Contribution of Different Movement Tasks to Differential Diagnosis of Parkinson's Disease

Attila Zoltán Jenei[1], Dávid Sztahó[1]

***Abstract*—Parkinson's disease is one of the most common neurological diseases, which, according to current knowledge, is incurable. Early detection is essential since, with appropriate therapy and medication, the progress of the disease can be slowed down and the quality of life maintained. The movement tasks described with the acceleration data are part of the intensive research area. This would make recognizing the disease and specific symptoms like tremors, rigidity, and bradykinesia possible. Many research studies focus on the selection of appropriate movement tasks. However, due to the diversity of the studies, no consensus has yet been reached. Therefore, in this research, we examine which of the movement tasks selected from the Unified Parkinson's Disease Rating Scale prove to be the best under the same procedure. Furthermore, we attempted to make a final decision based on the predictions obtained on the movement forms using voting procedures. 37 patients with Parkinson's disease and 47 healthy individuals participated in this study. 3-axial acceleration data from the wrist-mounted sensor was acquired, from which times-series features were determined. Binary classifications were done by Support Vector Machine (SVM). Soft, hard, and SVM-based prediction fusions were also explored. The results indicated that the PRONATION task has the highest balanced accuracy (76.2%). Among the voting approaches, the soft achieved the highest balanced accuracy (79.8%) compared to the best task. Based on the results, fewer movement tasks can be used to recognize the disease, among which PRONATION is essential. Further voting approaches can improve the performance.**

***Index Terms*—Acceleration, Body sensor, Classification, Parkinson's disease, Support Vector Machine, Voting ensemble**

I. INTRODUCTION

Parkinson's disease (PD) is one of the most common neurological diseases, which typically affects the ageing population. According to current clinical knowledge, PD is incurable, highlighting the importance of early recognition and personalized treatment.

It influences 1-2 people out of 1,000, but its prevalence increases with age: 1% of the population over 60 may be affected [1]. Other risk factors are male sex, environmental circumstances (industrialization, toxic chemicals), and genetics.

The disease occurs with the destruction of dopamine-producing neurons in the substantia nigra pars compacta region. When the first motor symptoms appear, up to 50-70% of the dopamine-producing cells have already died. Intraneuronal protein aggregates, so-called Lewy bodies, can also be observed. [2]

Concerning the disease, motor and non-motor symptoms can be distinguished. Non-motor symptoms can appear up to 10 years earlier than motor symptoms [3]. Such non-motor symptoms can be daytime sleepiness, difficulty sleeping, digestive problems, loss of smell, and pain. Depression and anxiety can often develop even before the actual diagnosis. These symptoms may persist and worsen even with the appearance of motor symptoms. [4]

The main motor symptoms are tremors at rest, bradykinesia, and rigidity [5]. These are the key symptoms that are used to diagnose the disease. Examining additional symptoms can be supportive (exclusion of other disease with similar symptoms). The neurologist primarily visually inspects these symptoms and assesses the patient's condition. Medication, labor tests and imaging procedures can also help establish a diagnosis [6].

Multiple severity measures are available to recognize and assess PD. One of the most commonly used scales for the clinical assessment of Parkinson's disease is the *Unified Parkinson's Disease Rating Scale* (UPDRS), with four main parts: 1) assessment of non-motor experiences during everyday life, 2) motor experiences during everyday life, 3) examination of motor symptoms, 4) motor complications [7]. Scores between 0 (means a healthy [normal] stage) and 4 (means the severe stage) can be given for each task.

Much scientific literature uses motor symptoms to support the diagnosis using machine learning algorithms. This may provide an opportunity for early disease detection and a semi-objective evaluation. Modalities capturing such motor symptoms are speech [8][9], writing/drawing [10][11] or limb movements [12][13].

Our research investigates **which form of hand movement effectively recognizes patients** with mild symptoms **and whether combining these forms of movement improves recognition.**

We detail the key findings in the *Literature Study* section and describe the experiment process in the *Methodology* section. In the *Results* section, we present the results respecting the hypothesis, and finally, in the *Conclusion* section, we summarise the work and the key findings of our research.

