Interspeech 2018
2-6 September 2018, Hyderabad

Multimodal i-vectors to Detect and Evaluate Parkinson's Disease

N. Garcia[1], *J.C Vásquez-Correa*[1,2], *J. R. Orozco-Arroyave*[1,2], *E. Nöth*[2]

[1]Faculty of Engineering, Universidad de Antioquia UdeA, Medellín, Colombia
[2]Pattern Recognition Lab, Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany

nicanor.garcia@udea.edu.co

Abstract

Parkinson's Disease (PD) is a neurodegenerative disorder characterized by a variety of motor symptoms. PD patients show several motor deficits, including speech deficits, impaired handwriting, and gait disturbances. In this work we propose a methodology to compute i-vectors extracted from three different bio-signals: speech, handwriting, and gait. These i-vectors are used to classify patients and healthy controls, and to evaluate the neurological state of the patients. Speech i-vectors are extracted from Mel-Frequency Cepstral Coefficients (MFCCs), handwriting i-vectors are extracted from kinematic features, and gait i-vectors are extracted from modified MFCCs computed from inertial sensor signals. Two fusion strategies are tested: concatenating the i-vectors of a subject to form a super-i-vector with information from the three bio-signals and score pooling. The super-i-vector fusion method leads to better classification results (accuracy of 85%) with respect to the separate analysis with each bio-signal.

Index Terms: Parkinson's Disease, i-vectors, pathological speech, handwriting processing, gait processing.

1. Introduction

Parkinson's Disease (PD) is the second most common neurodegenerative disease worldwide after Alzheimer's, and its incidence and prevalence are rising [1]. PD is often associated with its primary motor symptoms, which include tremor, akinesia, bradykinesia, and postural instability [2]. These symptoms result in gait impairments and handwriting deficits. Additionally, secondary motor symptoms include speech disorders such as hypokinetic dysarthria. PD is a progressive disorder, that is, symptoms get worse over time. All these problems lead to a lower quality of life for the patients, hindering their communication skills and mobility. Additionally, they make burdensome for the patient to attend medical checks and therapy appointments [3]. To diagnose assess the progression of the disease, neurologists apply different tests to the patient. The most common is the Movement Disorder Society Unifified Parkinson's Disease Rating Scale (MDS-UPDRS). However, the tests are subject to the expertise of the clinician, therefore their inter- and intra-rater variability could be high.

In recent years, the scientific community has analyzed different bio-signals aiming to detect and evaluate the progression of the disease. The final goal is to provide neurologists with methods to support their diagnosis objectively and patients with tools to improve their therapy and treatment. Hopefully, these automatic methods will improve the understanding of the disease and lead to improved quality of life for the patients [3]. Among the analyzed bio-signals, inertial signals from gait, online handwriting signals, and speech recordings are highlighted as non-intrusive. Regarding the assessment of speech impairments of PD patients, in [4] the authors classified the speech of PD patients vs. healthy controls (HC) speakers using the Gaussian Mixture Models-Universal Background Model (GMM-UBM) approach and i-vector models reporting classification accuracies up to 87%. Studies performed at our lab also suggest that i-vectors can be used to model the changes in speech due to PD: in [5] i-vectors extracted from features that characterized different dimensions of speech were used to obtain a Spearman's correlation of up to 0.63 with the third part of the MDS-UPDRS neurological evaluation score . Online handwriting signals were analyzed in [6]. The authors used kinematic and pressure features to characterize online handwriting signals captured with a Wacom digitizer tablet. They classified PD patients and age balanced HC, and reported an accuracy of up to 81.3%. In [7] the authors performed a handwriting assessment using a smart-phone application where the patients draw a spiral computing several features: the kurtosis of the speed stroke, the length of the spiral drawing curve, the area of the spiral in each loop, the time of the drawing, among others. The authors evaluated different items of the Unified Parkinson's Disease Rating Scale (UPDRS) score related to the upper limbs, and reported correlations ranging from 0.47 to 0.52 combining handwriting features with finger-taping measures. In [8] the authors proposed two novel interpretable features to assess gait impairments in PD patients: The peak forward acceleration in the loading phase and peak vertical acceleration around heel-strike, which encodes the engagement in stride initiation and the hardness of the impact at heel-strike, respectively. The new features were correlated with the UPDRS-III score of 98 PD patients. The results indicate that the proposed features correlate with the disease progression and the loss of postural agility/stability of the patients.

