2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Spectral Electroencephalographic and Heart Rate Variability features enhance identification of medicated/non-medicated Parkinson's disease patients

MariNieves Pardo-Rodríguez[1], Erik Bojorges-Valdez[2] and Oscar Yanez-Suarez[3]

Abstract— **Parkinson's Disease is a neuropathy that produces changes in several biomarkers, these changes could be used to evaluate even sub–clinical conditions. This paper presents an evaluation of indices extracted from electroencephalography and Heart Rate Variability (HRV), when used to classify a sample of subjects from three groups: control (healthy), medicated and non medicated subjects diagnosed with Parkinson's disease. Classification performance was measured using accuracy over these classes and a cross validation scheme was used to assess repeatability for the classification process. Results tend to prove that inclusion of an autonomic index derived from HRV analysis enhances classification, suggesting that Parkinson's disease could be related with unperceptible to mild alterations of the Autonomic Nervous System.**

I. INTRODUCTION

Parkinson's disease (**PD**) is a neurodegenerative disorder that alters the dopaminergic system [1]. In advanced stages it produces motor symptoms (tremor) and diverse cognitive impairments. Some hypothesis point to associate this neurodegenerative process with alterations on the Autonomic Nervous System (ANS), impacting on heart rate [2] and gait patterns [3].

Supervised machine learning (ML) algorithms label new unseen information based on data distributions derived from prior information. This is a two step process: first, the classifier is trained (calibrated) with labeled data producing a model, and then, new data are assigned to a class according to this calibration setup. The generated model describes a usually non–convex mapping, even delimiting non contiguous regions. Despite the model used, for a reliable classification process data must have differentiated distributions for each class either in the original or transformed feature space. If that condition is satisfied, then such features could be considered as a discriminative data descriptor. The Gaussian Naive Base classification algorithm applies the Bayes' theorem with the "naive" assumption the distribution of the features is Gaussian [4].

ML models for medical data are often used as a diagnostic tool. For example, assuming two classes (healthy and diseased), any physiological variable (like glucose concentration in blood, heart rate, blood pressure, etc.) that changes by presence of illness is suitable to be considered as input for ML. Because of the high complexity of the biological systems, disease could produce non–evident and non–simple changes on several of the observed variables.

In their literature review and meta–analysis about **PD**, Bočková and Rektor [5] considered EEG studies and observed that β ([14-30]Hz) band shows an excessive synchrony and also some alterations on γ ([30-100] Hz) on cortical–sub-cortical loops. Melgari et al. [6] have shown that levadopa (L-dopa) administration on Parkinson's patients produces significant increase over spectral power on β and α ([8-14] Hz). They hypothesize that dopaminergic networks are implicated on abnormal oscillatory patterns reflected on such bands. Extending this idea, these alterations could impact on the structure of several other networks like the Central Autonomic Network (CAN) which contributes to maintain homeostasis [7].

This paper presents an evaluation of the discriminating power of diverse spectral features used to identify PD patients with and without L-dopa intake, and healthy controls, from recordings in a publicly available EEG database [8]. The features were extracted from EEG and ECG signals digitized during three minutes while in a resting condition. The working hypothesis assumes that EEG features bring information about cortical neurons, and from ECG, the HRV patterns will incorporate information about ANS. Feature pair selections among EEG and HRV measures were conducted either separately per signal type or combined. The features analyzed provide information about dynamics and mean values of the EEG band power in several channels, and HRV features come from different components to analyze sympathetic and parasympathetic isolated contributions.

II. MATERIALS AND METHODS

A. Database

Data used was obtained from the "Resting State EEG Data from Patients with Parkinson's Disease" public dataset collected at the University of San Diego and curated by Alex Rockhill at the University of Oregon [8]. EEG and ECG from 14 subjects with Parkinson's and 16 healthy subjects, all right handed, 17 female, ages 50 through 82, were analyzed. Data from Parkinson's subjects were obtained on two sessions, **On** and **Off** their medication, while healthy subjects (**HC**) just went through one session. No specific tasks were performed during sessions.

