



HAL
open science

Detecting Parkinson's Disease with Sustained Phonation and Speech Signals using Machine Learning Techniques

Jefferson Almeida, Pedro Pedrosa Rebouças Filho, Tiago Carneiro, Wei Wei, Robertas Damaševičius, Rytis Maskeliūnas, Victor Hugo Costa de Albuquerque

► To cite this version:

Jefferson Almeida, Pedro Pedrosa Rebouças Filho, Tiago Carneiro, Wei Wei, Robertas Damaševičius, et al. Detecting Parkinson's Disease with Sustained Phonation and Speech Signals using Machine Learning Techniques. *Pattern Recognition Letters*, 2019, 125, pp.55-62. 10.1016/j.patrec.2019.04.005 . hal-02380596

HAL Id: hal-02380596

<https://hal.science/hal-02380596v1>

Submitted on 26 Nov 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Detecting Parkinson's Disease with Sustained Phonation and Speech Signals using Machine Learning Techniques

Jefferson S. Almeida^a, Pedro P. Rebouças Filho^{a,b}, Tiago Carneiro^{a,e}, Wei Wei^c, Robertas Damaševičius^d, Rytis Maskeliūnas^d, Victor Hugo C. de Albuquerque^{b,**}

^aInstituto Federal de Educação, Ciência e Tecnologia do Ceará, Maracanaú, Brazil

^bPrograma de Pós-Graduação em Informática Aplicada, Universidade de Fortaleza, Fortaleza, Ceará, Brazil

^cSchool of Computer Science and Engineering, Xi'an University of Technology, Xi'an 710048, China

^dFaculty of Informatics, Kaunas University of Technology, Kaunas 51368, Lithuania

^eInstitut National de Recherche en Informatique et en Automatique (INRIA) Lille - Nord Europe, Villeneuve d'Ascq, France

ABSTRACT

This study investigates the processing of voice signals for detecting Parkinson's disease. This disease is one of the neurological disorders that affect people in the world most. The approach evaluates the use of eighteen feature extraction techniques and four machine learning methods to classify data obtained from sustained phonation and speech tasks. Phonation relates to the vowel /a/ voicing task and speech to the pronunciation of a short sentence in Lithuanian language. The audio tasks were recorded using two microphone channels from acoustic cardioid (AC) and a smartphone (SP), thus allowing to evaluate the performance for different types of microphones. Five metrics were employed to analyze the classification performance: Equal Error Rate (EER) and Area Under Curve (AUC) measures from Detection Error Tradeoff (DET) and Receiver Operating Characteristic curves, Accuracy, Specificity, and Sensitivity. We compare this approach with other approaches that use the same data set. We show that the task of phonation was more efficient than speech tasks in the detection of disease. The best performance for the AC channel achieved an accuracy of 94.55%, AUC 0.87, and EER 19.01%. When using the SP channel, we have achieved an accuracy of 92.94%, AUC 0.92, and EER 14.15%.

1. Introduction

Recently, the World Health Organization (WHO) described neurological disorders as one of the most significant threats to public health. Among the most common disorders are Parkinson's disease (PD), stroke, multiple sclerosis, headache disorders, dementia, epilepsy and other diseases. At this time it is estimated that 16 out of 60 people are suffering due to some neurological disease. Parkinson's disease (PD), initially called shaking palsy, first was described by James Parkinson in 1817 [1]. This disease is neurological and degenerative, characterized by deficiency dopamine production in the midbrain region called black substance. Figure 1 brings an Illustrative diagram showing a region of the brain affected by Parkinson's disease.

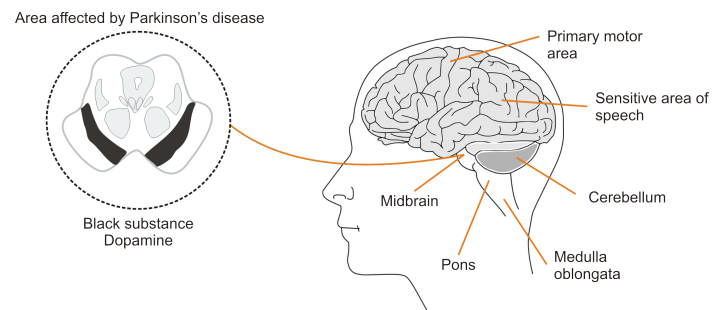


Fig. 1. Illustrative diagram showing a region of the brain affected by Parkinson's disease. This condition affects the production of the neurotransmitter dopamine in the midbrain region. Own representation based on Parkinson's Disease: Non-Motor and Non-Dopaminergic Features.

The deficiency of dopamine is caused by the degeneration of neurons in black substance and may be related to age, family history, skull trauma and even contact with some pesticides. Dopamine is an essential neurotransmitter, because it assists

**Corresponding author:

e-mail: victor.albuquerque@unifor.br (Victor Hugo C. de Albuquerque)

involuntary movements and its reduction below normal levels causes symptoms, which can be divided into two groups: motor and non-motor. The motor group manifests by involuntary tremulous motions, lessened muscular power that makes it difficult for performing simple activities like buttoning a shirt, or putting on a shoe. In addition to muscle rigidity due to the lack of dopamine in the body, the muscles fail to receive the signals to relax. Other symptoms are the loss of facial expression, changes in speech and handwriting. The non-motor group may also be present with dementia, depression, anxiety, altered sleep, and slow thinking.