Although there are several studies considering different bio-signals to assess the motor impairments of PD patients, most of these studies consider only one bio-signal (modality) at a time. Multimodal analyses, i.e., considering information from different sensors, have not been researched extensively [9]. In previous studies [10] we found that the combination of three bio-signals (speech, handwriting and gait) can improve the results of the automatic assessment of the motor capabilities of the patients relative to the performance when only one bio-signal is considered. The results also improved the accuracy of classification of PD patients vs. HC subjects.

In recent years the i-vector approach [11], which was initially conceived for speaker verification tasks, has become the state of the art in many other speech analysis tasks, including PD speech analysis. This approach has also been adapted to perform biometric verification tasks using handwriting and gait signals. In [12] the authors performed the identification of a person from gait signals captured with a smart-phone, while online handwritten signature verification was performed with i-vectors in [13]. The results of these works suggest that i-vectors are able to capture the traits of a person in different bio-signals.

10.21437/Interspeech.2018-2295

This means, that the i-vectors could also capture the effects of PD in handwriting and gait bio-signals. This study proposes to test i-vectors extracted from gait, handwriting, and speech signals to perform automatic analysis of PD. Additionally, we propose two fusion methods to combine these i-vectors. The performance of the proposed model is tested in two aspects: i) the classification between PD patients and HC subjects, and ii) the evaluation of the neurological state of the patients. In this case, we consider the third section of the Movement Disorder Society Unified Parkinson's Disease Rating Scale (MDS-UPDRS-III) [14] to assess the neurological state of the patients. This section evaluates the motor skills of the patient through different exercises. The results indicate that the fusion of i-vectors from different modalities improve the classificaton of PD vs. HC subjects up to 6%.

The rest of this work is divided as follows: Section 2 describes the data considered in this study. Section 3 explains the methodology followed in this study, including the steps to build the multimodal i-vectors. Section 4 describes the experiments and shows the results of this work. Finally, section 5 presents some conclusions and future work.

2. Data

The experiments were performed considering data from the following subjects:

- 49 PD patients with average age 60 ± 10.0 years. Most in early to mid-stages of the disease. MDS-UPDRS-III scores are available for 39 of them.
- Elderly HC: 41 with average age 65.1 ± 10.8.

For each subject, speech, handwriting, and gait signals were captured during the same recording session. In addition, signals of 40 young HC (average age 24.3 ± 4.0) were used to increase the amount of data available to train the Universal Background Model and i-vector extractor. These young HC were not included in any test group. All the subjects are Colombian Spanish native speakers.

During the recordings all the subjects followed the speech protocol described in [15]. This study only considers the speech signals corresponding to ten short sentences. The speech signals have a sampling frequency of 16 kHz. Speech signals for 30 of the patients were recorded with professional equipment under controlled acoustic conditions by using an acoustically treated box to isolate some of the acoustic noise. The rest of the patients and HC were recorded with a generic Logitech headset in quiet places where possible. It should be noted that the acoustic conditions make this task more difficult. Gait signals were captured using the eGait platform[1] [16] which has two inertial sensors (tri-axial accelerometers and gyroscopes). The sensors are attached to the lateral heel of the shoes. The signals were captured with a sampling frequency of 102 Hz. The tasks performed by the subjects included a 20 meters walk with a stop at 10 meters and a 40 meters walk stopping every 10 meters. The same 30 patients mentioned before were recorded using different shoes than the rest of the subjects[2]. The handwriting signals were captured using a Wacom Cintiq 13-HD digitizing tablet with a sampling frequency of 180 Hz. In this tablet the subject writes/draws over a screen which provides instant visual feedback of their strokes. The tablet captures the x, y, and z positions, azimuth and altitude angles, and the pressure of the pen on the tablet screen. The tasks included writing their name, their ID, the numbers from 0 to 9, the alphabet, and spontatenous sentence, and drawing geometric shapes e.g., an Archimedian spiral, a cube, and the Rey-Osterrieth figure [17], among others. Due to their education background not all patients could perform the writing tasks.