*This work was supported by División de Investigación y Posgrado of Universidad Iberoamericana Ciudad de México

[1] Biomedical Engineering Bachelor program at Universidad Iberoamericana Ciudad de México mnnipi@yahoo.com.mx

[2] Engineering Studies for Innovation Department, Universidad Iberoamericna Ciudad de México erik.bojorges@ibero.mx

[3] Electrical Engineering Department, Universidad Autónoma Metropolitana–Iztapalapa oyanez@izt.uam.mx

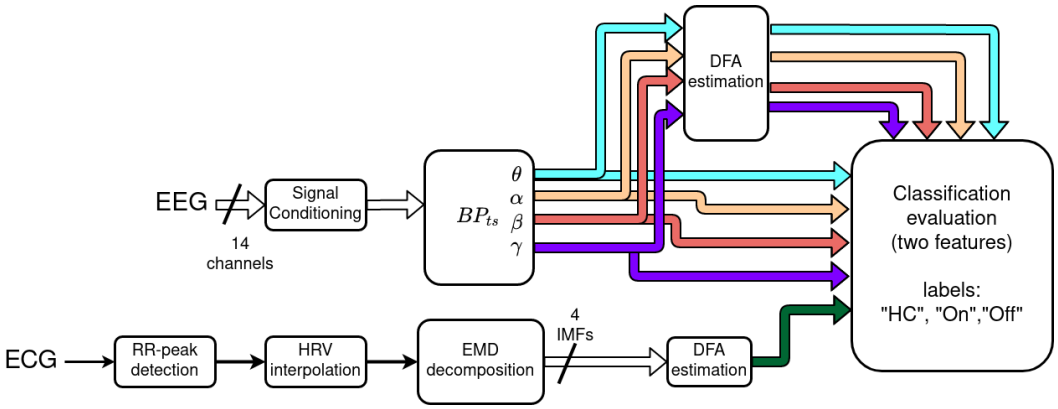

Fig. 1. Preprocessing, feature extraction and classification pipeline. EEG features are measured from band power time series (**BPts**) over different bands, while HRV features are determined as DFA estimates of the EMD decomposition of the interpolated RR series (details in the text).

All participants provided written consent in accordance to the Institutional Review Board of the University of California, San Diego and the Declaration of Helsinki [9].

B. Signal processing and feature extraction

The ECG signal was processed following the QRS detection procedure described in [10], to obtain the RR interval series. The HRV signal was then interpolated at a sampling frequency of 10 sps using a cubic spline. The HRV signal was decomposed into IMFs using the Empirical Mode Decomposition method and lastly the Detrended Fluctuation Analysis (DFA) indices were estimated for each IMF. DFA reflects long range temporal correlation of a time series, it is based on fractal analysis of time-series' variance along several windows of different lengths [11]. As a brief description, if the DFA exponent is close to 0.5 the series corresponds to white noise while values close to 1.5 indicate that the series has a presumably brownian motion behavior.

From EEG signals the band power time series $\mathbf{BP}_{ts}$ [12] were estimated for alpha ([8-12]Hz), beta ([14-30]Hz), gamma ([30-100]Hz) and theta ([4-8]Hz) bands with a sliding window of two seconds and a sliding step of 0.1 seconds using Welch periodogram estimators. The $\mathbf{BP}_{ts}$ computation consists of obtaining the signal's power spectral density and calculating the relative power for each frequency band, therefore the $\mathbf{BP}_{ts}$ show the distribution of power according to the time evolution of the components of the signal. The DFA indices were estimated for the $\mathbf{BP}_{ts}$ of channels Fp1, Fp2, F3, Fz, F4, C3, Cz, C4, P3, Pz, P4, O1, Oz, and O2.

C. Classification

Finally, for the classification task, a Gaussian Naive Bayes algorithm was used from the scikit-learn implementation [4]. The features were evaluated in pairs following a greedy strategy, random pairs of features were evaluated from the sets: **BPts** dynamics, **BPts** mean values, and IMF dynamics. Performance was evaluated with accuracy index in a 70% training-30% testing, 15–fold cross–validation scheme . The pairs of features were examined both in intra- and inter–set combinations, but not in an exhaustive manner. After some intra–set random choices, if band or dynamics did not achieve reliable performance the set was discarded an explored on inter–set evaluation.

Processing pipeline, from feature extraction to classification is summarized in Fig. 1.

III. RESULTS

A one-way ANOVA test was run across groups: **On**, **Off** and **HC** for the DFA values of the **BPts** of all bands and all channels; none of the groups had a significantly different population mean even though the **On** group tended towards lower values. When combining two of these features as inputs, the classifiers did not achieve an accuracy value above chance, and therefore were no longer included in the classification task.

When combining the mean **BPts** values across bands and channels as input features, combinations with and between beta channels achieved the highest accuracy values, followed by theta, gamma and alpha respectively. The best classification was achieved when using channels O1, Cz, Oz, C3, O2 and P3, these channels and regions had been previously identified as significant markers of **PD** [13].

