Machine Learning based Detection of Parkinson's Disease Using Beta-band PLV and Coherence Measures from EEG

Şevval Savaş[1], İbrahim Kaya[1,*]
[1]Department of Biomedical Engineering, İzmir Katip Çelebi University,İzmir, Türkiye
ORCIDs: 0009-0006-2321-816X, 0000-0003-0802-4376
E-mails:sewall.saavas@gmail.com, ibrahim.kaya@ikcu.edu.tr
*Corresponding author.

Abstract*—Parkinson's disease (PD) is a progressive neurodegenerative disorder that impairs motor functions and disrupts functional brain connectivity. Electroencephalography (EEG), with its high temporal resolution, enables investigation of these disruptions in resting-state brain activity. This study aims to differentiate PD patients from healthy controls by analyzing beta-band (12–35 Hz) EEG connectivity through two widely used synchronization metrics: Phase Locking Value (PLV) and magnitude-squared coherence. We utilized an open-access EEG dataset from the University of Iowa, comprising 13 individuals with PD and 13 age-matched healthy controls, recorded during eyes-open resting state. EEG signals were filtered in the beta band, and both PLV and coherence metrics were calculated across 63 channels. Connectivity heatmaps were generated for whole-brain and motor-related regions (F3, C3, Cz, C4, F4, FC3, FC4). While PD subjects showed increased phase synchronization (PLV), they demonstrated reduced coherence values compared to controls, suggesting abnormal hypersynchrony alongside decreased linear coupling. Statistical comparisons using independent t-tests revealed significant differences in specific connections (e.g., Cz–F4, F3–C3), particularly within motor areas. These features were used to train six supervised machine learning classifiers including Fine Tree, LDA, Linear SVM, Fine KNN, Naive bayes and neural network. Fine Tree model achieved promising classification accuracy of 84.7%, highlighting the potential of EEG-based features in aiding early diagnosis of Parkinson's disease. In conclusion, our findings demonstrate that PLV and coherence in the beta band, especially in motor-related networks, can serve as meaningful biomarkers. Combined with machine learning, this approach offers a non-invasive tool for distinguishing PD from healthy controls

Keywords*—Parkinson's disease; EEG; PLV; coherence; machine learning

I. Introduction

Parkinson's disease (PD) is a progressive neurodegenerative disorder affecting motor control and cognitive function, often disrupting functional connectivity in motor and executive networks [1], [2]. EEG, a non-invasive technique with high temporal resolution, enables detection of abnormal brain dynamics in PD [4]. Beta-band activity (12–35 Hz), linked to motor processing, is frequently altered in PD, especially in motor-related cortical areas [1], [5], [25]. Previous studies report increased phase-locking and decreased coherence in PD, varying by region and method [4], [6], [19], [28]. This study analyzed beta-band connectivity in frontal and central regions (F3, C3, Cz, C4, F4, FC3, FC4) identified in prior literature as abnormal in PD [1], [2], [4], [25]. We evaluated whether phase-locking value (PLV) and coherence from these regions could distinguish PD patients from controls and serve as features for machine learning models [10], [22].

II. Materials & Methods

A. EEG Dataset

This study utilized an open-access EEG dataset originally published by Anjum et al. [31], developed for the LEAPD (Linear Predictive Coding EEG Algorithm for Parkinson's Disease) study. The full dataset includes recordings from 41 Parkinson's disease (PD) patients and 41 demographically-matched healthy controls, collected at two institutions: the University of New Mexico (UNM) and the University of Iowa. For the present analysis, only the Iowa cohort was used, which consists of 14 PD patients and 14 healthy controls, all recorded during a resting-state, eyes-open condition. EEG recordings were acquired using a 64-channel BrainVision system, sampled at 500 Hz, with a hardware bandpass filter of 0.1–100 Hz. The online reference electrode was Pz in the Iowa dataset. As this electrode was used as a reference, its data was not available, resulting in 63 usable EEG channels for all subjects. During preprocessing, two subjects (PD1681 and Control1411) were excluded due to technical issues, including incomplete recordings and high artifact contamination. Thus, the final dataset comprised 26 participants (13 PD and 13 control). Eye-blink artifacts were removed using Independent Component Analysis (ICA). Only the eyes-open condition was analyzed in this study, as it provides a more stable and informative neural baseline for resting-state EEG connectivity analysis [31].