In all, three steps are required to diagnose the disease. The first one contemplates the clinical consultation. After the symptoms are confirmed, the next step is performing drug therapy, in which the patient with suspected PD undergoes dopamine supplementation. If there are improvements, there is a high probability of the individual having the disease. However, a third step is required to confirm the condition, in case of doubts. In this stage, laboratory tests are performed, which may not be accessible to all people, especially in the developing world. The loss of dopaminergic neurons evolves rapidly, reaching about 50% at the moment of clinical diagnosis, and can quickly progress over five years [2]. Often the patient faces long queues awaiting examination. Therefore, having more tests available in short time are of great importance, as they allow streamlining of the diagnostic and treatment process. The voice and handwriting laboratory tests are among the most accessible ones. These tests are relevant because they are applied based on the first symptoms that arise in the individual with suspected PD. Handwriting tests can be conducted on paper, aided by computer vision systems [3, 4]. Finger touch tests can be employed to evaluate tremor symptoms, such as for patients with Huntington's disease, which has similar symptoms to PD [5]. The voice tests contemplate sustained phonation and speech signals, and it is a promising area of research.

The current research considers the database proposed by Vaiciukynas *et al.* [6] with the same. The main contribution of our approach evaluates the performance with the use of other classifiers and processing audio without fusion of feature sets. The clinical impact is the possibility for the doctor to use the PD classification implemented on a smartphone. In the area of Computer Science the impact is related to the evaluation of classical classification methods, which do not require relatively high processing capacity. We performed a complex study with several feature extractors and classifiers not yet used with this database.

The remainder of this paper is organized as follows. Section 2 overviews the related works in literature. Section 3 describes the dataset. In Section 4, the methodology is presented. The results are shown in Section 5. In section 6, a discussion is presented. Section 7 states the conclusion and future works.

2. Related works

In 1998, it was concluded that the voice is the primary deficit, most often affected in the early stages of PD when compared to articulation and fluency [7]. Studies have shown that PD

can be detected in its incipient phase [8]. Statistical techniques were used to evaluate the residual signal auto-correlation in Parkinson's disease patients and significant differences in pitches amplitude between patients and the healthy group was identified [9].

A measure of dysphonia was proposed for discriminating people with PD and healthy people from variations in voice frequency [10]. In 2013, sustained vowels, words, and sentences from a set of speaking tasks were found to carry PD-discriminative information that can be utilized using machine learning tools [11]. The works using voice signal continued in the following years with the contribution of new feature extractors and new data sets [12, 13, 14]. The most notable trends are: the use of smartphone technology for the recording of speech samples in everyday life, capturing signs of speech impairment in persons [15], and the development of software for analysing pathological speech signals, taking into account phonation, articulation, prosody, and intelligibility.

Recently, other approaches have been investigated for the detection of PD. [16] studied speech recordings using feature extracted from several dimensions of speech, including phonation, articulation and other human characteristics. To improve the diagnosis of Parkinson's disease, [17] introduced an improved and optimized version of the Crow Search Algorithm. Convolutional Neural Networks (CNN) were applied to hand-written exams to investigate their performance of PD detection, achieving 95% of accuracy [18].

[19] considered articulatory features for plosive consonants using a database composed of 27 individuals diagnosed with PD and 27 healthy controls. They used A sequential backward feature selection to find an optimal feature subset, and Support Vector Machine (SVM) for classification. [20] trained XGBoost models to predict the changes in clinical scores of 51 PD patient using their phonation data. [21] employed chaotic bacterial foraging optimization (CBFO) with an enhanced fuzzy k-nearest neighbor (FKNN) classifier for early diagnostics of PD using vocal recording data. [22] proposed Modified Grey Wolf Optimization (MGWO) as a search strategy for feature selection, and Random forest, k-nearest neighbor classifier and decision tree for classification. [23] proposed to use weights derived from a Multi-Layer Perceptron (MLP) for feature selection, while a Lagrangian Support Vector Machine (LSVM) was used for classification, aiming for early diagnosis of PD in a medical setting. [24] used feature selection procedures based on Pearson's and Kendall's correlation coefficients, principal component analysis (PCA), and self-organizing maps (SOM), and multiple feed-forward artificial neural networks (ANNs) for classification, reaching an accuracy of 86.47%.

In another approach, the cuttlefish algorithm was optimized for diagnosis of disease, but with an accuracy of approximately 94% [25]. [26] studied new classification schemes using machine learning and different phonetic units. CNN continue to be used to classify the disease into voice signals, but the results still do not exceed 90% of accuracy [27]. For more works related to PD, the reader can check [28].

3. Dataset

The dataset used in this work [6] was generated from two vocal tasks and organized into two different modalities: *Phonation* and *Speech*. The audio tasks were recorded using two channels from acoustic cardioid (AKG Perception 220, frequency range 20-20000Hz) and a smartphone (an internal microphone of Samsung Galaxy Note 3). The microphone was positioned at about 10 cm distance from the subject's source of audio. The audio format used was mono PCM wav (16 bits at 44.1 kHz sampling rate).

The Phonation modality contains a sustained voicing of vowel /a/ vocalized at a loudness level and comfortable pitch. The task duration was about 5s and was repeated 3 times. In turn, the Speech modality contains a pronunciation of a phrase in the Lithuanian language by a native speaker, "turėjo senelė žilą oželį". Praat software toolbox was used to split in audio speech, generating voiced and unvoiced parts. The database contains data recorded from 99 subjects, both male and female, for acoustic cardioid (AC) and smartphone (SP) recordings. Mean age of healthy control (HC) subjects was 41.8 years, while mean age of PD patients was 61.5 years. The health state of PD patients was evaluated at stages 1 to 2.5 according to Hoehn and Yahr [29], indicating early-to-mild stages of the disease. The PD patients had symptoms of hypokinetic dysarthria, manifesting in decrease in phonation time, reduced breath support, increased acoustic noise, and harsh breathy voice.