When dynamics of the HRV components are used as input features to the classification, and specifically IMF2, the accuracy values for all cases was improved. It is important to mention that due to the length of the recordings, IMF3 and IMF4 should not be considered as reliable indices of sympathetic or ultra-low frequency components. Nevertheless, IMF2 power spectral density content corresponds mainly to the low frequency band (**LF**, [0.04– 0.15]), known to be linked majorly to the sympathetic branch of ANS activity.

Fig. 2 displays the overall best combination of features and their distributions. The plots along the diagonal show the marginal distribution for each group of the individual features: mean $\mathbf{BP}_{ts}$ on beta for channels Cz, O1 and Oz and the DFA values for IMF2 ($IMF2_{DFA}$), respectively. The heatmaps show the mean and standard deviation of the normalized confusion matrices computed from 15–fold cross–validation.

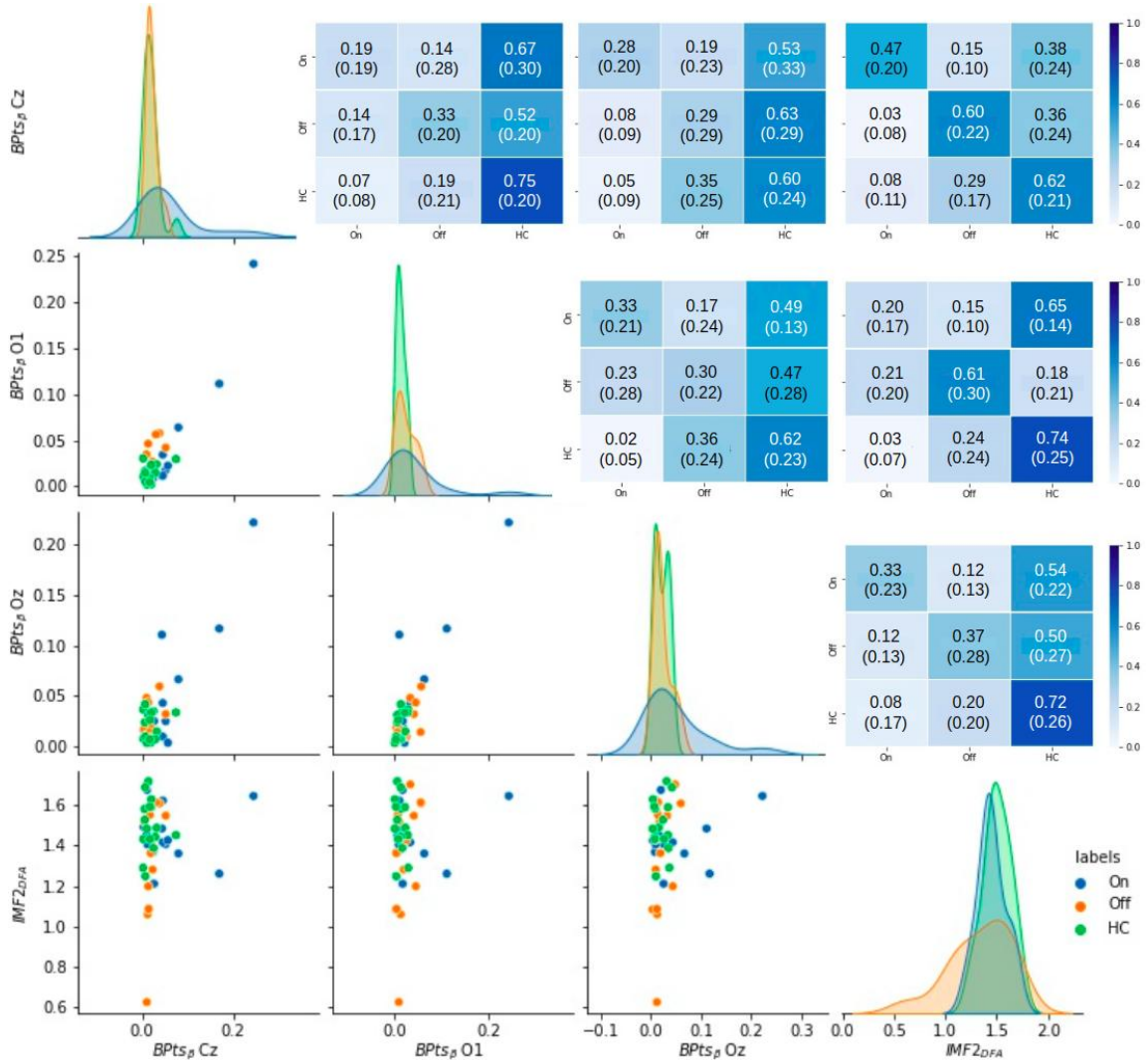

Fig. 2. Overall best feature combinations found: mean $\mathbf{BP}_{ts}$ in beta for channels Cz, O1 and Oz and DFA values for HRV IMF2. The plots along the diagonal show the marginal distribution for each group (**On**, **Off** and **HC** in blue, orange and green respectively) of the individual features. The scatter plots show the distribution between each pair of features. The heatmaps show the mean and standard deviation of the normalized confusion matrices after cross-validation.