[1]Department of Telecommunications and Media Informatics, Faculty of Electrical Engineering and Informatics, Budapest University of Technology and Economics, Budapest, Hungary (e-mail: jenei.attila.zoltan@vik.bme.hu, sztaho.david@vik.bme.hu; orcid id: 0000-0003-1007-9907, 0000-0002-7361-4260).

https://doi.org/10.3311/WINS2024-011

II. Literature Study

Limb movement data is primary for diagnosing the disease since motor symptoms interfere with regulating and coordinating movement. Some researchers identify the movement process with the help of a video camera and indicate if there is an abnormality (irregular movement or fall) [14]. However, the most common approach is the set of signals provided by wearable sensors. One such device is a wrist-mounted sensor that provides acceleration data [15].

Moore et al. [16] implemented freezing of gait (FoG) recognition with 11 PD patients wearing an ankle-mounted sensor. This sensor transmitted acceleration data wirelessly. They detected 78% of FoG events, but the system also classified 20% of stands as FoG. Individual calibration improved the detection rate to 89%, with 10% false positive cases.

Lonini's research [15] examined the necessary amount of sensors for PD recognition. 20 PD patients were recorded in one day and again after two weeks with six flexible wearable sensors. Participants performed 13 tasks that clinicians assessed according to bradykinesia and tremor. With the help of Convolutional Neural Networks (CNN), they pointed out that just one sensor is sufficient to capture the two symptoms, and this is the sensor placed on the back of the hand. They also pointed out that recording the same person several times did not improve the recognition of symptoms (and the disease).

Shawen and his colleagues [17] used skin-mounted and smart watch sensors to detect the disease. 13 PD patients were included in the research, who had to perform standard motor tasks. Clinicians assessed tremors and bradykinesia during task performance. With their results, they pointed out that the type of sensor did not significantly affect recognition. Also, it was shown that reducing the sampling frequency worsened the recognition of tremors (under 30 Hz), while that of bradykinesia did not.

Talitckii and his colleagues [18] investigated the optimal sets of movements to detect PD. A total of 15 tasks were developed for the study, which can be grouped into the following four categories: 1) *Gross Motor* (everyday activities), 2) *Clinical Evaluation* (activities that clinicians use), 3) *Fine Motor* (require fine motor coordination) and 4) *Tremor at Rest* (activities to reveal tremor). The data was recorded at 100 Hz with a sensor attached to the hand. The database included 42 patients with PD, 17 patients with other neurological diseases associated with tremors, and 24 individuals as healthy controls (HC). Dimensionally reduced standard and spectral features were used with multiple machine learning algorithms. They pointed out that walking and sitting down are essential forms of movement in recognizing the two symptoms, and additional movements that can help recognition can be defined.

Rovini et al. [19] studied 40 healthy and 40 PD individuals who wore sensors on all four limbs. Triaxial accelerations and triaxial angular rates were acquired from these sensors. Tasks were selected from several rating scales, including the UPDRS. Such movements for the upper limbs are finger tapping (3.4), hand movements (3.5), pronation-supination movements of hands (3.6), hand rest tremor amplitude (3.17) and postural tremor of hands (3.15). For the lower limbs, toe-tapping (3.7), leg agility (3.8) and gait (3.10) were recorded from the UPDRS scale. The tests achieved an average accuracy of 93.6%-96.0%. 100% accuracy was also achieved when signals from all limbs were used with a Support Vector Machine (SVM).

Albert and his colleagues [20] examined 18 HC and 8 PD patients regarding recognizing different forms of movement (walking, standing, sitting, holding, or not wearing the phone). Standard characteristics were extracted from the time series and classified using machine learning algorithms (SVM, regularised logistic regression). Their results pointed out that the movement patterns of PD patients can be separated from each other with a lower performance (75.1% accuracy) than that of healthy people (86.0% accuracy).

Based on the results, it can be seen that many movement tasks are widespread in the recognition of PD. The most common of these are defined by severity measurement scales like UPDRS. This is to be expected since these tasks were created so that the symptoms of PD are captured as efficiently as possible. Furthermore, it can be observed that the combined use of several forms of movement helps recognition. At the same time, *there is no consensus on whether the combination of sensors significantly improves performance (how many sensors are sufficient?)*. However, there are also several attempts in other areas to fuse sensor data, like clinical imaging [21], radar imaging [22], or acceleration and video data [23]. These studies report better recognizability by data fusion. The comparison can be difficult due to the variation in the position of the sensors, the amount and type of data, the nature of the processing, and the classification method.