More information on the dataset is shown in Table 1.

Table 1. Number of subjects in the dataset used in this work. In the Phonation modality, the voice test was repeat 3 times for each subject (see the number in parentheses).

	Phonation		Speech	
	AC	SP	AC	SP
HC male	11 (33)	11 (33)	11	11
HC female	24 (72)	24 (72)	24	24
PD male	30 (89)	30 (90)	29	30
PD female	34 (101)	34 (102)	34	34
Total	99 (295)	99 (297)	98	99

Notes: PD - Parkinson's disease patient, HC - healthy control subject, Microphone: AC - acoustic cardioid, SP - smart phone.

4. Methodology

For detecting PD, the proposed approach is based on motor symptoms explicitly related to voice. Figure 2 shows the diagram describing the steps followed in this approach.

The voice examination begins in step 1, with two vocal tasks (see Figure 2). Vocal tasks include phonation and speech; respectively voicing of vowel and pronunciation of sentence. Step 2 represents the voice signal acquisition equipment. Thus, we can evaluate the performance of the classifiers with audio data recorded by AC and SP. Step 3 is preprocessing audio signals acquired in the first step. Audio speech is partitioned into two audio files, voiced and unvoiced, with Praat Software Toolbox.

In step 4, the raw data was processed, and the feature extraction was applied, which generates other files for each feature set and modalities. In all, 144 extraction files were generated. This approach uses a hold-out technique to evaluate the

database and four classifiers, as shown in step 5. Ten iterations were used. The hyper-parameters for the k-Nearest Neighbour (kNN), Multi-Layer Perceptron (MLP), Optimum-Path Forest (OPF) and Support Vector Machine (SVM) were optimized using a hold-out technique.

kNN, MLP and SVM were trained using the toolbox (OpenCV 2.49) and default settings for the hyper-parameters. Three different function kernels were used for SVM classifier.

MLP was trained using the back-propagation algorithm and the range of hidden units varies from 1 to 50 with grid search. In turn, OPF was trained with various distance metrics. In step 6, a selection by Equal Error Rate (EER) and Accuracy metrics was applied to evaluate the results and choose the best combination of extractor and classifier.

Receiver Operating Characteristic (ROC) and Detection Error Tradeoff (DET) curves were applied to obtain Area Under Curve (AUC) and Equal Error Rate (EER). EER was used for two reasons: it allows identifying the combination of extractor and classifier that best presents a balance between the false negative rate (FNR) and the false positive rate (FPR). Therefore, the ERR metric is used as the standard metric to compare the results of the present work with the results of [6].

5. Results

Here we present the results achieved by the machine learning methods on *eighteen* data sets generated by different feature extraction techniques. Based on the data, we select the best configuration of each classifier.

The intent is to obtain a classifier that presents the best accuracy (ACC). We observe that in most cases selection by accuracy is characterized by a greater balance in specificity (Spe) and sensitivity (Sen). Generally, by selecting in this way we can get the best classifier performance in terms of AUC and EER. Of these, we chose the one that obtained the best performance, as shown in Table 2 (the best results are highlighted in bold font style).

To perform the statistical analysis of the results, we applied N-way ANOVA, Freidman/Kruskal-Wallis and Nemenyi tests to the EER results. There was a significant effect at the $p < .05$ level found for audio types [$F(2, 649) = 28.47, p < 0.001$] and classifiers [$F(13, 649) = 16.56, p < 0.001$], but no significant effect for channel modalities [$F(1, 649) = 0.03, p = 0.859$] and feature extractors [$F(5, 649) = 0.43, p < 0.826$]. The results were confirmed by the Nemenyi post-hoc tests, which found no significant difference in EER between the AC and SP channel modalities ($p = 0.78, CD = 0.06$), and between feature extractors ($p = 0.914, CD = 2.49$), but found significant difference between Phonation (P), Speech (S), Unvoiced part (U) and Voiced part (V) audio types ($p < 0.001, CD = 0.21$) and classifiers ($p = 0.01, CD = 1.65$). The results of the Nemenyi tests are presented in Figure 3. In further analysis, we focused on the AC channel data with P audio only, and on the SP channel data with P audio only as well. In this case, the best feature/classifier pairs were TS/SR for AC/P and EL/K1 for SP/P. However, the results were not significantly different (see Figure 4).

We have obtained satisfactory results when compared with Vaiciukynas's approach [6]. The compared results refer to the