B. Preprocessing

We specifically focused on the beta frequency band (12–35 Hz), as it has been widely associated with motor control and is known to be affected in Parkinson's disease [1], [5], [25]. We applied a bandpass filter to the EEG signals, removing lower

(delta, theta, alpha) and higher (gamma) frequency components. This ensured that all subsequent connectivity analyses, such as PLV and coherence, reflected only beta-band dynamics [8], [28]. We focused on electrode pairs in the frontal and central scalp regions, specifically F3, C3, Cz, C4, F4, FC3, and FC4. These were chosen based on prior studies that identified significant beta-band connectivity disruptions in motor and premotor areas in Parkinson's disease patients. Rather than limiting the analysis to a single electrode pair, we evaluated connectivity across multiple pairwise combinations within this region of interest. The pairs included: F3–C3, F3–Cz, F3–C4, F3–F4, F3–FC3, F3–FC4, C3–Cz, C3–C4, C3–F4, C3–FC3, C3–FC4, Cz–C4, Cz–F4, Cz–FC3, Cz–FC4, C4–F4, C4–FC3, C4–FC4, F4–FC3, F4–FC4, and FC3–FC4. These electrodes correspond to cortical areas involved in motor initiation, sensorimotor integration, and interhemispheric coordination—all of which are impacted in PD [4], [6], [19].

C. Connectivity Measures

1) Phase Locking Value (PLV): To assess functional connectivity based on phase synchronization between EEG channels, we employed the Phase-Locking Value (PLV) metric [7], [8]. PLV quantifies the consistency of phase differences between two time-series signals over time. The Hilbert transform was applied to each channel to obtain the instantaneous phase. PLV was then computed pairwise between all EEG channels according to the following formula:

$$\mathrm{PLV}_{i,j} = \left| \frac{1}{N} \sum_{n=1}^{N} e^{\,i(\phi_1(n)-\phi_2(n))} \right| \tag{1}$$

where $\phi_1(n)$ and $\phi_2(n)$ denote the instantaneous phases of channels i and j, respectively, and N represents the total number of time points. Equation 1 [7]. and further discussed by Aydore et n was first introduced by Lachaux et al. al. [8].

PLV values range from 0 (indicating no phase synchronization) to 1 (perfect phase locking). Individual PLV matrices were computed for each subject, separately for the Parkinson's Disease (PD) and Control groups, and used in subsequent statistical and machine learning analyses.

2) Coherence: Magnitude-squared coherence was used to evaluate functional connectivity in the frequency domain [18], [29]. Coherence reflects the linear correlation between two signals at specific frequencies, accounting for both phase and amplitude relationships. Coherence between each pair of EEG channels was estimated using Welch's method, which involves segmenting the signal, applying windowing, and averaging the periodograms. The coherence function $C_ij(f)$ was defined as:

$$\mathrm{C}_{ij} = \frac{|S_{ij}(f)|^2}{S_{ii}(f)\, S_{jj}(f)} \tag{2}$$

where Sij(f) denotes the cross-spectral density between channels i and j, and Sii(f), Sjj(f) are their respective auto-spectral densities. This formulation is based on the method originally proposed by Welch [29]. and widely used in EEG coherence studies [19], [30]. Coherence values were computed over the beta frequency band and averaged to obtain a single coherence value per channel pair. These values were used to generate subject-level coherence matrices for both PD and Control groups, which were then included in further analyses such as heatmap visualizations and classification modeling.