Therefore, with our research, we wish to contribute to this field by examining several forms of movement separately and in combination. The forms of movement were selected based on the UPDRS scale because of what was discussed in previous research. We want to investigate ***1)*** *which form of movement carries the most information for the recognition of the disease* and ***2)*** *whether the combined use of the forms of movement improves the recognition of the disease*.

III. Methodology

A. Database

The database contained movement data of 37 PD patients and 47 HC persons. All participants were informed in advance about the research details and gave their consent by signing a signed consent form.

There are 33 men and four women in the PD class. Their average age is 63.9 years, with a standard deviation of 14.3 years. 18 people were taking medication, and 14 people were receiving deep brain stimulation (DBS) (8 people belonged to both categories). The severity score of PD patients was determined using selected tasks from the UPDRS. Resting tremor (3.17), postural tremor (3.15), rigidity (3.3), and finger tapping (3.4) tasks were evaluated by the neurologist. The mean severity was 1.00. Severity scores for the right hand were given, as data from the right hand were examined. The neurologist observed at least right-sided symptoms in all patients.

There were 20 men and 27 women in the HC class. Their average age was 56.7 years, standard deviation of 15.2. According to their acknowledgement, they did not have Parkinson's disease or any other disease affecting their movement.

Acceleration data (in X, Y, and Z directions) was recorded using a wrist-mounted sensor (Mbientlab MMS) with a sampling frequency of 50 Hz. The forms of movement were selected based on the UPDRS scale[1]: gait (3.10), kinetic tremor of the hands (3.16), postural tremor of the hands (3.15), rest tremor amplitude (3.17), pronation-supination (3.6) movements of the hands, and drawing an Archimedes spiral on a tablet using a template. The last task is not a part of the UPDRS but is commonly used in PD detection.

B. Preprocessing and feature extraction

We subtracted the average from the data series in all three directions of each recording so that they moved around a uniform value of 0. Then, we calculated vector lengths (l_t) from the time series along the time domain (t) using Eq. 1. So, the resultant vector had the same length as the X, Y, and Z data. In the equation, x_t, y_t, z_t are the accelerations in the 3D coordinate system at a specific t timepoint.

$$l_t = \sqrt{x_t^2 + y_t^2 + z_t^2} \tag{1}$$

In the literature, manual features are extracted from the time series or image representations are created to benefit convolutional neural network automatic feature extraction. In this research, we explored the *Time Series Feature Extraction Library* (*TSFEL v0.1.4)* [24] Python package. Features were defined using the *time_series_features_extractor* with basic settings. This resulted in 60 different features in the statistical, temporal and spectral domain as a total of 390 descriptors[2]. These features are not specific for PD but act as a general description of the recordings. Among these features, those that did not contain a value for any subject (NaN) or those with a constant value for all subjects were removed. After that, the feature vectors were normalized between 0 and 1.

C. Classification and model evaluation

SVM algorithms with linear and radial basis function (rbf) kernels were used to classify PD and HC. The other parameters of the models were left at the basic settings according to *sklearn svm.SVC* [25]. With this setting, the algorithm provides generalized performance across movement tasks and also serves as a baseline for parameter optimization. The evaluation was carried out using the leave-one-out cross-validation (LOOCV) procedure, which maximizes the training set's size and provides an average performance across the whole dataset. Sensitivity, specificity, balanced accuracy and area under the curve (auc) were used to describe the model performances. Receiver Operating Characteristic (ROC) curves were generated for all experiments to give visual insight into the threshold settings.

Two experiments were tested according to the hypothesis mentioned in section II:

a) the use of movement forms separately for classification,

b) fusion of the predictions obtained on the movement forms with voting approaches.

Soft, hard and SVM-based voting were explored in the examination *b)*. In the soft voting case, the predictions per subject were averaged as a final prediction. In the hard voting, class assignment was done for each form of movement (HC or PD) first, and then the decision was made for the majority class. In the case of SVM-based voting, the input was the vectors of the predictions obtained for the movements by the previous models. Using these estimates as "*features*", we made a final prediction also using linear and rbf kernels.