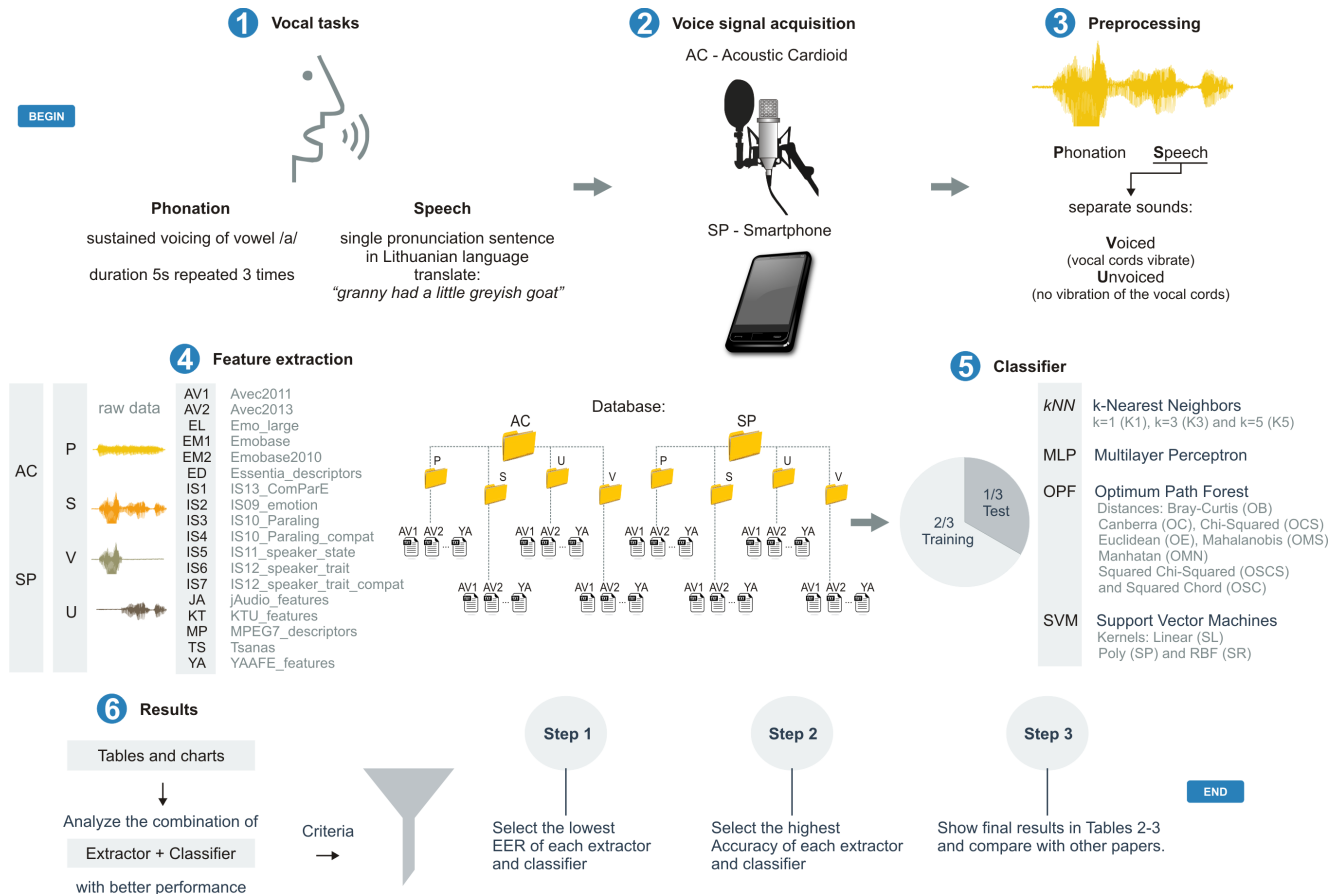


Fig. 2. Diagram of the proposed approach has 6 steps: (1) The voice examination begins; (2) Voice signal acquisitions equipment; (3) Preprocessing audio signals; (4) Feature extraction with eighteen feature sets; (5) Hold-out validation to training the data set using four classifiers; and (6) Evaluation of the results and selection of the best combination of feature extractor and classifier.

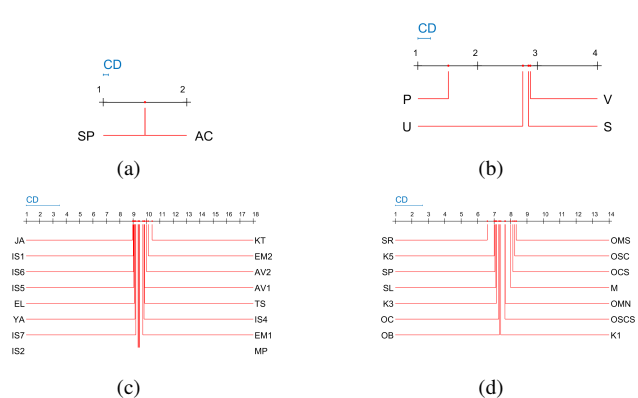


Fig. 3. The results of Nemenyi test: (a) channels, (b) modalities, (c) features and (d) classifiers

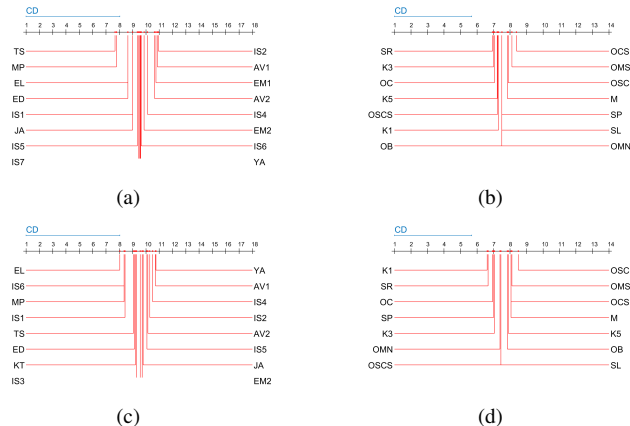


Fig. 4. The results of Nemenyi test: (a) features for AC/P, (b) classifiers for AC/P, (c) features for SP/P, and (d) classifiers for SP/P

