Biomedical Signal Processing and Control 31 (2017) 288–294

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

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

The impact of feature extraction for the classification of amyotrophic lateral sclerosis among neurodegenerative diseases and healthy subjects

Suleyman Bilgin*

Department of Electrical and Electronics Engineering, Faculty of Engineering, Akdeniz University, Antalya, Turkey

ARTICLE INFO

Article history:
Received 15 April 2016
Received in revised form 2 June 2016
Accepted 19 August 2016

Kewords:
Amyotrophic lateral sclerosis
Discrete wavelet transform
Linear discriminant analysis
Naïve bayes classification
Neurodegenerative diseases

ABSTRACT

Neurodegenerative diseases (NDD) including Amyotrophic Lateral Sclerosis (ALS), Parkinson's disease (PD) and Huntington disease (HD) can be defined as the degeneration in the structure of neurons in human body. It is mentioned in the related literature that NDD may cause various clinical symptoms disrupting gait dynamics. The characterization of gait analysis is crucial for early diagnosis, efficient treatment planning and monitoring of ALS progression and other NDD. The database consisting of 64 one-minute recordings of Compound Force Signal (CFS) obtained from 13 ALS, 15 PD, 20 HD and 16 healthy subjects was used in the study. CFS is the composition of force signals for both left and right feet of each subject during the gait. CFS was decomposed for determination of features using 6-level Discrete Wavelet Transform (DWT) with different wavelets in the study. The obtained features were evaluated using the means of 20-trials for five-fold cross-validation (FFCV) in Linear Discriminant Analysis (LDA) and Naïve Bayesian Classifier (NBC). As a result, D5 (4.6875–9.375 Hz) in all classifications, D4 (9.375–18.75 Hz) in ALS vs. PD, ALS vs. PD + HD and ALS vs. Co + PD + HD classifications while D2 (37.5–75 Hz) and D6 (2.3438–4.6875 Hz) in ALS vs. Co. and ALS vs. HD classifications were determined as the most significant frequency bands in CFS for discrimination of ALS among healthy and other NDD subjects in the end of the study.

1. Introduction

PD is a progressive disorder of the nervous system that causes stiffness or slowing down of the movements and HD is a hereditary disease causing the degeneration of nerve cells in the brain while it is known that ALS is a mortal and idiopathic NDD of motor system in the human body [1]. Although clinical analysis consisting of Electromyogram (EMG) analysis, blood tests, genetic tests and spinal tap are available to diagnose NDD, these types of analyses may result in the problems of time-consumption and misinterpretation for three neurodegenerative conditions at the first stage [2]. Therefore, the discrimination of ALS among other NDD and healthy subjects is so crucial for early diagnosis and making a timely intervention to help patients suffering from this fatal disease.

NDD may demonstrate the various clinical symptoms: ALS causes muscular atrophy and degeneration while PD brings about bradykinesia, tremor, rigidity and posture disturbances and HD causes cognitive and personality disorders. These irregular motions cause difficulties during the coordination of legs and affect the gait dynamics. Furthermore, many studies in the literature have focused on the gait dynamics analysis to investigate the effects of NDD on the locomotion of human [3–9]. Multi Scale Entropy (MSE), Wavelet analysis [10–12] and non-stationary analysis based on statistical and Time-Frequency (TF) methods are also available to observe the relationship between the gait dynamics and NDD in the literature [13,14]. Zheng et al. focused on the control vs. PD classification using statistical methods [13]. Sugavaneswaran et al. studied the control vs. ALS classification using TF based algorithm [14]. Liao et al. researched on the gait asymmetry of stance time fluctuations using multi-resolution entropy analysis [15]. Sang-Hong et al. investigated for the classification of PD using DWT [16]. Wu et al. focused on the classification of control vs. ALS using signal turn count based algorithm [17–19] and the classification between control and PD using Probability Density Function (PDF) [20]. Baratin et al. proposed a computer-aided automatic characterization method in the control vs. ALS classification using unbiased cross-validation strategy [21]. Sarbaz et al. researched on the control vs. HD classification using spectral analysis [22]. Xia et al.

* Corresponding author.
E-mail address: suleymanbilgin@akdeniz.edu.tr

http://dx.doi.org/10.1016/j.bspc.2016.08.016
1746-8094/

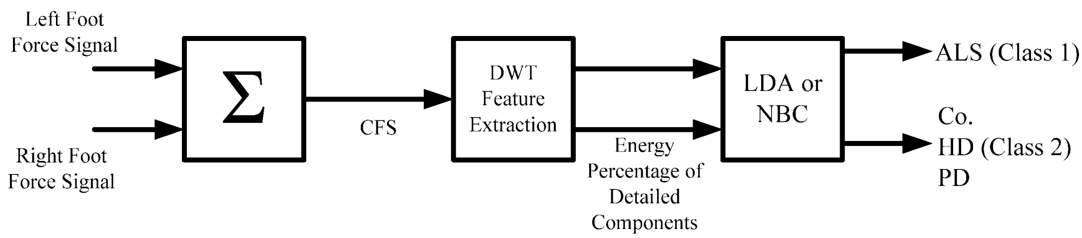

Fig. 1. The outline of study.