D. Connectivity Analysis

1) Heatmap Generation: Heatmaps based on Phase-Locking Value (PLV) and coherence were generated to visualize functional connectivity between brain regions. Color intensity indicates connectivity strength between EEG electrode pairs, offering a spatial view of inter-regional synchronization [7], [29], [31]. PLV measures phase consistency [7], [8], while coherence reflects linear frequency-domain correlation [18], [29]. The resulting symmetric, square matrices have axes corresponding to the selected EEG channels [4], [20], [31]. Group-level connectivity differences were examined by generating separate PLV and coherence heatmaps for PD and control groups [10]. Individual connectivity matrices were averaged to produce whole-brain maps (all EEG channels) and focused maps (motor-related electrodes: F3, C3, Cz, C4, F4, FC3, FC4) [2], [4], [6], [31]. In total, eight heatmaps were created PLV and coherence for both whole brain and motor regions in each group enabling both global and region-specific comparisons of functional connectivity.

2) Statistical Analysis: To identify statistically significant differences in functional connectivity between the Parkinson's disease (PD) and control groups, independent two-sample t-tests were performed on precomputed connectivity features [10], [18], [28]. The specific channel pairs were selected based on their relevance to motor and premotor brain regions [2], [4], [6], [12], [28]: F3–C3, F3–Cz, F3–C4, F3–F4, F3–FC3, F3–FC4, C3–Cz, C3–C4, C3–F4, C3–FC3, C3–FC4, Cz–C4, Cz–F4, Cz–FC3, Cz–FC4, C4–F4, C4–FC3, C4–FC4, F4–FC3, F4–FC4, and FC3–FC4. This yielded a p-value for each connection in both PLV and coherence datasets, enabling the identification of statistically meaningful group differences [10]. The p-values obtained from t-tests were used to determine the statistical significance of group-level differences in functional connectivity between the Parkinson's disease (PD) and control groups. For each selected electrode pair, a p-value was calculated separately for both phase-locking value (PLV) and coherence measures [7], [8], [18], [29]. A significance threshold of $p < 0.05$ was applied. Channel pairs that yielded p-values below this threshold were considered to exhibit statistically significant differences in beta-band connectivity between the two groups. These results were later visualized using bar plots and box plots to highlight key connections with potential diagnostic relevance [31]. For visual representation of the differences in connectivity strength for selected connections, box plots and bar plots were generated. These plots provided an intuitive view of distribution differences across the groups, particularly for those with significant or near-significant p-values [10], [31].

E. Machine Learning Classification

To increase dataset size and maintain uniformity, each EEG recording from PD and control subjects was divided into non-overlapping 20-second epochs [10], [22]. Segments shorter

than 20 seconds at the end of recordings were discarded. This ensured fixed-length inputs, improving the robustness and reliability of subsequent machine learning classification [22]. This stratiied split ensured that samples from both Parkinson's disease (PD) and control groups were proportionally represented in both the training and testing sets. By keeping the class distribution balanced across the split, we aimed to minimize bias and obtain more reliable performance metrics during classification [10], [24]. For the classification task, we employed six supervised machine learning algorithms: Fine Decision Tree, Linear Discriminant Analysis (LDA), Naive Bayes, Linear Support Vector Machine (SVM), Fine K-Nearest Neighbors (KNN), and a Narrow Neural Network. These models were selected to capture a diverse range of decision boundaries—from linear to highly nonlinear patterns—while comparing their performance in distinguishing between Parkinson's disease (PD) and control EEG segments [10], [22], [24]. Moreover, various features of EEG data have been successfully utilized in related applications [11]–[17].