data obtained without combining different audio types. In our approach, the audios were classified individually. When analyzing Table 2, our approach has selected two classifiers with best performance. In the AC modality, using Phonation (P) audio type, YA extractor and K1 classifier allowed to achieve an accuracy of 94.55% and EER of 19.01%. Vaiciukynas [6] used these extractors with the Random Forest (RF) classifier and obtained an EER of 20.30% with (ED) extractor. In the SP modal-

ity, we highlight the audio type (P) and (KT) extractor and (K1) classifier with 92.94% of accuracy and EER 14.15%. In AC and SP modality, our approach obtained the best performance. Moreover, our approach is simpler, because we do not use a fusion of all feature sets, so we obtained good performance with individual audio type classifier. Using fusion of all feature sets Vaiciukynas's approach obtained EER of 19.27% for the AC

and 23.00% for the SP channel.

As expected and mentioned in the literature, the k NN classifier is characterized by a long time in the test and shorter time of training. However, the k NN classifier implemented in the OpenCV library uses parallel processing techniques, making it relatively faster than other implementations. When comparing with the times obtained with different classifiers in Table 2, we can see that the k NN classifier obtained a test time of 34.2 ms, which is an acceptable and sufficient time for a medical examination.

The quality of the audio did not significantly influence the results. The difference between the AC and SP channels did not exceed 5%. Thus, the use of a professional microphone is optional, and the internal microphone of the smart-phone can also be used. We observed that for both modalities the task of phonation had more relevance and allowed to reach better results. There were differences only in the feature extractors, as shown in Table 2. For the AC channel, the best feature extractor of characteristics was the YA, whereas for the SP channel it was the KT. It is true that an acoustic and controlled environment is required to perform this test and obtain good results as noise pollution can interfere significantly.

The results showed that it is possible to use this approach with a smart-phone collected data since it has many advantages. Among these advantages, we can highlight the robustness of the tool to aid in the medical diagnosis of PD. In addition, the advantage of a smart-phone is portability and accessibility. These results open new horizons for the development of embedded software focused on detecting diseases that can be identified by voice related characteristics.

The data analyzed has a relatively large dimension. For dimensional reduction, we used t-distributed Stochastic Neighbor Embedding (t-SNE). It's a tool to view high-dimensional data. In this case, we applied t-SNE to view the samples of the YA and KT feature sets marked in bold type font in Table 2. One can see in Figure 5 the resulting t-SNE visualization presents the clusters of PD and HC. Note that we used t-SNE only used for visualization.

The confusion matrix of classification results using best performing methods is shown in Table 3. We have observed that in the AC modality, the number of false positives and false negatives were presented in a balanced way and this could be an attractive characteristic. While in the SP mode, the number of false positives was relatively higher. This means that a larger number of healthy subjects was confused with subjects having PD. However, the number of false negatives was smaller. It means that the probability of an individual with PD being diagnosed as healthy is lower.

To evaluate, if difference between confusion matrices obtained for database of AC by P/YA classifier K1, and for database of SP obtained by P/KT classifier K1 is significant, we have used the McNemar's test, which gives the statistical significance of differences between the accuracies achieved by different classifiers. This test is based on the standardized normal test statistic $|Z|$. At the 5% level of significance, the difference of accuracies between the classifiers is statistically significant if $|Z| > 1.96$. In this case, the difference is $Z = 2.92$, at

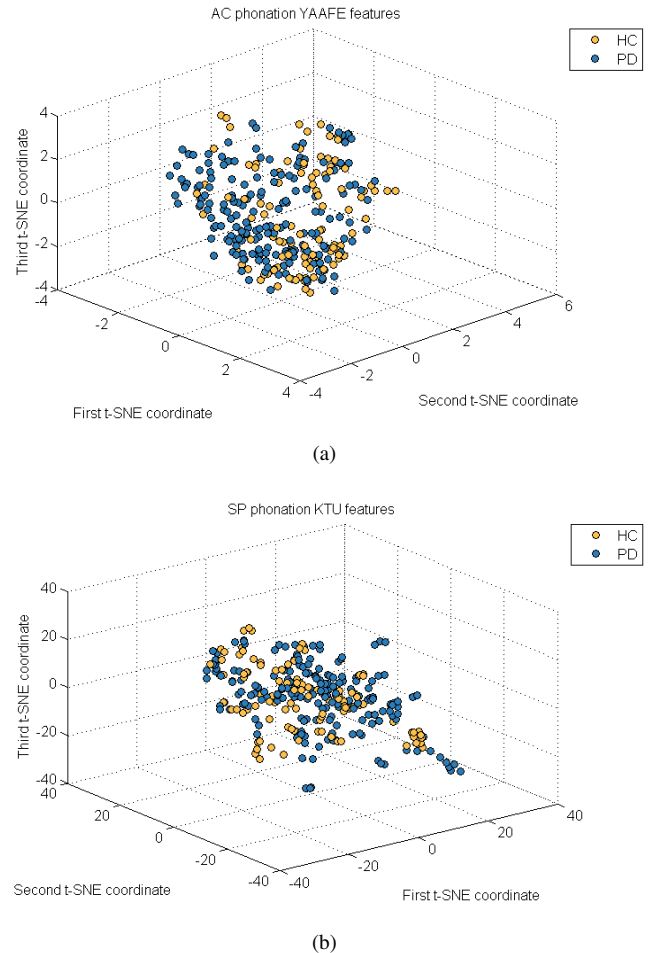


Fig. 5. Visualization of data by the t-SNE. (a) AC - acoustic cardioid, phonation and YA features. (b) SP-smart phone, phonation and KT features. Recording from: PD - Parkinson's disease patient and HC - healthy control subject.

$$p = 0.0035.$$

6. Discussion

The primary goal of the present work is to accelerate the process of diagnosis of PD, thus, allowing patients to start treatment as soon as possible. However, we can not tolerate a misdiagnosis. That is why the academic community has been looking for constant evolution in this area of research. Specifically, these approaches aid the task of medical diagnosis.