IV. RESULTS

A. Separate classification of movement forms

TABLE I shows the results achieved with different forms of movement. The table can be divided horizontally into the linear and rbf kernels results. The rows present the different movement forms (the names correspond to the tasks in section III, subsection A). The last four columns describe the sensitivity (sens), specificity (spec), balanced accuracy (bacc) and auc value, respectively. The best results are marked with **bold** style.

TABLE I
CLASSIFICATION OF DIFFERENT MOVEMENT TASKS WITH LINEAR AND RBF KERNEL SVM.

kernel	movement form	sens	spec	bacc	auc
linear	GAIT	51.4%	78.7%	66.7%	0.680
	KINETIC	73.0%	76.6%	75.0%	0.745
	POSTURAL	75.7%	63.8%	69.0%	0.755
	REST	64.9%	70.2%	67.9%	0.809
	PRONATION	**75.7%**	**76.6%**	**76.2%**	**0.809**
	SPIRAL	56.8%	63.8%	60.7%	0.704
rbf	GAIT	62.2%	78.7%	71.4%	0.745
	KINETIC	78.4%	68.1%	72.6%	0.806
	POSTURAL	100.0%	42.6%	67.9%	0.781
	REST	75.7%	57.4%	65.5%	0.766
	PRONATION	**75.7%**	**76.6%**	**76.2%**	**0.800**
	SPIRAL	59.5%	70.2%	65.5%	0.737

The PRONATION achieved the highest performance with either linear or rbf kernel (76.2% bacc). The spiral movement representation gave the lowest result: 60.7% bacc with linear kernel and 65.5% bacc with rbf kernel. With the other forms of movement, we achieved results between these extreme values. For the linear kernel, the averaged sensitivity for movement forms is

[1] Detailed description of the scale is available here (last accessed: 22th of January, 2024): https://www.movementdisorders.org/MDS-Files1/Resources/PDFs/MDS-UPDRS.pdf

[2] List of the features can be seen here (last accessed: 22th of January, 2024): https://tsfel.readthedocs.io/en/latest/descriptions/feature_list.html

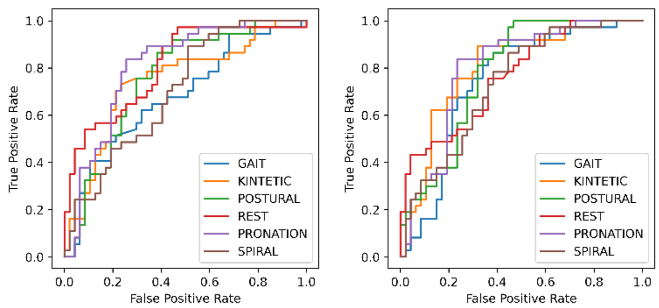

Fig. 1. ROC curves from the predictions of the linear (left) and rbf (right) kernel models. The different colors represent the different forms of movement.

66.2%, the specificity is 71.6%, and the auc value is 0.750. With the rbf kernel, the average sensitivity is 75.2%, the specificity is 65.6%, and the auc value is 0.772. The average bacc was almost the same for both kernels: 69.2% for linear and 69.8% for rbf kernel. Based on these, it can be seen that the model with a linear kernel biased the samples more in the direction of the HC class, while the rbf kernel biased the samples more in the direction of the PD class. The latter case may be more desired in the health care because the false negative error has higher impact and cost for the patient life.

Fig. 1 shows the development of ROC curves with linear and rbf kernels. The *True Positive Rate* (sensitivity) is shown on the vertical axis, and the *False Positive Rate* (1 - specificity) on the horizontal axis. The colours represent the different forms of movement.

The course of the curves was similar for both kernels. In the linear case, the PRONATION and REST tasks were at the top. They also have the highest auc value (0.809). Meanwhile, the GAIT and SPIRAL can be seen in the lower area. The auc value of these two is the smallest (0.680, 0.704). Regarding the rbf kernel, PRONATION and KINETIC progress at the top with auc values of 0.800 and 0.806. In the lower part is the SPIRAL with an auc value of 0.737.