III. Results

A. Heatmap Analysis

1) Whole-Brain PLV and Coherence: When examining whole-brain EEG connectivity, distinct patterns emerged in both PLV and coherence analyses between the Parkinson's disease (PD) and control groups [4], [10], [31]. Visual inspection of whole-brain PLV heatmaps reveals slight connectivity differences. In the Parkinson's disease (PD) group ((Fig. 1)), more yellow regions indicate stronger beta-band phase synchronization between certain EEG channel pairs. The control group shows more evenly distributed and overall lower connectivity. Visual inspection of whole-brain coherence heatmaps shows distinct differences: the control group has more high-intensity (yellow) regions, indicating stronger overall coherence, whereas the PD group displays fewer yellow areas, reflecting reduced coherence. When focusing on motor-related EEG regions (F3, C3, Cz, C4, F4, FC3, FC4), different patterns emerged in PLV and coherence analyses between Parkinson's disease (PD) and control groups [2], [4], [6], [19].

2) Focused Region PLV and Coherence: In the PLV heatmaps, only slight differences were observed; the PD group exhibited a mild increase in phase synchronization in some connections, but the overall distribution remained comparable to the control group (Fig. 2). In the coherence heatmaps, a distinct difference was observed between the control and Parkinson's disease (PD) groups. The PD group showed a marked reduction in these yellow regions, reflecting a substantial decrease in functional connectivity. This decline was particularly evident in the focused motor-related areas (F3, C3, Cz, C4, F4, FC3, FC4), suggesting a disruption in inter-regional communication within motor networks commonly affected by Parkinson's disease (Fig. 3) [3], [4], [6], [20].

B. Statistical Test Results

T-test results revealed significant group differences in PLV (Cz–F4, p = 0.0073; Cz–FC3, p = 0.0213) and coherence