[1] eGaIT - embedded Gait analysis using Intelligent Technology, `http://www.egait.de/`

[2] These 30 patients were the first patients to be considered. For technical reasons we had to change some of the capture equipment between recording sessions. We are aware this could influence the results. The measures taken to minimize this influence are described in Section 4.

3. Methods

The methodology followed in this work comprises four steps: (1) features are extracted from the bio-signals, (2) the i-vectors are extracted for the computed features, (3) i-vectors extracted from each bio-signal are combined, and (4) the i-vector are used to classify the PD vs. HC subjects, or to assess the neurological state of the patients.

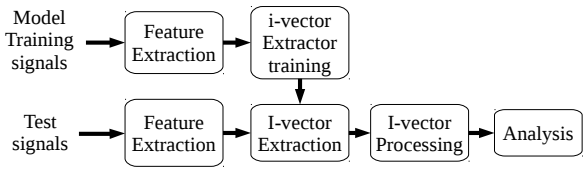

Figure 1: *Methodology for the i-vector analysis*

3.1. Feature extraction

3.1.1. Speech Modeling

To characterize these signals, 20 MFCCs (including $MFCC_0$) and their first and second derivatives were extracted from frames of 25 ms length with a time-shift of 10 ms. This is the standard set of features used in the i-vector approach. They also have been successfully used to model articulation in PD [18].

3.1.2. Handwriting Modeling

The raw signals captured with the tablet, and their first and second derivatives, were used as kinematic features to characterize the handwriting process. These constitute a set of 18 features. The signals were scaled to be in a range of [0,1], except the y-position which was scaled to keep a the same aspect ratio of the table (16:9).This set of features is a subset of the kinematic features proposed in [6].

3.1.3. Gait Modeling

As mentioned before, the inertial sensors include a triaxial (x, y, and z) accelerometer and gyroscope attached to each foot, for a total of 12 gait signals. Each of the gait signals is characterized with the modified version of the MFCCs proposed in [12]. For this case, 8 coefficients per signal were used to keep the number of features low. Frames of 0.32 s with an overlap of 50% were used. This frame length was selected to obtain enough frames for the i-vector modeling as limited amounts of data are available.

3.2. i-vector approach

In the i-vector approach, a factor analysis model is used as a feature extractor [11]. A new low-dimensional vector space known as the Total Variability space is defined. For speech signals this

vector space models the inter- and intra-speaker variabilities as well as channel effects. The inter-speaker variability carries the information about traits of the speech of an individual, in this case we hope to capture the effects due to the disease. On the other hand, intra-speaker variability could provide insight on the progression of the disease over time. Analysis beyond the scope of this work is needed to provide similar interpretations to the i-vectors extracted from handwriting and gait signals.

The factor analysis model is expressed by Equation 1, where $\mathbf{M}$ is the GMM supervector for a given signal, where $\mathbf{m}$ is the speaker and channel independent GMM supervector (usually taken from a UBM), $\mathbf{T}$ is the Total Variability matrix, and $\mathbf{w}$ is the i-vector, which is a standard normally distributed latent variable.

$$\mathbf{M} = \mathbf{m} + \mathbf{T}\mathbf{w} \tag{1}$$

For this work the dimension of the i-vector, $\dim_w$, was chosen based on the number of components in the UBM. M and the dimension of the original feature vector, $\dim_f$, following the relation $\dim_{\mathbf{w}} = \log_2(M)\dim_f$.