IV. DISCUSSION

Fig. 2 shows how medication spreads beta $\mathbf{BP}_{ts}$ values on all channels causing a larger variance on its distribution with respect to the other groups (**Off** and **HC**), while also right shifting its mean value. Despite this being a known effect of the administration of medication, $\mathbf{BP}_{ts}$ DFA indices showed no significant differences between groups. Maybe no effect was found because of the small sample size rather than because the HRV dynamic remains unchanged since, as seen on Fig. 3, **On** group's indices do tend toward lower values and, possibly there are altered communication networks when medication is administered. This in turn suggests that the long range temporal correlations of central connections are not a clear biomarker of the disease, even though L-dopa administration seems to alter them.

Classification based merely on mean beta $\mathbf{BP}_{ts}$ between channels tends to label all groups as **HC**. The fact that $\mathbf{BP}_{ts}$ values on O1 of patients off their medication **Off** are slightly less concentrated than those of **HC** helps alleviate the confusion among these two groups, specially when combined with the $IMF2_{DFA}$ values.

When considering two groups -healthy and diseased-, these last two features (beta $\mathbf{BP}_{ts}$ on O1 and the $IMF2_{DFA}$) turn out to be the best to classify healthy and medicated patients (**HC** and **On**) from the non-medicated ones (**Off**); this consideration implies taking into account values labeled accurately but also **HC** labeled as **On** and viceversa. On the other hand, when considering the three groups individually -therefore taking into account only the accuracy values- beta $\mathbf{BP}_{ts}$ on Cz and the $IMF2_{DFA}$ indices are the best features for classifying each group correspondingly.

On Fig. 2 the last column shows how the $IMF2_{DFA}$ indices improve classification accuracy overall, suggesting **PD** displays affectations on the ANS dynamic involved with heart rate. The distribution of these values is similar between patients on medication (**On**) and **HC**, since they

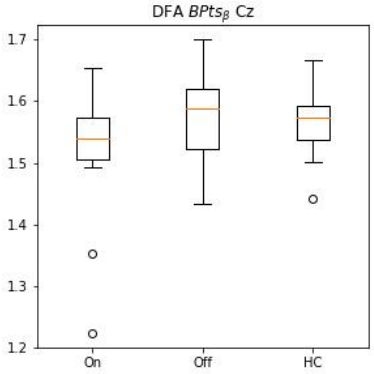

Fig. 3. DFA estimates of the band power time series (**BPts**) over beta band for each group (**On**, **Off** and **HC**)

both tend towards higher DFA indices, meaning the HRV dynamic of the components related to sympathetic and parasympathetic activity of these groups is more brownian than the one of patients off their medication (**Off**). In other words, administration of medication does seem to have a long range temporal effect on the **LF** components of the ANS, specifically on those impacting HRV. The fact that including this index enhances classification, despite of the feature it is paired with, suggests that **PD** produces sub–clinical autonomic dysfunctions, given that patients off their medication (**Off**) can be set apart based on their HRV behavior.

It is of great importance to remark that the database used is not very large therefore a greater set of patients and HC would be required in order to confirm or rule out these hypothesis. Also, longer recordings are needed in order to analyze and determine whether ultra-low frequency components related to cardiac activity and sympathetic ANS activity are useful indicators of the disease. Future work could include a stratified cross-validation scheme as well as HRV-only classification.

V. CONCLUSIONS

Classification results show **PD** exhibits affectations on the ANS activity related to the **LF** components that alter heart rate. Beta band power is clearly affected upon medication administration, nevertheless this effect seems not to display modifications of the signal dynamics that could be used as input features for the classification task. On the other hand, medication does seem to alleviate the effects of **PD** on the HRV **LF** dynamics, making this index a useful marker to discriminate non-medicated Parkinson's patients. It seems as beta $\mathbf{BP}_{ts}$ mean value and dynamics of IMF2 relationship is affected at the autonomic level after L-dopa administration, particularly this relationship tends to be closer with the **HC**. However further studies should be carried out in order to learn more about the effects of Parkinson's medication on the ANS.