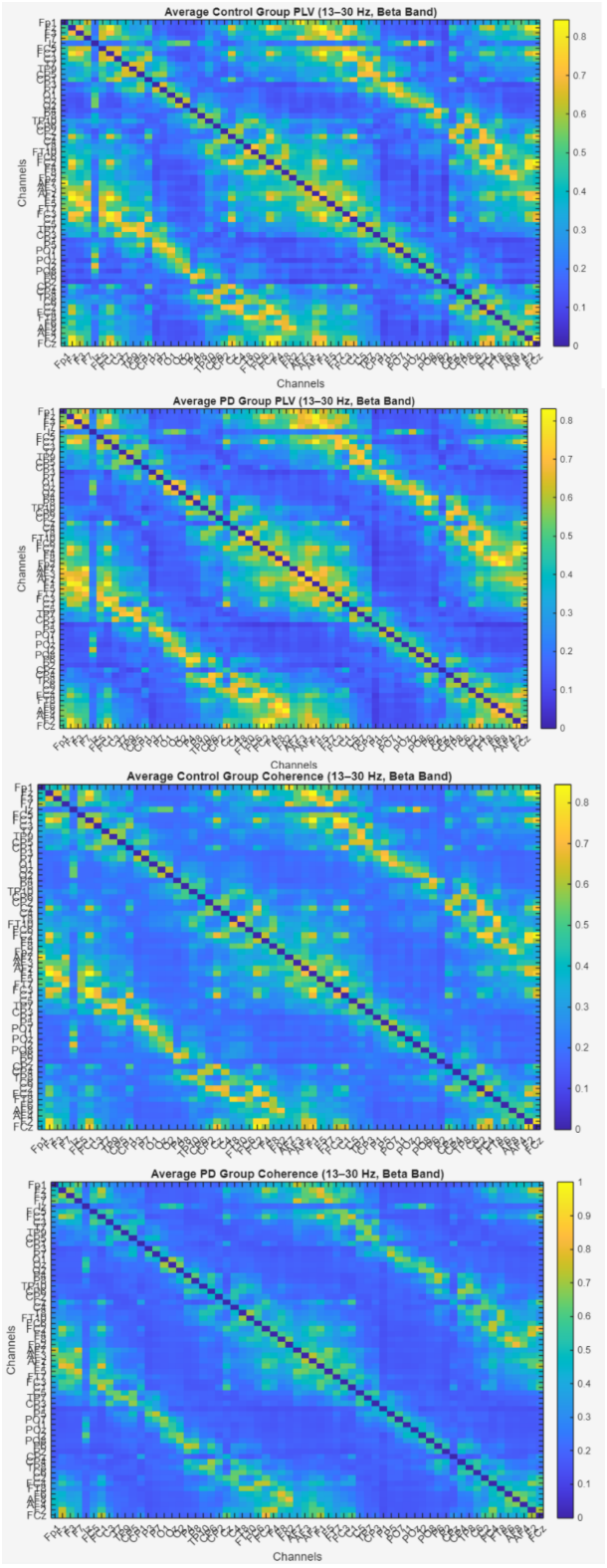

Figure 1: a)Average Control Group PLV Beta Band b)Average PD Group PLV Beta Band c)Average Control Group Coherence Beta Band d)Average PD Group Coherence Beta Band

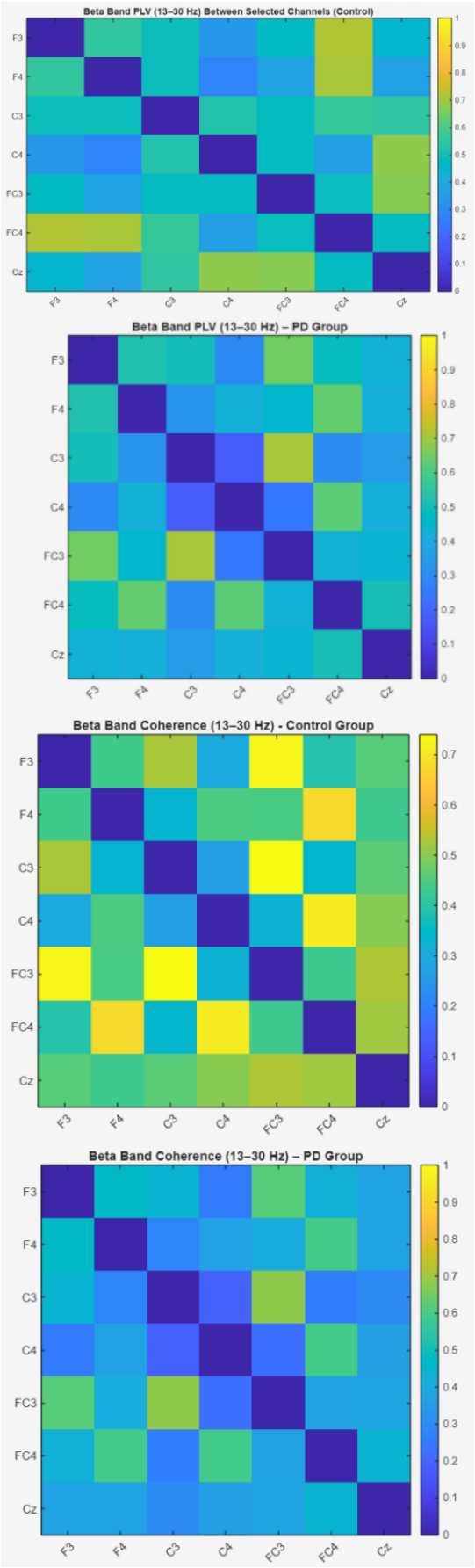

Figure 2: a)Average Control Group PLV Beta Band b)Average PD Group PLV Beta Band c)Average Control Group Coherence Beta Band d)Average PD Group Coherence Beta Band