Once the i-vectors are extracted, we process them in two steps: (1) i-vectors of the different tasks of a given subject are averaged to obtain one i-vector per subject, and (2) Principal Component Analysis (PCA) is applied to the subject i-vectors to perform a whitening transformation [19].

3.3. Super i-vectors

A multimodal super i-vector is formed by concatenating the i-vectors from each bio-signal. This is expressed by Equation 2, where H, G, and S are the dimensions of the handwriting, gait and speech i-vectors.

$$\mathbf{w}_f = \begin{bmatrix} \mathbf{w}_h \\ \mathbf{w}_g \\ \mathbf{w}_s \end{bmatrix}_{(H+G+S)\times 1} \tag{2}$$

3.4. Classification

A soft margin Support Vector Machine (SVM) with Gaussian kernel is used to classify PD vs. elderly HC subjects. This classifier was chosen as it has obtained good results when analyzing PD bio-signals [20]. Two hyper-parameters need to be optimized in this classifier: the margin cost C and the bandwidth of the Gaussian kernel γ. Details of the optimization procedure are given in the next section.

3.5. Evaluation of the neurological state

To perform evaluation of the neurological state of the patient we correlate the MDS-UPDRS-III label with the average cosine distance between the subject's i-vector and a set of N reference i-vectors. The average cosine distance for the $j-th$ test subject i-vector is given by Equation 3.

$$\mathrm{d}(\mathbf{w}_{\mathrm{test},j}) = \frac{1}{N}\sum_{i=1}^{N}\left(1 - \frac{\mathbf{w}_{\mathrm{test},j}\cdot\mathbf{w}_{\mathrm{ref},i}}{||\mathbf{w}_{\mathrm{test},j}||||\mathbf{w}_{\mathrm{ref},i}||}\right) \tag{3}$$

The Spearman's rank correlation coefficient is used to assess the correlation between the average cosine distances of the patients with their MDS-UPDRS-III labels. Three sets of reference i-vectors were used in this case: i) young HC i-vectors, ii) elderly HC i-vectors, iii) PD patients i-vectors.The reference sets will be described in more detail in the next section.

A multimodal fused score is obtained by averaging the cosine distance of the i-vectors of each bio-signal. This is referred to as Score fusion.

4. Experiments and results

4.1. Validation

A five-fold cross-validation scheme was implemented for the classification experiment, i.e., 80% of the data are used to train the models and the remaining 20% are considered to test the system. The hyper-parameters are optimized by performing a nested three-fold cross-validation over 80% training data. The hyper-parameters that need to be optimized are the number of Gaussians in the UBM, the number of components in PCA, and the parameters of the SVM classifier (C, γ). To minimize possible bias due to the different microphones and shoes used to capture the signals, the patients of each fold were balanced according to this condition.

4.2. Experiments

For the classification task three experiments were performed: (1) signals from PD, elderly HC, and young HC were used to train the i-vector extractor, (2) the i-vector extractor was trained using only the signals from young HC as these subjects are easier to find and enlist, and (3) the i-vector extractor was trained with the signals from young HC and a small number of eldelry HC. In all experiments only PD patients and elderly HC subjects are considered in the test set. The young HC are not considered. The results are given in terms of the the Area Under the ROC-curve (AUC), and the accuracy (Acc.), sensitivity (Sens.), and specificity (Spec.) of the classifier

4.2.1. Experiment 1

Table 1: *Classification results PD vs. HC, experiment 1*

Signal	Acc. (%)	AUC	Sens. (%)	Spec. (%)
Gait	76.9 ± 9.1	0.83	77.1 ± 11.5	76.8 ± 12.5
Handwriting	75.1 ± 3.7	0.82	79.3 ± 7.4	70.0 ± 17.0
Speech	79.4 ± 7.8	0.87	83.1 ± 15.2	75.0 ± 17.7
Super i-vector	85.0 ± 9.6	0.92	81.3 ± 12.4	89.6 ± 9.5