The average auc value was 0.750 for models with a linear kernel and 0.772 for rbf models. This difference represents one person (the rbf model correctly identified one person more than the linear model).

B. Voting ensembles

TABLE II summarises the voting results achieved on the predictions on the movement forms. The first column indicates whether soft, hard or SVM voting took place (in the case of SVM, both linear and rbf kernels were used). The second column shows which kernel was used to classify movement forms previously: *linear*, *rbf*, *all* (the results of both kernels were used). As in TABLE I, the last four columns show the sensitivity, specificity, balanced accuracy and the auc value. The cases where better results were achieved than PRONATION according to TABLE I (the best-performing form of movement) were marked in *italic*. The best result was marked in **bold**. The auc values were omitted for the hard voting type because it's predictions are 0 and 1 only (rather than continuous between 0 and 1).

In the case of *soft* and *hard* voting, all with one exception, achieved a better result than the best performance according to TABLE I. Comparing the balanced accuracy, the improvement is 2.8% on average. This means three people on average (another three people are correctly classified by voting process rather than using only the PRONATION).

TABLE II
CLASSIFICATION BY VOTE-BASED MERGING OF PREDICTIONS.

kernel/ voting type	previous kernel	sens	spec	bacc	auc
soft	**linear**	**83.8%**	**76.6%**	**79.8%**	**0.872**
	rbf	*83.8%*	*74.5%*	*78.6%*	*0.874*
hard	linear	64.9%	83.0%	75.0%	-
	rbf	*81.1%*	*76.6%*	*78.6%*	-
linear	all	64.9%	80.9%	73.8%	0.813
	linear	64.9%	78.7%	72.6%	0.832
	rbf	*75.7%*	*78.7%*	*77.4%*	*0.823*
rbf	all	67.6%	72.3%	70.2%	0.760
	linear	70.3%	76.6%	73.8%	0.781
	rbf	73.0%	78.7%	76.2%	0.819

Regarding the SVM voting, only one case shows an improvement (linear kernel SVM on rbf predictions) compared to PRONATION. In balanced accuracy, this means an improvement of 1.2%. This improvement represents a person (another correct recognition of a person).

Considering soft and hard voting on average, 78.4% sensitivity, 77.7% specificity and 78.0% balanced accuracy can

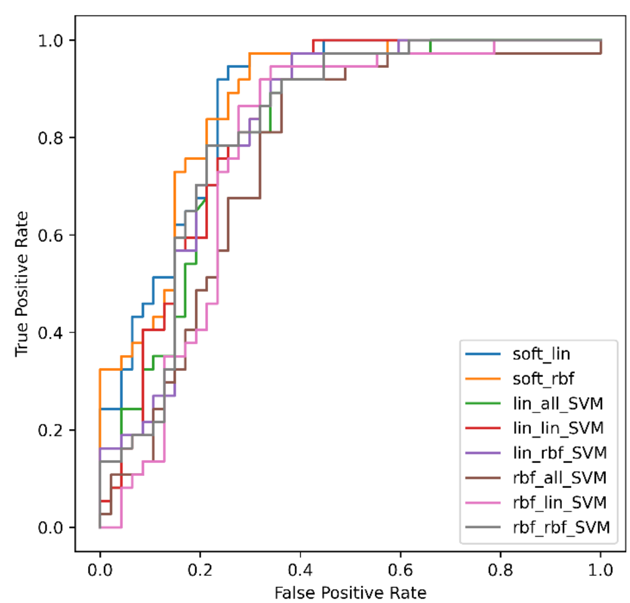

Fig. 2. ROC curves from the voting predictions. The different colors represent the different voting approaches. The names correspond to TABLE II.

be observed. SVM-based voting has an average sensitivity of 69.4%, specificity of 77.7% and balanced accuracy of 74.0%. It can be seen that the difference between the two approaches is, on average, 9% in sensitivity. In other words, the soft and hard voting approach maintained a balanced classification, while the SVM bias shifted in the healthy direction.