(F3–C3, p = 0.0053; F3–Cz, p = 0.0223; F3–FC4, p = 0.0340; Cz–C4, p = 0.0154; Cz–F4, p = 0.0316), highlighting altered motor-related connectivity in Parkinson's disease. Bar plots were created for each PLV and coherence feature, showing mean connectivity values for Parkinson's disease and control groups across selected channels, with red dots marking significant differences ($p < 0.05$) and error bars showing standard deviation. These plots complement statistical tests by visually highlighting group differences, especially for features with significant results. Boxplot analysis revealed significantly higher PLV in Parkinson's disease patients for Cz–F4 (p = 0.0073) and Cz–FC3 (p = 0.0213) links, indicating abnormal hypersynchronization in frontocentral regions that may relate to reduced motor flexibility. Coherence analysis revealed significantly reduced connectivity in Parkinson's disease patients for F3–C3 (p = 0.0053), F3–Cz (p = 0.0223), F3–FC4 (p = 0.0340), and Cz–C4 (p = 0.0154), indicating impaired communication in motor and executive networks. Cz–F4 also showed reduced coherence (p = 0.0316), aligning with PLV findings and suggesting consistent disruptions across measures.

C. Machine Learning Results

Using PLV and coherence features, several machine learning models were tested to distinguish Parkinson's disease from controls. The Fine Tree classifier achieved the highest test accuracy (84.72%), outperforming Linear SVM (70.83%), Fine KNN (76.39%), Neural Network (73.61%), Naive Bayes (69.44%), and LDA (68.06%).

IV. DISCUSSION

This study compared beta-band functional connectivity in Parkinson's disease (PD) and healthy controls using EEG-derived PLV and coherence. PD patients showed increased PLV, particularly in frontocentral regions, indicating pathological hypersynchronization related to motor rigidity and bradykinesia. In contrast, coherence was broadly reduced, especially in fronto-motor connections, reflecting impaired large-scale network integration and reduced coordination associated with disease progression [2], [4], [6], [20], [21]. Statistical testing revealed significant PLV (e.g., Cz–F4) and coherence (e.g., F3–C3) alterations in regions linked to motor planning, sensorimotor integration, and executive control, functions typically impaired in PD [3], [6], [18]. Visualization techniques such as box plots and heatmaps highlighted connectivity differences, a common neuroengineering approach to reveal alterations hidden in raw EEG data [8]. Machine learning classifiers (e.g., Fine Decision Tree, KNN, SVM) based on these EEG features showed promising diagnostic potential, with the Fine Tree achieving 84.72% accuracy, consistent with previous studies using synchronization measures for neural disorder detection [10], [22], [24]. Combining spectral (coherence) and phase-based (PLV) metrics offers a fuller view of connectivity changes, as coherence captures linear relationships while PLV reflects phase consistency [7], [19].