The results in Table 1 show that the best results are obtained when considering the super-i-vector fusion method. This fusion method also achieves better classification results than analyzing the i-vectors of each bio-signal separately. This method specially improves the average specificity, that is, the capability of the system to recognize HC from PD. The results also show that speech is the bio-signal with better classification performance, and probably the signal that contributes most information in the fusion. This could be expected as the i-vector approach was initially thought to analyze speech signals.

4.2.2. Experiment 2

Table 2: *Classification results PD vs. HC, experiment 2*

Signal	Acc. (%)	AUC	Sens. (%)	Spec. (%)
Gait	70.4 ± 5.5	0.70	66.9 ± 13.5	75.0 ± 22.4
Handwriting	74.9 ± 7.1	0.80	74.9 ± 18.1	75.0 ± 13.7
Speech	80.7 ± 9.0	0.88	83.6 ± 8.3	77.1 ± 21.4
Super i-vector	81.9 ± 6.2	0.89	83.8 ± 4.6	79.6 ± 9.8

The results in Table 2 show that the classification performance of gait signal decreases when the i-vector extractor is only trained with young HC signals. In this case, the sensitivity is highly affected, while the specificity only decreases slightly. This result could be explained by the fact that gait problems are usually more prominent in latter stages of the disease [21]. The lower accuracy in gait signals also affects the results of the super i-vectors which also show a lower performance compared to the experiment 1. The other two signals in the fusion keep the classification accuracy, but the specificity is more affected than the sensitivity. The classification results considering speech signals were not affected. This means that robust models for speech can be trained using only signals from young HC.

4.2.3. Experiment 3

Table 3: *Classification results PD vs. HC, experiment 3*

Signal	**Acc. (%)**	**AUC**	**Sens. (%)**	**Spec. (%)**
Gait	69.2 ± 5.3	0.77	65.3 ± 10.0	73.9 ± 14.8
Handwriting	71.8 ± 7.3	0.78	81.3 ± 12.4	60.0 ± 26.7
Speech	72.5 ± 7.5	0.83	87.3 ± 12.7	53.9 ± 12.3
Super i-vector	83.8 ± 9.3	0.90	79.3 ± 7.4	89.3 ± 13.2

The results in Table 3 show that training the extractor with signals from old HC and young HC does not improve the classification results, moreover, in the case of handwriting and speech the results are worse than in the previous experiments. However, the proposed fusion method was not affected. The results of experiments 2 and 3 indicate that the Total Variability Matrix is sensible to the data used to train it. A robust extractor needs to be trained with both signals of PD patients and HC in order to take into account the variability due to the disease. This means that when analyzing PD bio-signals with the i-vector approach is not enough to add more data. In order to obtain robust models a careful selection of the data added needs to be done, and probably signals from PD patients are needed.

4.2.4. Neurological evaluation experiment

For the evaluation of the neurological state, the i-vector extractor was trained only with signals of young HC. The test set for this experiment comprises the 39 patients for which MDS-UPDRS-III scores are available. The YHC reference set is obtained by extracting the i-vectors from signals of young HC, however this reference could be biased, as these were the same signals used to train the extractor, that is, these i-vectors comprise the same vector space as the rows of the Total Variabilty Matrix. The OHC reference set comprises the i-vectors extracted from elder HC signals. Finally, the PD reference is obtained from the i-vectors of the signals of the 10 patients not included in the test set.