Fig. 2 shows the evolution of the ROC curves for the voting approaches. The markings of the axes correspond to the markings in Fig. 1. Different colors represent different voting approaches. The names of the labels correspond to the designations in TABLE II. We have omitted the hard voting approach from the plot because it outputs values 0 and 1 only (rather than continuously between 0 and 1). The figure shows that the soft voting curves are moving at the top, while the rbf kernel voting SVM is at the lower area. In a compact description, the auc values also show this: in the case of soft voting, the auc value is 0.872 and 0.874, while in the case of SVM voting, the best auc value is 0.823. When voting with the rbf kernel, the *all* and the *linear* cases take an auc value below 0.8 (this curves are at the bottom of the ROC plot).

Fig. 3 shows the importance of movement forms for the *linear kernel voting SVM*. On the left side is the importance obtained with predictions from *linear* SVM on movements separately, while on the right is the importance obtained with predictions from *rbf* SVM on movements separately. The markings of the movement tasks are according to TABLE I.

In both cases, it can be seen that PRONATION is the most essential movement task. This aligns with the results achieved per movement (where PRONATION provided the best performance). However, the two kernels differ in their additional importance: REST and SPIRAL were essential for the linear model, while GAIT and KINETIC were important for the rbf model (in the top 3 positions). However, still, there is no significant difference in the importance scores.

It can be further seen that the REST and GAIT tasks are interchangeable. So when REST is selected as an important task, GAIT becomes a less dominant feature for the model.

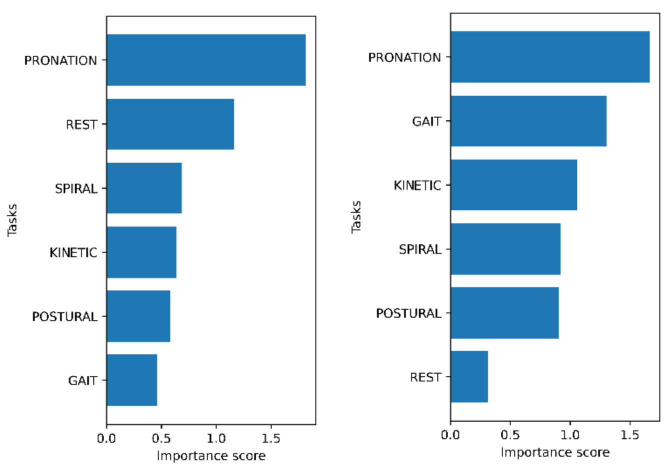

Fig. 3. The importance of movement patterns in the case of linear kernel voting SVM. On the left side, the voting took place on the predictions achieved with the linear model, on the right side on the predictions achieved with the rbf model.

V. Summary and Conclusion

Facilitating the diagnosis of PD has become an intensive research area since the disease is incurable. This emphasizes early recognition in order to start the appropriate therapy. Personalized therapy and medication are essential to slow down the progress of the disease and maintain a quality life.

Solutions using motor symptoms use several modalities, which can be based on speech, handwriting/drawing, and movement. They are based on the fact that the symptoms appear during fine motor movements and may be seen by computer vision early enough before a neurology examination. They can also be recorded using several methods, such as a video camera (body-mounted markers) or sensors attached to the body.

Previous research has shown that using only one sensor on the upper limbs may be sufficient for effective PD recognition among the sensors attached to the body. Another research proved that although the recognition achieved in the upper limbs is indeed significant, an improvement can be achieved by supplementing it with the measurement of the lower limb. However, this may cause discomfort for the patient.

Further research examined several forms of movement to determine which are essential for diagnosing the disease. They pointed out that there are conventionally important movements, such as walking and sitting, which can be supplemented with additional movements to identify specific symptoms.

Based on these, our research aimed to investigate the recognition of the six-movement tasks selected based on the UPDRS scale in patients with mild symptoms. Also, we attempted to increase recognition by using movement forms together with prediction fusion approaches.

In our experiments, an automatic feature extraction was used. This was not tuned or specified for PD but gave a general feature set to analyze time series. With the resulting features, SVM-based classification was performed to recognize the disease (HC or PD class).

PRONATION provided the highest performance (76.2% balanced accuracy), examining the movement forms separately. The second highest performance is achieved with the KINETIC task (75.0% balanced accuracy for linear kernel and 72.6% balanced accuracy for rbf kernel).