Our approach is essential to PD diagnosis because it promises a more accessible examination in comparison to other types of examination. However, there is the concern in correctly diagnosing the number of people affected by the disease. In dealing with the computer-aided diagnosis, some errors need to be minimized or preferably extinct, which can be mentioned: (1) patients with PD classified as healthy, and (2) healthy patients classified as PD. The most serious is the situation (1), because in this case, the individual will have the chance of treatment delayed until a new diagnosis is made and the disease is confirmed. The situation (2) is also serious, because it also

characterizes the incorrect diagnosis. In this sense, the selection criteria were adopted to select the extractor and classifier combination that presents a balance between situations (1) and (2) previously exemplified to minimize the occurrence of these cases. These occurrences can be identified by analyzing the confusion matrix. This equilibrium was achieved by performing the selection using the criteria shown in Figure 2 in step 6 of our approach.

In addition, there is the problematic of the classification system covering a single language. In order to carry out the training of learning algorithms in a specific language, we would need new databases of speech in this language. The development of a multi-lingual disease classification system is still a major challenge.

Having unbalanced data is very common in general, but in this case had more samples with the disease than healthy control samples. Since an unbalanced dataset will influence the prediction model for the most common class, I do not see a significant problem with the dataset being unbalanced in this regard. We used 75% of the samples for training and 25% for the test using the hold-out technique. Note that we did not use cross-validation because of the large data dimension, which would take more time for training.

7. Conclusions

We presented an approach for Parkinson's disease detection using consolidated feature set extraction and machine learning techniques. We applied 18 feature sets and 14 classifiers on voice dataset and evaluated the results in terms of ERR, and Accuracy. The results of statistical testing on ERR results using N-way ANOVA and Nemenyi test showed that there is no significant statistical difference between the audio cardioid (AC) and smart-phone (SP) microphones, thus supporting the idea that PD can be detected using smart-phone. The Phonation (P) audio has provided significantly better results than Speech (S), Voiced part (V) and Unvoiced part (U) audio.

The best individual feature set and classifier was Yaffe (YA) feature set with 1-nearest neighbour (K1) classifier when using phonation modality of the audio cardioid (AC) microphone, and KTU (KT) feature set with K1 when using phonation (P) modality of the smartphone (SP) microphone, achieving an accuracy of 94.55% and 92.94%, AUC of 87.84% and 92.40%, EER of 19.01% and 14.15%, respectively. Our approach obtained different results from Vaiciukynas's approach [6] and proved to be relatively better. The performance between the AC and SP modalities showed to be coherent in terms of accuracy.

Our main contribution is the use of other classifiers for the detection of PD and the investigation of an approach without the use of audio fusion. The clinical impact is the possibility for the doctor to use the PD classification method application installed on a smart-phone. In the area of Computer Science, the impact is related to the evaluation of classical classification methods. We performed a complex study with several features sets and classifiers not yet used with this data set. Based on the results shown, we can state that the phonation task is the most suitable for use, besides being the one that allowed to obtain the best performance.

In future work, we will apply cross-validation to solve the problem of the unbalanced dataset. Another future work will include testing and optimization of feature sets performing the pruning of characteristics that do not contribute much to the detection of PD. Besides, we will propose a new feature extraction tool for detection of PD.

8. Ethical statement

The study protocol related to this data set was approved by the Regional Kaunas Bioethics Committee (P2-24 / 2013). Written informed consents were obtained from study participants, patient identifiers were removed to ensure anonymity.

9. Acknowledgment

JSA received support from the Cearense Foundation for the Support of Scientific and Technological Development (FUN-CAP), and the Federal Institute of Education, Science and Technology of Ceará, in Brazil. VHCA received support from the Brazilian National Council for Research and Development (CNPq, Grants No.304315/2017-6 and 430274/2018-1). WW received support from the National key R&D Program of China under Grant NO. 2018YFB0203900 and also from the Key Research and Development Program of Shaanxi Province (No.2018ZDXM-GY-36).