Table 4: *Spearman's correlation between the cosine distance and the MDS-UPDRS-III*

Signal	ρ **- YHC ref.**	ρ **- EHC ref.**	ρ **- PD ref.**
Gait	-0.14	-0.11	-0.25
Handwriting	0.20	-0.07	-0.18
Speech	-0.14	0.30	-0.33
Super-i-vector	0.03	-0.08	-0.26
Score fusion	0.31	0.20	-0.41

The results in Table 4 show a negative correlation between the PD reference average cosine distances and MDS-UPDRS-III (last column) in all cases. This means that as disease progresses, the distance to an affected reference decreases. The opposite could be said about the correlation between the average cosine distance for HC references, however, in most cases these correlations are low or negative. This could indicate that these references are not adequate for this task. The results also show that the proposed super-i-vector fusion was not able to improve the evaluation results. This may indicate that the super-i-vectors do not behave like normal i-vectors. An additional fusion method was proposed to try to improve the results. In this method, referred to as score fusion, the average cosine distances from the three different bio-signals are averaged. The score fusion method provides a slight improvement in terms of the Spearman's correlation.

5. Conclusion and future work

In this study, detection and evaluation of PD was performed using three different bio-signals. Speech signals were characterized using MFCCs, online handwriting signals with kinematic features, and gait signals with modified MFCCs. Then, each of these signals was modeled using the i-vector approach. Classification between PD and age balanced HC was done using a SVM classifier. Finally, the average cosine distance to a reference set of i-vectors was correlated to the MDS-UPDRS-III labels of patients.

The results for each individual bio-signal show that i-vectors extracted from speech perform better in both analyses. This result makes sense because the i-vector approach was initially proposed for speech analysis. The results from experiments 2 and 3 show that just adding more data to train the i-vector extractor does not always leads to better results. We think that the results can be improved by choosing data from different tasks (speech, gait and handwriting) such that contain more variability. A fusion scheme based on concatenating the i-vectors extracted from different bio-signals was proposed. The resulting vector is called super-i-vector. This fusion method improved the classification performance: however, it did not improve the correlation with the neurological state of the patients. Another fusion method was proposed for this correlation task. It consisted of averaging the scores over the i-vectors of the three bio-signals per patient. The method improved the correlation with the neurological labels that the neurologist assigned to each patient.

Future work we will refine the feature sets, especially those that represented gait and handwriting, since the results were not satisfactory, thus we think that the selected features did not capture the patient's condition properly. On the other hand, since the i-vector approach was initially developed for speech signals, it might need to be adapted to model other bio-signals. Additionally, other fusion methods to improve the performance in the neurological evaluation of the patient could be proposed. Finally, we are collecting bio-signals using smart-phones instead of dedicated platforms.

6. Acknowledgements

This work was financed by CODI from University of Antioquia by grants Number PRG-2015-7683 and PRV-16-2-01. This project has received funding from the European Unions Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie Grant Agreement No. 766287

7. References

[1] S. Sveinbjornsdottir, "The clinical symptoms of Parkinson's disease," *Journal of Neurochemistry*, no. 139, pp. 318–324, jul 2016.

[2] A. A. Moustafa, S. Chakravarthy, J. R. Phillips, A. Gupta, S. Keri, B. Polner, M. J. Frank, and M. Jahanshahi, "Motor symptoms in Parkinson's disease: A unified framework," *Neuroscience and Biobehavioral Reviews*, vol. 68, pp. 727–740, 2016.

[3] J. A. Stamford, P. N. Schmidt, and K. E. Friedl, "What engineering technology could do for quality of life in Parkinson's disease: A review of current needs and opportunities," *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 6, pp. 1862–1872, nov 2015.

[4] L. Moro-Velázquez, J. A. Gómez-García, J. I. Godino-Llorente, J. Villalba, J. R. Orozco-Arroyave, and N. Dehak, "Analysis of speaker recognition methodologies and the influence of kinetic changes to automatically detect parkinson's disease," *Applied Soft Computing*, vol. 62, pp. 649 – 666, 2018.