proposed a method for the binary classification in PD vs. Co. and HD vs. Co. [23]. Zeng et al. proposed a method for the classification of NDD [33] and Xia et al. have identified new features for the measurement of altered gait variability using standard deviation statistics and temporal structural characteristics [34]. The previous studies in the literature point out that the gait signals cause a fluctuation depending on the type of NDD [13,14,17–23,33,34]. These research studies in the literature focus on the analysis of stride interval time series and stance time fluctuations in five-minute gait records. Conversely, the present study ensures faster analysis and diagnosis by means of one-minute gait record analysis while others research on the analysis of five-minute gait records. The main objective of this study is to obtain critical frequency bands on the CFS for fast discrimination of ALS among other NDD and control subjects. Hence, we propose a method based on DWT, LDA and NBC in the present study.

2. Materials and methods

The outline of the study is shown in Fig. 1. In the research, first, the summation of the force signals recorded from the right and left feet was obtained from the subjects, respectively. These cumulative signals are mentioned as CFS in this study. Afterwards, the CFS was decomposed into reconstructed approximation and detailed components using DWT to obtain features [9]. These features were evaluated considering their classification accuracy performances using LDA and NBC in the last stage of the study. Finally, the crucial features were determined and reported in the study. The classes were selected as ALS (class 1) vs. control (class 2), ALS (class 1) vs. PD (class 2), ALS (class 1) vs. HD (class 2), ALS (class 1) vs. PD + HD (class 2) and ALS (class 1) vs. PD + HD + Co. (class 2) at the output of the algorithm during analysis.

2.1. Database description

The gait dynamics in NDD database on www.physionet.org website were used in this study [24]. This database includes 64 force signals recorded from right and left feet belonging to 13 (10 male, 3 female) ALS subjects (age range: 36–70), 15 (10 male and 5 female) PD subjects (age range: 44–79), 20 (6 male and 14 female) HD subjects (age range: 29–71) and 16 (2 male and 14 female) healthy control subjects (age range: 20–74) [5,25]. The raw data in the database was recorded using force-sensitive resistors placed under the right and left feet [25–29]. Also, the observation of the coordination of legs is crucial for gait analysis. Since the only one foot movement signal cannot be enough, CFS is a more useful signal for the modeling of the coordination between two feet in the study. The 6-s episodes of signals recorded from the left foot and the right foot signals and CFS are shown in Fig. 2. All of the raw data were sampled at 300 Hz sampling frequency. There are several limitations [6] in the database: 1-The groups do not have proper matching with regard to gender and age. Nevertheless, further research studies are needed to execute the effects of age and gender on the gait rhythm dynamics, it is explained that NDD do not usually vary across gender and age groups on the gait dynamics in the literature [5–7,30]. 2-Subjects were also not eligibly matched considering their height. This limitation is related to increasing of stride time fluctuation in the gait dynamics but it may not affect the CFS patterns and analysis considering stride time fluctuation. 3-The relatively small number of subjects within each group causes the poorly matching of clustering during classification and evaluation process. In this case, cross-validation processing parameters must be selected in a balanced manner for validation of classification performance. Therefore, FFCV was used for the measurement of performance in the study. 4-The database in the literature includes

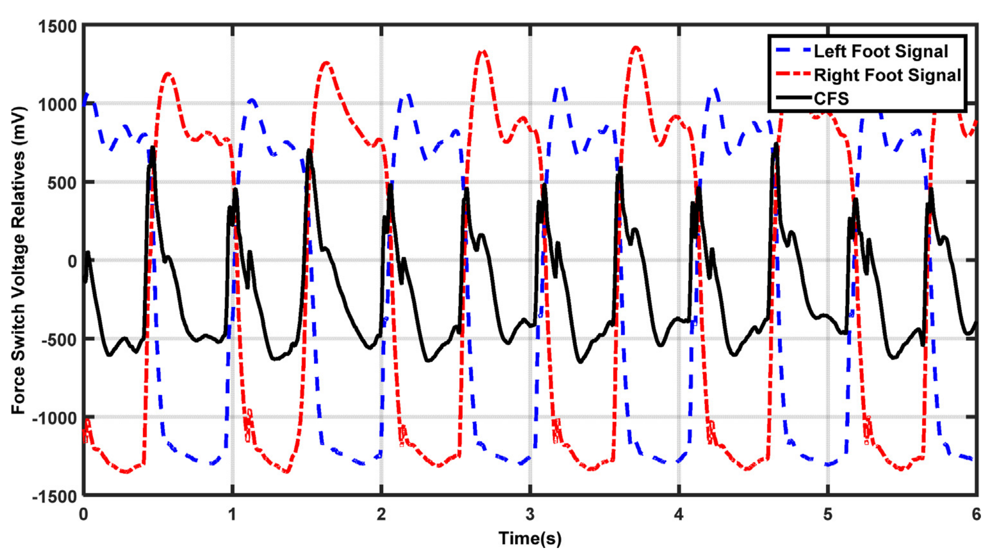

Fig. 2. The Force signal of the left foot, the right foot and CFS.