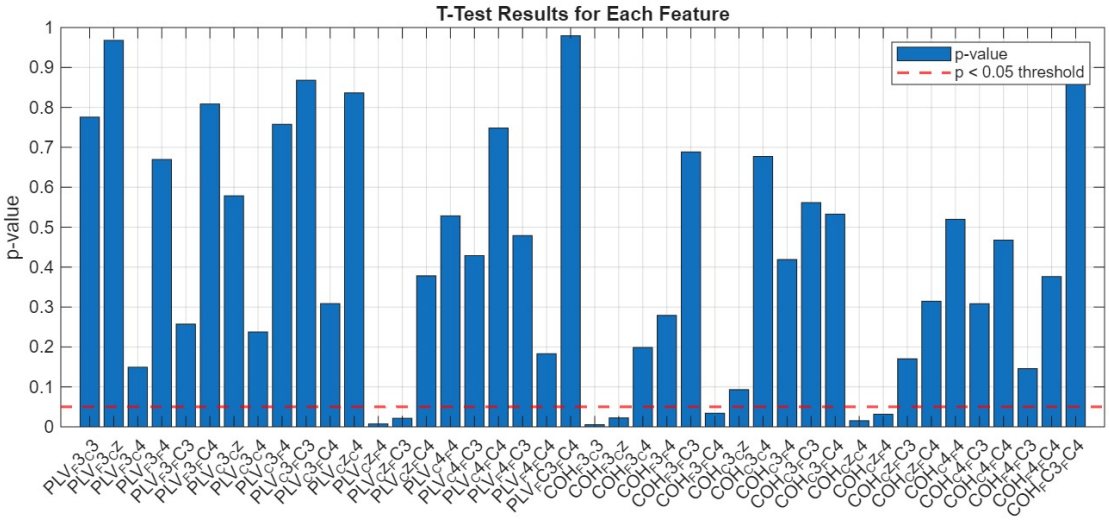

Figure 3: Focused-Region PD Group Coherence Beta Band

V. CONCLUSIONS

This study demonstrates that beta-band EEG connectivity (increased PLV, decreased coherence) can effectively differentiate Parkinson's disease patients from healthy controls, particularly in motor-related regions. Significant frontocentral alterations and strong Fine Decision Tree classification performance highlight the potential of these metrics for non-invasive diagnosis and real-time clinical decision support. Future research should assess their longitudinal changes and integrate EEG with modalities like MRI to enhance diagnostic precision

AUTHOR CONTRIBUTIONS

This paper is an original research article by two authors. Sevval Savas, preparing and organizing the manuscript and performing statistical analysis, feature extraction and selection, and code block development on Matlab. Ibrahim Kaya: finding the dataset, supervision of the research, and edits.

REFERENCES

[1] A. Oswal, P. Brown, and V. Litvak, "Synchronized neural oscillations and the pathophysiology of Parkinson's disease," Brain, vol. 139, no. 2, pp. 518–530, 2016.

[2] A. A. Kühn et al., "Pathological synchronisation in the subthalamic nucleus of patients with Parkinson's disease relates to both bradykinesia and rigidity," Experimental Neurology, vol. 215, no. 2, pp. 380–387, 2009.

[3] A. Singh et al., "Frontal theta and beta oscillations during lower-limb movement in Parkinson's disease," Clinical Neurophysiology, vol. 131, no. 3, pp. 694–702, 2020.

[4] J.-P. Lachaux, E. Rodriguez, J. Martinerie, and F. J. Varela, "Measuring phase synchrony in brain signals," Human Brain Mapping, vol. 8, no. 4, pp. 194–208, 1999.

[5] S. Aydore, M. Pantazis, and R. M. Leahy, "A note on the phase locking value and its properties," NeuroImage, vol. 74, pp. 231–244, 2013.

[6] C. de Hemptinne et al., "Therapeutic deep brain stimulation reduces cortical phase-amplitude coupling in Parkinson's disease," Nature Neuroscience, vol. 18, no. 5, pp. 779–786, 2015.

[7] A. Mishra et al., "A machine learning approach for detection of Parkinson's disease using phase locking value and EEG signals," Biomedical Signal Processing and Control, vol. 65, Art. no. 102350, 2021.

[8] M. E. Bastos and P. Fries, "Rhythms for cognition: Communication through coherence," Neuron, vol. 88, no. 1, pp. 220–235, 2015.

[9] B. C. M. van Wijk, J. D. Beudel, M. J. J. van den Heuvel, C. M. J. van Laar, and C. F. van Putten, "Low-beta cortico-pallidal coherence decreases during movement in Parkinson's disease," Journal of Physiology, vol. 594, no. 5, pp. 1539–1551, 2016.

[10] M. Olde Dubbelink et al., "Disrupted brain network topology in Parkinson's disease: A longitudinal magnetoencephalography study," Brain, vol. 137, no. 1, pp. 197–207, 2014.

[11] M. Y. Özdemir and Y. İşler, "Nonlinear feature analysis for EEG-based biometric authentication via machine learning," Journal of Intelligent Systems with Applications, vol. 6, no. 2, pp. 27–33, 2023.

[12] A. Bulut, G. Ozturk, and I. Kaya, "Classification of sleep stages via machine learning algorithms," Journal of Intelligent Systems with Applications, vol. 5, no. 1, pp. 66–70, 2022.

[13] M. Degirmenci, Y. K. Yuce, and Y. Isler, "Motor imaginary task classification using statistically significant time domain and frequency domain EEG features," Journal of Intelligent Systems with Applications, vol. 5, no. 1, pp. 49–54, 2022.