Aggregating the predictions was done with *soft*, *hard* and *SVM-based* voting. Based on the results, soft and hard voting performed better than SVM-based voting. The improvement was measured compared to the PRONATION task (which achieved the best performance). With the improvement, the balanced accuracy was 79.8%. In other words, this voting approach recognized three more people correctly with the final decision. This also points out that it is unnecessary to set a learning algorithm for the final prediction; instead, averaging the predictions is sufficient (considering the current setup).

This preliminary work suggests that the PRONATION includes the majority of PD information. However, especially at an early stage, the patient may not exclusively present fine motor symptoms during all the examination tasks. For this, the post-fusion of the predications (votings) may help the doctor to see the overall PD probability.

The sex difference may indicate a limitation in these results. In the [26] work, the authors found no statistically significant difference in how men and women drew the spiral under guided drawing conditions. We believe there is little or no difference in gender when doing the movement tasks similarly in a guided condition. However, this may require further examination.

For future work, we aim to explore hyperparameters for SVM models that ensure more optimal learning. Additionally, examining movements with other feature extraction methods can help validate these results because the models' performance and the movements' importance may depend on the applied feature extraction procedure.

REFERENCES

[1] Z. Saman and Y. Sridhara S., "Parkinson Disease." Treasure Island (FL): StatPearls Publishing, Jan. 2023. Accessed: Nov. 21, 2023. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK470193/

[2] D. K. Simon, C. M. Tanner, and P. Brundin, "Parkinson Disease Epidemiology, Pathology, Genetics, and Pathophysiology," *Clinics in Geriatric Medicine*, vol. 36, no. 1, pp. 1–12, Feb. 2020, doi: 10.1016/j.cger.2019.08.002.

[3] A. Schrag, L. Horsfall, K. Walters, A. Noyce, and I. Petersen, "Prediagnostic presentations of Parkinson's disease in primary care: a case-control study," *The Lancet Neurology*, vol. 14, no. 1, pp. 57–64, Jan. 2015, doi: 10.1016/S1474-4422(14)70287-X.

[4] S. Sveinbjornsdottir, "The clinical symptoms of Parkinson's disease," *Journal of Neurochemistry*, vol. 139, no. S1, pp. 318–324, Oct. 2016, doi: 10.1111/jnc.13691.

[5] A. A. Moustafa *et al.*, "Motor symptoms in Parkinson's disease: A unified framework," *Neuroscience & Biobehavioral Reviews*, vol. 68, pp. 727–740, Sep. 2016, doi: 10.1016/j.neubiorev.2016.07.010.

[6] P. Rizek, N. Kumar, and M. S. Jog, "An update on the diagnosis and treatment of Parkinson disease," *CMAJ*, vol. 188, no. 16, pp. 1157–1165, Nov. 2016, doi: 10.1503/cmaj.151179.

[7] Movement Disorder Society Task Force on Rating Scales for Parkinson's Disease, "The Unified Parkinson's Disease Rating Scale (UPDRS): Status and recommendations," *Movement Disorders*, vol. 18, no. 7, pp. 738–750, Jul. 2003, doi: 10.1002/mds.10473.

[8] K. Polat, "A Hybrid Approach to Parkinson Disease Classification Using Speech Signal: The Combination of SMOTE and Random Forests," in *2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT)*, Istanbul, Turkey: IEEE, Apr. 2019, pp. 1–3. doi: 10.1109/EBBT.2019.8741725.

[9] H. Gunduz, "Deep Learning-Based Parkinson's Disease Classification Using Vocal Feature Sets," *IEEE Access*, vol. 7, pp. 115540–115551, 2019, doi: 10.1109/ACCESS.2019.2936564.

[10] S. Rosenblum, M. Samuel, S. Zlotnik, I. Erikh, and I. Schlesinger, "Handwriting as an objective tool for Parkinson's disease diagnosis," *J Neurol*, vol. 260, no. 9, pp. 2357–2361, Sep. 2013, doi: 10.1007/s00415-013-6996-x.

[11] C. Kotsavasiloglou, N. Kostikis, D. Hristu-Varsakelis, and M. Arnaoutoglou, "Machine learning-based classification of simple drawing movements in Parkinson's disease," *Biomedical Signal Processing and Control*, vol. 31, pp. 174–180, Jan. 2017, doi: 10.1016/j.bspc.2016.08.003.