[5] N. Garcia, J. R. Orozco-Arroyave, L. F. D'Haro, N. Dehak, and E. Nöth, "Evaluation of the neurological state of people with parkinson's disease using i-vectors," in *Proceedings of the 18th INTERSPEECH (In Press)*, 2017.

[6] P. Drotár, J. Mekyska, I. Rektorová, L. Masarová, Z. Smékal, and M. Faundez-Zanuy, "Evaluation of handwriting kinematics and pressure for differential diagnosis of parkinson's disease," *Artificial Intelligence in Medicine*, vol. 67, no. C, pp. 39 – 46, 2016.

[7] S. Aghanavesi, D. Nyholm *et al.*, "A smartphone-based system to quantify dexterity in Parkinson's disease patients," *Informatics in Medicine Unlocked*, vol. 9, pp. 11–17, 2017.

[8] J. Hannink, H. Gaßner *et al.*, "Inertial sensor-based estimation of peak accelerations during heel-strike and loading as markers of impaired gait patterns in PD patients," *Basal Ganglia*, vol. 8, p. 1, 2017.

[9] Q. W. Oung, H. Muthusamy *et al.*, "Technologies for assessment of motor disorders in Parkinson's disease: a review," *Sensors*, vol. 15, no. 9, pp. 21 710–21 745, 2015.

[10] J. C. Vásquez-Correa, J. R. Orozco-Arroyave *et al.*, "Multiview representation learning via gcca for multimodal analysis of Parkinson's disease," in *Proceedings of the 42nd ICASSP*, 2017, pp. 1–6.

[11] N. Dehak, P. J. Kenny, R. Dehak, P. Dumouchel, and P. Ouellet, "Front-End Factor Analysis for Speaker Verification," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 19, no. 4, pp. 788–798, may 2011.

[12] R. San-Segundo, J. M. Montero, R. Barra-Chicote, F. Fernndez, and J. M. Pardo, "Feature extraction from smartphone inertial signals for human activity segmentation," *Signal Processing*, vol. 120, no. Supplement C, pp. 359 – 372, 2016.

[13] H. Zeinali and B. BabaAli, "On the usage of i-vector representation for online handwritten signature verification," in *Proocedings of the 14th ICDAR*, 2016.

[14] C. G. Goetz et al., "Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Scale presentation and clinimetric testing results," *Movement Disorders*, vol. 23, no. 15, pp. 2129–2170, 2008.

[15] J. R. Orozco, J. D. Arias, J. F. Vargas, M. C. González Rátiva, and E. Nöth, "New Spanish speech corpus database for the analysis of people suffering from Parkinson's disease," in *Proceedings of the 9th LREC*, 2014, pp. 342–347.

[16] J. Klucken, J. Barth *et al.*, "Unbiased and mobile gait analysis detects motor impairment in Parkinson's disease," *PloS one*, vol. 8, no. 2, p. e56956, 2013.

[17] M. Shin, S. Park, S. Park, S. Seol, and J. Kwon, "Clinical and empirical applications of the Rey–Osterrieth complex figure test," *Nature Protocols*, vol. 1, no. 2, pp. 892–899, 2006.

[18] J. R. Orozco-Arroyave et al., "Automatic detection of Parkinson's disease from words uttered in three different languages," *Journal of Acoustical Society of America*, vol. 139, no. 1, pp. 481–500, jan 2016.

[19] D. Garcia-Romero and C. Espy-Wilson, "Analysis of i-vector Length Normalization in Speaker Recognition Systems." in *Proceedings of the 12th INTERSPEECH*, sep 2011.

[20] J. Orozco-Arroyave, *Analysis of speech of people with Parkinson's disease*, 1st ed. Berlin, Germany: Logos-Verlag, 2016.

[21] P.-H. Chen, R.-L. Wang, D.-J. Liou, and J.-S. Shaw, "Gait disorders in parkinson's disease: Assessment and management," *International Journal of Gerontology*, vol. 7, no. 4, pp. 189 – 193, 2013.