[14] M. Degirmenci, Y. K. Yuce, and Y. Isler, "Classification of multi-class motor imaginary tasks using Poincare measurements extracted from EEG signals," Journal of Intelligent Systems with Applications, vol. 5, no. 2, pp. 74–78, 2022.

[15] M. Degirmenci, Y. K. Yuce, M. Perc, and Y. Isler, "Classification of finger movements from statistically-significant time-domain EEG features," Journal of the Faculty of Engineering and Architecture of Gazi University, vol. 39, no. 3, pp. 1597-1610, 2024.

[16] M. Degirmenci, Y. K. Yuce, M. Perc, and Y. Isler, "EEG channel and feature investigation in binary and multiple motor imagery task predictions," Frontiers in Human Neuroscience, vol. 18, article 1525139, 2024.

[17] M. Degirmenci, Y. K. Yuce, M. Perc, and Y. Isler, "Statistically significant features improve binary and multiple Motor Imagery task

predictions from EEGs," Frontiers in Human Neuroscience, vol. 17, article 1223307, 2023.

[18] M. Stoffers et al., "Slowing of oscillatory brain activity is a stable characteristic of Parkinson's disease without dementia," Brain, vol. 130, no. 7, pp. 1847–1860, 2007.

[19] J. Pereira, H. X. Xie, D. Cunha, and D. P. C. Vieira, "Parkinson's disease classification using machine learning and resting-state EEG," in Proc. IEEE EMBC, Honolulu, HI, USA, pp. 2048–2051, 2018.

[20] H. Abásolo, R. Hornero, D. García, M. López, and C. Gómez, "Nonlinear analysis of EEG background activity in Alzheimer's disease patients," Physiological Measurement, vol. 27, no. 4, pp. 373–383, 2006.

[21] F. Trambaiolli, A. C. Sato, P. S. Balardin, and J. A. P. de Oliveira, "Functional connectivity and machine learning: A review for clinical neuroimaging," IEEE Reviews in Biomedical Engineering, vol. 13, pp. 135–148, 2020.

[22] J. Jenkinson and M. Brown, "New insights into beta oscillations in Parkinson's disease," Brain, vol. 134, no. 3, pp. 623–636, 2011.

[23] J. C. Brittain and P. Brown, "Oscillations and the basal ganglia: Motor control and beyond," NeuroImage, vol. 85, pp. 637–647, 2014.

[24] A. Little and P. Brown, "The functional role of beta oscillations in Parkinson's disease," Lancet Neurology, vol. 13, no. 6, pp. 557–559, 2014.

[25] D. M. Palva and S. Palva, "Functional connectivity of neuronal oscillations in the human cortex," NeuroImage, vol. 61, no. 4, pp. 1035–1046, 2012.

[26] P. Welch, "The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms," IEEE Trans. Audio Electroacoust., vol. 15, no. 2, pp. 70–73, 1967.

[27] T. Brunner et al., "Coherence analysis of intracranial EEG signals reveals focal seizure networks," Clinical Neurophysiology, vol. 125, no. 1, pp. 129–138, 2014.

[28] D. Bassett and E. Bullmore, "Small-world brain networks," The Neuroscientist, vol. 12, no. 6, pp. 512–523, 2006.

[29] B. H. Fleisher et al., "Deep learning-based EEG classification system for Parkinson's disease," Scientific Reports, vol. 12, pp. XX–XX, 2022.

[30] T. Roy, A. Garg, and P. Prasad, "A hybrid deep learning model for detection of Parkinson's disease using EEG signals," IEEE Access, vol. 9, pp. 103960–103969, 2021.

[31] M. F. Anjum, S. Dasgupta, R. Mudumbai, A. Singh, J. F. Cavanagh, and N. S. Narayanan, "Linear predictive coding distinguishes spectral EEG features of Parkinson's disease," Parkinsonism Relat. Disord., vol. 79, pp. 79–85, 2020.