[12] W. Zeng, F. Liu, Q. Wang, Y. Wang, L. Ma, and Y. Zhang, "Parkinson's disease classification using gait analysis via deterministic learning," *Neuroscience Letters*, vol. 633, pp. 268–278, Oct. 2016, doi: 10.1016/j.neulet.2016.09.043.

[13] A. Som, N. Krishnamurthi, M. Buman, and P. Turaga, "Unsupervised Pre-trained Models from Healthy ADLs Improve Parkinson's Disease Classification of Gait Patterns," in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, Montreal, QC, Canada: IEEE, Jul. 2020, pp. 784–788. doi: 10.1109/EMBC44109.2020.9176572.

[14] B. Lin *et al.*, "Bradykinesia Recognition in Parkinson's Disease via Single RGB Video," *ACM Trans. Knowl. Discov. Data*, vol. 14, no. 2, pp. 1–19, Apr. 2020, doi: 10.1145/3369438.

[15] L. Lonini *et al.*, "Wearable sensors for Parkinson's disease: which data are worth collecting for training symptom detection models," *npj Digital Med*, vol. 1, no. 1, p. 64, Nov. 2018, doi: 10.1038/s41746-018-0071-z.

[16] S. T. Moore, H. G. MacDougall, and W. G. Ondo, "Ambulatory monitoring of freezing of gait in Parkinson's disease," *Journal of Neuroscience Methods*, vol. 167, no. 2, pp. 340–348, Jan. 2008, doi: 10.1016/j.jneumeth.2007.08.023.

[17] N. Shawen *et al.*, "Role of data measurement characteristics in the accurate detection of Parkinson's disease symptoms using wearable sensors," *J NeuroEngineering Rehabil*, vol. 17, no. 1, p. 52, Dec. 2020, doi: 10.1186/s12984-020-00684-4.

[18] A. Talitckii *et al.*, "Defining Optimal Exercises for Efficient Detection of Parkinson's Disease Using Machine Learning and Wearable Sensors," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–10, 2021, doi: 10.1109/TIM.2021.3097857.

[19] E. Rovini, C. Maremmani, and F. Cavallo, "A Wearable System to Objectify Assessment of Motor Tasks for Supporting Parkinson's Disease Diagnosis," *Sensors*, vol. 20, no. 9, p. 2630, May 2020, doi: 10.3390/s20092630.

[20] M. V. Albert, S. Toledo, M. Shapiro, and K. Kording, "Using Mobile Phones for Activity Recognition in Parkinson's Patients," *Front. Neur.*, vol. 3, 2012, doi: 10.3389/fneur.2012.00158.

[21] Y. Dai *et al.*, "Multi-Focus Image Fusion Based on Convolution Neural Network for Parkinson's Disease Image Classification," *Diagnostics*, vol. 11, no. 12, p. 2379, Dec. 2021, doi: 10.3390/diagnostics11122379.

[22] S. A. Shah *et al.*, "Sensor Fusion for Identification of Freezing of Gait Episodes Using Wi-Fi and Radar Imaging," *IEEE Sensors J.*, vol. 20, no. 23, pp. 14410–14422, Dec. 2020, doi: 10.1109/JSEN.2020.3004767.

[23] F. Heidarivincheh *et al.*, "Multimodal Classification of Parkinson's Disease in Home Environments with Resiliency to Missing Modalities," *Sensors*, vol. 21, no. 12, p. 4133, Jun. 2021, doi: 10.3390/s21124133.

[24] M. Barandas *et al.*, "TSFEL: Time Series Feature Extraction Library," *SoftwareX*, vol. 11, p. 100456, Jan. 2020, doi: 10.1016/j.softx.2020.100456.

[25] F. Pedregosa *et al.*, "Scikit-learn: Machine Learning in Python," 2012, doi: 10.48550/ARXIV.1201.0490.

[26] M. San Luciano *et al.*, "Digitized Spiral Drawing: A Possible Biomarker for Early Parkinson's Disease," *PLoS ONE*, vol. 11, no. 10, p. e0162799, Oct. 2016, doi: 10.1371/journal.pone.0162799.