healthcare sphere is increasing rapidly. Technology can play a crucial role in improving and promoting health and healthcare. It can, in fact, encourage healthy behaviours to prevent and reduce pathologies, facilitate continuous health monitoring and support the early detection of specific diseases, such as dysphonia.

Dysphonia is an alteration of voice production due to a morphological or functional alteration of the pneumo-articulatory apparatus. It is a disease that affects a great number of people, with about one third of adults suffering from this disorder at least once in their lifetime [6]. Unfortunately, people often underestimate the symptoms of dysphonia not taking the appropriate countermeasures to avoid a worsening of their health. A computer-based system can be an appropriate instrument to attract the interest of people through the screening and early detection of risk factors and specific symptoms of the pathology, as well as to support the diagnosis of the disease.

The assessment of voice disorders is a multidimensional and multiparametric investigation, composed of several clinical-instrumental analyses. A team instituted by the Italian Society of Phoniatrics and Logopaedics compiled a protocol, called the SIFEL (Società Italiana di Foniatria e Logopedia) protocol [7], of basic clinical-instrumental investigations for the subjective and objective assessment of dysphonia according to the guidelines recommended by the Committee of Phoniatrics of the European Laryngological Society. The protocol recommends several examinations to diagnose dysphonia, such as laryngoscopy, an invasive analysis useful to observe the morphological and functional alterations of the vocal tract, and the logopaedic evaluation of the pneumophonic-articulatory accordance, posture and muscular tensions. Another important examination is the acoustic analysis, thanks to which it is possible to estimate the state of health of the voice evaluating characteristic parameters extracted from the voice signal, which can be indexes of possible laryngeal alterations.

Among the indicated examinations, the acoustic analysis offers several advantages due to its non-invasive nature and its potential to provide quantitative data about the clinical state of the vocal tract, with an appropriate analysis time. Smart computer-based systems, able to perform a correct acoustic analysis, can offer a valid instrument for the automatic evaluation of dysphonia. The developed algorithms and/or systems must be tested with realistic data, and for their refinement the tests must be repeated and reproducible. For this reason, the available databases are essential resources for developers and evaluators of algorithms for the analysis of physiological data.

Currently, there are few databases of voice signals existing in literature and often they are not easily available. One of the most used databases is the Massachusetts Eye and Ear Infirmary (MEEI) Voice and Speech Laboratory [8]. It contains voice recordings of subjects, healthy or pathological, with a wide variety of organic, neuralgic, traumatic and psychogenic voice disorders. Even if it is considered as the basis of many studies of voice pathology assessment [9], it has some limitations, such as for example the fact that it is not freely available for the scientific community, and that different environments and sample frequencies are used to record the healthy and pathological voices.

Another frequently used database is the Saarbruecken Voice Database (SVD) [10–12], collected by the Institute of Phonetics of the University of Saarland in collaboration with the Department of Phoniatrics and Ear, Nose and Throat (ENT) at the Caritas clinic St. Theresia in Saarbruecken. The database contains a collection of voice recordings of subjects suffering from several pathologies, including both functional and organic ones.

The main features of the different voice databases currently on the market are summarised in Table 1.

Based on these considerations, VOICED could be considered as the second freely available voice database, soon to be accessible on physionet.org, and a new element in research on automatic voice disorder detection and classification. Although the proposed database consists of a smaller sample size than the other indicated databases, it is important to highlight that this smaller size might be useful in the detection of relevant findings in the planning and interpretation of testing studies. Additionally, there is no statistical reason why a result achieved in a trial including a wider sample size should be given more credibility than a trial including a smaller sample size. However, it should be remarked here, that VOICED is the first database that contains, in addition to the voice signals and their diagnosis, also information about life habits (voice use, smoking and alcohol abuse), the patient's character and previous or