Physica A 471 (2017) 98–105

Contents lists available at ScienceDirect

Physica A

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

Parkinson's disease detection based on dysphonia measurements

Salim Lahmiri

Department of electrical engineering, École de Technologie Supérieure, Montreal, Canada

HIGHLIGHTS

- Seven categories of dysphonia measurements on Parkinson disease detection are investigated.
- Vocal fundamental frequency statistics yield to high accuracy.
- Accuracy improves when all categories of dysphonia measurements are employed.
- Refinement of dysphonia measurements allows achieving the highest accuracy.

ARTICLE INFO

Article history:
Received 8 August 2016
Received in revised form 24 October 2016
Available online 19 December 2016

Keywords:
Parkinson disease
Dysphonia measurements
Dysphonic symptoms
Voice disorder
Support vector machine
Classification

ABSTRACT

Assessing dysphonic symptoms is a noninvasive and effective approach to detect Parkinson's disease (PD) in patients. The main purpose of this study is to investigate the effect of different dysphonia measurements on PD detection by support vector machine (SVM). Seven categories of dysphonia measurements are considered. Experimental results from ten-fold cross-validation technique demonstrate that vocal fundamental frequency statistics yield the highest accuracy of 88% $\pm$ 0.04. When all dysphonia measurements are employed, the SVM classifier achieves 94% $\pm$ 0.03 accuracy. A refinement of the original patterns space by removing dysphonia measurements with similar variation across healthy and PD subjects allows achieving 97.03% $\pm$ 0.03 accuracy. The latter performance is larger than what is reported in the literature on the same dataset with ten-fold cross-validation technique. Finally, it was found that measures of ratio of noise to tonal components in the voice are the most suitable dysphonic symptoms to detect PD subjects as they achieve 99.64% $\pm$ 0.01 specificity. This finding is highly promising for understanding PD symptoms.

1. Introduction

Parkinson's disease (PD) is a neurodegenerative disorder that frequently affects elderly individuals and causes cognitive deficits. In this regard, a large amount of works have been proposed to study the characteristics related to PD symptoms; including analysis of hand movements [1–5], finger tremor [6], gait [7–14], and voice disorder [15–21]. The latter, is a frequent symptom as most of subjects affected with PD show vocal impairment [15], and also because voice measurement is a simple and noninvasive technique to adopt in detecting PD subjects [15]. Moreover, PD detection based on dysphonia measurements can be applied in telemedicine [16]. In general, there are several dysphonia measurements based on speech analysis which are helpful to detect PD in subjects [15]; including for instance the fundamental frequency or pitch of vocal oscillation, absolute sound pressure level, jitter, shimmer, noise-to-harmonics ratios, statistics from nonlinear dynamical

E-mail address: salim.lahmiri.1@ens.etsmtl.ca.

http://dx.doi.org/10.1016/j.physa.2016.12.009
0378-4371/© 2016 Published by Elsevier B.V.

systems theory, recurrence period density entropy, detrended fluctuation analysis, and pitch period entropy. The later was introduced by the authors in [15] as sensitive measure to observed changes in speech specific to PD.

The purpose of our paper is to study statistical characteristics of different dysphonia measurements, to investigate the effectiveness of each category of dysphonia measurements in PD detection, and how they can affect accuracy when employed jointly as main patterns. Indeed, whilst previous studies based on general voice disorders [15–21] have demonstrated the usefulness of dysphonic symptoms to detect PD in subjects, the problem of investigating the accuracy of different categories of dysphonia measurements in PD detection have not received attention. Therefore, the main goal of this study is to examine the influence of each category/type of dysphonia measurements on the performance of a predictive machine learning model; namely, the support vector machine (SVM) classifier [22] in distinguishing between healthy controls and unhealthy subjects affected with PD. Indeed, the SVM is constructed following statistical learning theory to execute the principle of structural risk minimization in order to perform a classification task. In this regard, it achieves global optimum. Furthermore, the SVM is appealing because of its ability to generalize well even with small training data [23]. Moreover, the SVM was successfully applied in PD assessment [24], detection of exudates in retina digital images [25], Alzheimer disease grading [26], cerebral arteriovenous malformations classification [27], time series prediction [28], Alzheimer detection [29], and glioma recognition [30].

The rest of our paper follows. Section 2 presents statistical tests, support vector machine classifier, and performance measures. Then, the obtained empirical results are provided and discussed in Section 3. Finally, Section 4 concludes our work.

2. Methods

Our methodology is briefly described as follows. First, two statistical test will be applied to dysphonia measurements to check whether variances are similar across healthy and PD subjects by using the Brown–Forsythe test [31]. In addition, in order to test whether dysphonia patterns exhibit similar distribution, the two-sample Kolmogorov–Smirnov (K–S) test [32,33] is applied to data from PD and healthy subjects. Indeed, the purpose is to explore variability and distribution differences in dysphonic symptoms across healthy and unhealthy patients. In other words, main statistical differences in characteristics of dysphonia measurements across PD and healthy subjects are examined in order to provide more insight on the statistical properties of the data. Second, three types of experiments will be conducted in our empirical work. First, the SVM classifier will be trained separately by each category of dysphonia measurements and its performance will be evaluated by computing accuracy, sensitivity, and specificity statistics. The goal is to find out which category of dysphonia measurements performs the best. In the second experiment, the SVM classifier will be trained with all dysphonia patterns to check whether performance can be improved. In an attempt to improve the SVM classifier performance, all dysphonia measurements with similar variability/distributions will be removed from the training and testing sets in the third experiment.

2.1. The Brown–Forsythe test

Since the standard F-test is highly sensitive to non-normality assumption in data, the Brown–Forsythe test [31] is chosen in this study to test the null hypothesis of equality of variance in dysphonia measurements between healthy and PD subjects. The Brown–Forsythe test performs a one-way analysis of variance on the absolute deviations from the median. Indeed, as Brown–Forsythe test uses the median, it is robust against non-normality in data. The Brown–Forsythe based F statistic is given by:

$$F = \frac{(N-p)\sum_{j=1}^{p} n_j \left(\tilde{z}_{\bullet j} - \tilde{z}_{\bullet\bullet}\right)^2}{(p-1)\sum_{j=1}^{p}\sum_{i=1}^{n_j}\left(\tilde{z}_{ij} - \tilde{z}_{\bullet j}\right)^2} \tag{1}$$

$$z_{ij} = \left|y_{ij} - \tilde{y}_j\right| \tag{2}$$

where N, p, n_j, and $\tilde{y}_j$ are respectively total number of observations, number of groups, number of observations in group j, and median of group j. The variables $\tilde{z}_{\bullet j}$ and $\tilde{z}_{\bullet\bullet}$ are respectively group mean and overall mean of z_{ij}. The decision rule to reject the null hypothesis of equal variance between samples is based on the comparison of the probability-value (p-value) with the statistical significance level of 5%. A p-value less than 5% yields to a rejection of the null hypothesis.

2.2. The Kolmogorov–Smirnov test

In our study, we employ the two-sample Kolmogorov–Smirnov (K–S) test [32,33] to verify whether or not a given type of dysphonia measurements in the PD patients significantly differ from those associated with healthy subjects. For instance, K–S test is applied to check whether data of two dysphonia measurements (PD versus healthy) are from the same