ACKNOWLEDGMENT

The funding for participation on this conference were granted by División de Investigación y Posgrado of Universidad Iberoamericana Ciudad de México through the project: "Investigación en Interfaces Cerebro Computadora".

REFERENCES

[1] S. Evangelisti, F. Pittau, C. Testa, G. Rizzo, L. L. Gramegna, L. Ferri, A. Coito, P. Cortelli, G. Calandra-Buonaura, F. Bisquoli, C. Bianchini, D. N. Manners, L. Talozzi, C. Tonon, R. Lodi, and P. Tinuper, "L-dopa modulation of brain connectivity in Parkinson's disease patients: A pilot EEG-fMRI study," Frontiers in Neuroscience, vol. 13, p. 611, 2019. [Online]. Available: https://www.frontiersin.org/article/10.3389/fnins.2019.00611

[2] M. Kallio, K. Suominen, A. M. Bianchi, T. Mäkikallio, T. Haapaniemi, S. Astafiev, K. A. Sotaniemi, V. V. Myllylä, and U. Tolonen, "Comparison of heart rate variability analysis methods in patients with Parkinson's disease," Medical and Biological Engineering and Computing, vol. 40, no. 4, pp. 408–14, 2002. [Online]. Available: https://www.proquest.com/scholarly-journals/comparison-heart-rate-variability-analysis/docview/661695835/se-2?accountid=37347

[3] J. M. Hausdorff, "Gait dynamics in Parkinson's disease: common and distinct behavior among stride length, gait variability, and fractal-like scaling," Chaos (Woodbury, N.Y.), vol. 19, no. 2, pp. 026 113–026 113, Jun. 2009. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/19566273

[4] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," Journal of Machine Learning Research, vol. 12, pp. 2825–2830, 2011.

[5] M. Bočková and I. Rektor, "Impairment of brain functions in parkinson's disease reflected by alterations in neural connectivity in EEG studies: A viewpoint," Clinical Neurophysiology, vol. 130, no. 2, pp. 239–247, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1388245718313762

[6] J.-M. Melgari, G. Curcio, F. Mastrolilli, G. Salomone, L. Trotta, M. Tombini, L. di Biase, F. Scrascia, R. Fini, E. Fabrizio, P. M. Rossini, and F. Vernieri, "Alpha and beta EEG power reflects l-dopa acute administration in parkinsonian patients," Frontiers in Aging Neuroscience, vol. 6, p. 302, 2014. [Online]. Available: https://www.frontiersin.org/article/10.3389/fnagi.2014.00302

[7] H. D. Critchley, Y. Nagai, M. A. Gray, and C. J. Mathias, "Dissecting axes of autonomic control in humans: Insights from neuroimaging," Autonomic Neuroscience, vol. 161, no. 1, pp. 34 – 42, 2011, special Section: A selection of papers from the 6th Congress of the International Society for Autonomic Neuroscience (ISAN), September, 2009. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1566070210001967

[8] A. P. Rockhill, N. Jackson, J. George, A. Aron, and N. C. Swann, ""uc san diego resting state eeg data from patients with Parkinson's disease"," 2020.

[9] J. S. George, J. Strunk, R. Mak-McCully, M. Houser, H. Poizner, and A. R. Aron, "Dopaminergic therapy in parkinson's disease decreases cortical beta band coherence in the resting state and increases cortical beta band power during executive control," NeuroImage: Clinical, vol. 3, pp. 261–270, 2013. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2213158213001034

[10] A. E. Johnson, J. Behar, F. Andreotti, G. D. Clifford, and J. Oster, "R-peak estimation using multimodal lead switching," in Computing in Cardiology 2014. IEEE, 2014, pp. 281–284.

[11] C. Peng, S. Havlin, H. E. Stanley, and A. L. Goldberger, "Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series," Chaos: An Interdisciplinary Journal of Nonlinear Science, vol. 5, no. 1, pp. 82–87, 1995. [Online]. Available: https://doi.org/10.1063/1.166141

[12] M. Pardo-Rodriguez, E. Bojorges-Valdez, and O. Yanez-Suarez, "Causal relationship analysis of heart rate variability and power spectral density time series of electroencephalographic signals," in 2019 Computing in Cardiology Conference (CinC), vol. 46, 2019, pp. 1–4.

[13] X. He, Y. Zhang, J. Chen, C. Xie, R. Gan, R. Yang, L. Wang, K. Nie, and L. Wang, "The patterns of EEG changes in early-onset Parkinson's disease patients," International Journal of Neuroscience, vol. 127, no. 11, pp. 1028–1035, 2017, pMID: 28281852. [Online]. Available: https://doi.org/10.1080/00207454.2017.1304393