Computers and Electrical Engineering 57 (2017) 257–265

Contents lists available at ScienceDirect

Computers and Electrical Engineering

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

An incremental method combining density clustering and support vector machines for voice pathology detection

Rimah Amami[a,*], Abir Smiti[b]

[a] Department of Electrical Engineering, National School of Engineering of Tunis, Tunis
[b] LARODEC, University of Tunis, Tunisia, Institut Supérieur de Gestion de Tunis, Tunis

ARTICLE INFO

Article history:
Received 14 January 2016
Revised 26 August 2016
Accepted 29 August 2016
Available online 6 September 2016

Keywords:
Voice pathology
Classification
SVM
DBSCAN
Incremental DBSCAN

ABSTRACT

Machine learning techniques are a valuable tool for discriminative classification. They have been applied to a diverse range of applications in speech processing, such as the analysis of pathological voices. We propose, in this paper, a novel policy, called Incremental DBSCAN-SVM in order to detect noises, to analyze and to classify pathological voice from normal voice. We use a modified density-based clustering algorithm named DBSCAN with an incremental learning in order to detect noisy samples. Then, the output model is submitted to Support Vector Machines (SVM) classifier with a Radial Basis Function (RBF) kernel to discriminate between normal and pathological voices. Our method has the ability to handle incremental and dynamic voices database which evolve over time. We support our approach with empirical evaluation using voices data set from the Massachusetts Eye and Ear Infirmary Voice and Speech Laboratory (MEEI) database to show the effectiveness of our method in terms of detection voice disorders.

1. Introduction

The recognition and classification of pathological voices are still a complex area of research in speech processing. Actually, the pathological voice refers to talking problems resulting from injure, abnormality of the speech organs, mental illness, hearing impairment, autism, paralysis or multiple disabilities. The presence of pathologies in the vocal folds affects the normal vibration pattern of the glottis and cause changes in voice quality. The classical means to spot voice pathology are prejudiced and not efficient. They are based essentially on the examination of the vocal folds and the study of the specialist that can generate different and confused evaluations. These techniques are costly, time consuming, distress for the patients and require many types of equipment.

In this paper we are interested in the classification between normal and pathological voices. There exists a wide fields of methods and systems which address this mission using a large range of databases and algorithms. Nevertheless, the majority is not forceful enough towards noisy data. Actually, the researchers in this area considered that pathological speech contains noisy data that affect badly the evaluation results. Based on this fact, we should first of all detect these noises from the pathological speech samples. Furthermore, the most practical applications are batch and not dynamic, they are not suitable for the frequently change information.

* Corresponding author.
E-mail addresses: rimah.amami@gmail.com, rimah.amami@yahoo.fr (R. Amami), smiti.abir@gmail.com (A. Smiti).

http://dx.doi.org/10.1016/j.compeleceng.2016.08.021
0045-7906/

To alleviate these potential issues, we propose, in this paper, a novel policy, named Incremental DBSCAN-SVM in order to analyze and to classify pathological voice from normal voice. Our method has the ability to handle noises and incremental voices database which evolve over time. We support our approach with empirical evaluation using voices data set to show the effectiveness of our method in terms of detection voice disorders.

The rest of this paper is organized as follows. In the next Section, we introduce some related works of speech pathology discrimination. In Section 3, we present two machine learning methods namely DBSCAN and SVM. Section 4 describes our proposed method. Experimental setup and results are given in Section 5. The paper concludes and presents some future works in Section 6.

2. Related works of speech pathology discrimination

In the last decade, machine learning techniques are often applied to medical applications. Among those practical applications, the evaluation the speech quality is the most complex and important tasks. This evaluation is performed in order to perceive the voice disorder and to build a system able to handle the voice pathologies. In this paper, we are interested in approaches, which detect and deal with the voice pathologies. T. Dubuisson et al. [1] proposed an analysis system which can present a discrimination between normal and pathological voices.

In their work, a discriminant correlation has been used to evaluate the performance within the MEEI database.

There are various classifiers used in the aim to classify the normal and pathological voices: Gaussian Mixture Model (GMM) is a supervised classification approach widely used for speaker detection. Fredouille et al. [2] used GMM to detect pathological and normal voices. Their study , 95% of normal samples are correctly classified and 18.3% of pathological samples are wrongly classied.

GMM for the detection of pathological have been, also, the subject of the study of Wang et al. [3] where one GMM maps all the pathological subjects. This approach provides a better performance than [2] with an error rate equal to 7% for normal samples and 1.4% for pathological samples.

The Support Vector Machines (SVM) was introduced first to solve pattern recognition problems. In the last years, SVM was successfully applied to many real-world applications. The authors in [4] proposed to use the SVM classifier in order to discriminate normal samples from pathological samples. The performance of SVM generates an average error rate equal to 05%.

In the automatic speech recognition area, Hidden Markov Model (HMM) appraoch is known to be the competitor of the most robust techniques in speech processing such as SVM. Thus, HMM were used, also, for the discrimination between normal and pathological speech samples.

Dibazar et al. presented an application of HMM using MEEI samples to detect pathological ones. Using HMM, the correct classification rates for the sustained vowels is equal to 99%.

Indeed, in the literature, various practical applications were done to deal with the discrimination between normal and pathological samples using batch data sets. In our knowledge, no successful empirical work was implemented to deal with incremental learning for this task. In this paper, we propose a method using an incremental learning at the first stage of our discrimination system in order to detect noises subjects using DBSCAN technique. Then, in the second part of our system, we use the model generated by the first step to classify the normal and pathological samples.

3. Abstract view of DBSCAN and SVM techniques

In this Section, we will focus on the DBSCAN clustering method and SVM technique since our paper deals with them:

3.1. DBSCAN clustering technique

The main idea of density based clustering method is to find regions of high density and low density, with high-density regions being separated from low-density regions. These approaches can make it easy to discover arbitrary clusters. More specifically, Density Based Spatial Clustering of Applications with Noise (DBSCAN) [5] is one of the most used and a typical density-based clustering algorithms (see Fig. 1).

The clustering algorithm DBSCAN shows several positive points: it can discover clusters of arbitrary shape, can distinguish noise, uses spatial access methods and is efficient even for large spatial databases.

DBSCAN is based on a main concept: density reachability (see Definition 2). This concept requires obligatory two important input parameters: *Eps* (see Definition 1) and *MinPts* the minimum number of points required to form a cluster.

Definition 1. The Eps neighborhood of a point x_i in database CB, denoted by $Eps(x_i)$, is defined by:
$Eps(x_i) = \{x_t \in CB | distance(x_i, x_t) \leq Eps\}$.