References

- [1] J. Parkinson, An essay on the shaking palsy, *The Journal of Neuropsychiatry and Clinical Neurosciences* 14 (2) (2002) 223–236, pMID: 11983801 (2002). doi:10.1176/jnp.14.2.223.
- [2] J. M. Fearnley, A. J. Lees, Ageing and parkinson's disease: substantia nigra regional selectivity, *Brain* 114 (5) (1991) 2283–2301 (1991). doi:10.1093/brain/114.5.2283.
- [3] J. W. M. de Souza, S. S. A. Alves, E. de S. Rebouças, J. S. Almeida, P. P. R. Filho, A new approach to diagnose parkinson's disease using a structural cooccurrence matrix for a similarity analysis, *Computational Intelligence and Neuroscience* 18 (2018) 8 (2018). doi:10.1155/2018/7613282.
- [4] C. R. Pereira, D. R. Pereira, F. A. Silva, J. P. Masieiro, S. A. Weber, C. Hook, J. P. Papa, A new computer vision-based approach to aid the diagnosis of parkinson's disease, *Computer Methods and Programs in Biomedicine* 136 (2016) 79 – 88 (2016). doi:10.1016/j.cmpb.2016.08.005.
- [5] A. Lauraitis, R. Maskeliūnas, R. Damaševičius, ANN and fuzzy logic based model to evaluate huntington disease symptoms, *Journal of Healthcare Engineering* 2018 (2018) 1–10 (2018). doi:10.1155/2018/4581272.
- [6] E. Vaiciukynas, A. Verikas, A. Gelzinis, M. Bacauskiene, Detecting parkinson's disease from sustained phonation and speech signals, *PLOS ONE* 12 (10) (2017) 1–16 (10 2017). doi:10.1371/journal.pone.0185613.
- [7] A. K. Ho, R. Iansek, C. Marigliani, J. L. Bradshaw, S. Gates, Speech impairment in a large sample of patients with parkinson's disease, *Behavioural Neurology* 11 (3) (1999) 131–137 (1999). doi:10.1155/1999/327643.
- [8] B. Harel, M. Cannizzaro, P. J. Snyder, Variability in fundamental frequency during speech in prodromal and incipient parkinsons disease: A longitudinal case study, *Brain and Cognition* 56 (1) (2004) 24–29 (oct 2004). doi:10.1016/j.bandc.2004.05.002.
- [9] J. C. Pereira, A. O. Schelp, A. N. Montagnoli, A. R. Gatto, A. A. Spadotto, L. R. de Carvalho, Residual signal auto-correlation to evaluate speech in parkinson's disease patients, *Arquivos de Neuro-Psiquiatria* 64 (4) (2006) 912–915 (dec 2006). doi:10.1590/s0004-282x2006000600004.

Table 2. Results obtained after applying the selection criteria. The font in bold represents the best result in each modality. AC - Professional microphone, SP - Smartphone; Audio types: P - Phonation, S - Speech, U - Unvoiced, V - Voiced; Feature sets and Classifiers labels: see Figure 2.

Modalities	Audio	Feature	Classifier	Acc (%)	Spe (%)	Sen (%)	AUC	EER (%)	Train Time (s)	Test Time (s)
AC Channel	P	YA	K1	94.55±3.33	94.26±3.96	94.55±3.32	0.87	19.01	0.002±0.005	0.034±0.033
	S	IS4	SR	72.65±4.76	55.16±10.77	72.65±4.75	0.84	22.61	2.900±0.035	0.008±0.007
	U	IS6	SR	76.47±5.89	65.57±7.75	76.47±5.88	0.78	27.42	4.557±0.292	0.008±0.007
	V	EM1	SR	68.53±6.18	49.88±10.70	68.53±6.17	0.78	27.64	0.656±0.012	0.001±0.000
SP Channel	P	KT	K1	92.94±2.55	89.21±3.62	92.94±2.54	0.92	14.15	0.002±0.004	0.025±0.024
	S	KT	SP	71.18±4.52	50.94±6.91	71.18±4.51	0.87	19.63	7.720±0.330	0.0001±0.000
	U	MP	K1	72.94±3.91	61.38±3.30	72.94±3.90	0.74	30.84	0.001±0.000	0.003±0.003
	V	EM1	SP	65.00±4.46	41.14±7.92	65.00±4.45	0.80	25.97	6.094±0.151	0.001±0.000

Table 3. Confusion matrix of classification results using best performing methods: (a) AC obtained by P/YA classifier K1 and (b) SP obtained by P/KT classifier K1.

AC Performance				SP Performance			
Actual class	HC	Predicted class		Actual class	HC	Predicted class	
		HC	PD			HC	PD
		33.80	2.20			30.50	5.50
PD	3.30	61.70	PD	1.70	64.30		

Notes: PD - Parkinson's disease patient; HC - Healthy control subject.

- [10] M. Little, P. McSharry, E. Hunter, J. Spielman, L. Ramig, Suitability of dysphonia measurements for telemonitoring of parkinsons disease, *IEEE Transactions on Biomedical Engineering* 56 (4) (2009) 1015–1022 (apr 2009). doi:10.1109/tbme.2008.2005954.
- [11] B. E. Sakar, M. Isenkul, C. O. Sakar, A. Sertbas, F. Gurgen, S. Delil, H. Apaydin, O. Kursun, Collection and analysis of a parkinson speech dataset with multiple types of sound recordings, *IEEE Journal of Biomedical and Health Informatics* 17 (4) (2013) 828–834 (jul 2013). doi:10.1109/jbhi.2013.2245674.
- [12] A. Tsanas, M. Little, P. McSharry, J. Spielman, L. Ramig, Novel speech signal processing algorithms for high-accuracy classification of parkinsons disease, *IEEE Transactions on Biomedical Engineering* 59 (5) (2012) 1264–1271 (2012). doi:10.1109/TBME.2012.2183367.
- [13] A. Verikas, A. Gelzinis, E. Vaidukynas, M. Bacauskiene, J. Minelga, M. Hällander, V. Uloza, E. Padervinskis, Data dependent random forest applied to screening for laryngeal disorders through analysis of sustained phonation: Acoustic versus contact microphone, *Medical Engineering & Physics* 37 (2) (2015) 210 – 218 (2015). doi:10.1016/j.medengphy.2014.12.005.
- [14] M. Cernak, J. R. Orozco-Arroyave, F. Rudzicz, H. Christensen, J. C. Vásquez-Correa, E. Nöth, Characterisation of voice quality of parkinson's disease using differential phonological posterior features, *Computer Speech & Language* 46 (2017) 196–208 (nov 2017). doi:10.1016/j.csl.2017.06.004.
- [15] J. Ruzs, J. Hlavnicka, T. Tykalova, M. Novotny, P. Dusek, K. Sonka, E. Ruzicka, Smartphone allows capture of speech abnormalities associated with high risk of developing parkinson's disease, *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 26 (8) (2018) 1495–1507 (aug 2018). doi:10.1109/tnsre.2018.2851787.
- [16] J. Vásquez-Correa, J. Orozco-Arroyave, T. Bocklet, E. Nöth, Towards an automatic evaluation of the dysarthria level of patients with parkinson's disease, *Journal of Communication Disorders* 76 (2018) 21 – 36 (2018). doi:https://doi.org/10.1016/j.jcomdis.2018.08.002.
- [17] D. Gupta, S. Sundaram, A. Khanna, A. E. Hassani, V. H. C. de Albuquerque, Improved diagnosis of parkinson's disease using optimized crow search algorithm, *Computers Electrical Engineering* 68 (2018) 412 – 424 (2018). doi:https://doi.org/10.1016/j.compeleceng.2018.04.014.
- [18] C. R. Pereira, D. R. Pereira, G. H. Rosa, V. H. Albuquerque, S. A. Weber, C. Hook, J. P. Papa, Handwritten dynamics assessment through convolutional neural networks: An application to parkinson's disease identification, *Artificial Intelligence in Medicine* 87 (2018) 67 – 77 (2018). doi:https://doi.org/10.1016/j.artmed.2018.04.001.
- [19] D. Montaña, Y. Campos-Roca, C. J. Pérez, A diadochokinesis-based expert system considering articulatory features of plosive consonants for early detection of parkinson's disease, *Computer Methods and Programs in Biomedicine* 154 (2018) 89 – 97 (2018). doi:https://doi.org/10.1016/j.cmpb.2017.11.010.
- [20] Z. Galaz, J. Mekyska, V. Zvoncak, J. Mucha, T. Kiska, Z. Smekal, I. Eliasova, M. Mrackova, M. Kostalova, I. Rektorova, M. Faundez-Zanuy, J. Alonso-Hernandez, P. Gomez-Vilda, Changes in phonation and their relations with progress of parkinson's disease, *Applied Sciences* 8 (12) (2018) 2339 (nov 2018). doi:10.3390/app8122339.
- [21] Z. Cai, J. Gu, C. Wen, D. Zhao, C. Huang, H. Huang, C. Tong, J. Li, H. Chen, An intelligent parkinson's disease diagnostic system based on a chaotic bacterial foraging optimization enhanced fuzzy KNN approach, *Computational and Mathematical Methods in Medicine* 2018 (2018) 1–24 (jun 2018). doi:10.1155/2018/2396952.
- [22] P. Sharma, S. Sundaram, M. Sharma, A. Sharma, D. Gupta, Diagnosis of parkinson's disease using modified grey wolf optimization, *Cognitive Systems Research* 54 (2019) 100–115 (may 2019). doi:10.1016/j.cogsys.2018.12.002.
- [23] L. Parisi, N. RaviChandran, M. L. Manaog, Feature-driven machine learning to improve early diagnosis of parkinsons disease, *Expert Systems with Applications* 110 (2018) 182–190 (nov 2018). doi:10.1016/j.eswa.2018.06.003.
- [24] L. Berus, S. Klančnik, M. Brezocnik, M. Ficko, Classifying parkinson's disease based on acoustic measures using artificial neural networks, *Sensors* 19 (1) (2018) 16 (dec 2018). doi:10.3390/s19010016.
- [25] D. Gupta, A. Julka, S. Jain, T. Aggarwal, A. Khanna, N. Arunkumar, V. H. C. de Albuquerque, Optimized cuttlefish algorithm for diagnosis of parkinson's disease, *Cognitive Systems Research* 52 (2018) 36 – 48 (2018). doi:https://doi.org/10.1016/j.cogsys.2018.06.006.
- [26] L. Moro-Velazquez, J. A. Gomez-Garcia, J. I. Godino-Llorente, J. Villalba, J. Ruzs, S. Shattuck-Hufnagel, N. Dehak, A forced gaussians based methodology for the differential evaluation of parkinson's disease by means of speech processing, *Biomedical Signal Processing and Control* 48 (2019) 205 – 220 (2019). doi:https://doi.org/10.1016/j.bspc.2018.10.020.
- [27] L. C. Afonso, G. H. Rosa, C. R. Pereira, S. A. Weber, C. Hook, V. H. C. Albuquerque, J. P. Papa, A recurrence plot-based approach for parkinson's disease identification, *Future Generation Computer Systems* 94 (2019) 282 – 292 (2019). doi:https://doi.org/10.1016/j.future.2018.11.054.
- [28] C. R. Pereira, D. R. Pereira, S. A. Weber, C. Hook, V. H. C. de Albuquerque, J. P. Papa, A survey on computer-assisted parkinson's disease diagnosis, *Artificial Intelligence in Medicine* (2018). doi:https://doi.org/10.1016/j.artmed.2018.08.007.
- [29] S. Fahn, R. L. Elton, UPDRS Development Committee (1987) Unified Parkinson's Disease Rating Scale, Macmillan: Florham Park, NJ, USA, 1987 (1987).