Markerless gait analysis and circular statistics: a case study in Parkinson's disease classification

Cesare Davide Pace[1,2][0009-0006-5678-1610], Abdul Rahman[1,2][0009-0008-3185-585X], Houyem Demni[1,2][0000-0002-7745-2393], Alessandro Marco De Nunzio[0000-0003-4862-6742] Claudio De Stefano[1,2][0000-0002-7654-6849], Francesco Fontanella[1,2][0000-0002-3242-0179], Mario Molinara[1,2][0000-0002-6144-0654], and Giovanni C. Porzio[1,2][0000-0003-1208-6991]

[1] University of Cassino and Southern Lazio, Cassino, 03043, Italy
[2] EUt+ Research Institute in Data Science, European University of Technology, European Union
[3] Department of Research and Development, LUNEX, Luxembourg
[4] Luxembourg Health & Sport Sciences Research Institute ASBL, Luxembourg

Abstract. Markerless gait analysis enables Parkinson's disease (PD) assessment directly from video, without markers or wearable sensors. Most pipelines describe joint-angle trajectories with standard summary statistics (means, standard deviations, . . .). However, joint angles are directional quantities that lie on a non-linear manifold, and summarising them with statistics designed for data on a line may misrepresent their distribution. This paper investigates whether circular statistical descriptors provide a representation of angular gait signals that is at least as informative as the standard one for supervised PD classification. Using a public vision-based gait dataset and ViTPose for pose estimation, joint-angle trajectories are summarised with two matched feature families: a standard (linear) family and a circular one. The two families are compared under one identical supervised protocol that couples filter-based selection, recursive feature elimination, and four classifiers, assessed over repeated subject-level train–test splits. The goal is not to maximise classification performance, but to test whether circular descriptors can match or improve upon the standard descriptors conventionally used for angular gait data. As a result, circular descriptor's performance is comparable to, and in places better than, the standard ones, with the most visible differences in the more difficult trial and in the identification of PD subjects. When angular distributions are affected by pose-estimation errors, the adopted robust circular descriptors appear particularly useful. Estimates are obtained over repeated splits, and differences should be read in light of run-to-run variability. Overall, results suggest circular statistics is a valuable, geometry-aware addition to markerless gait analysis.

Keywords: Circular median absolute deviation · Pose estimation · Robust statistics.

1 Introduction

Parkinson's disease (PD) is a progressive neurodegenerative disorder commonly associated with motor symptoms such as bradykinesia, rigidity, tremor, postural instability, and gait impairment [3]. Because walking requires the coordinated control of multiple body segments, alterations in gait can provide clinically useful information for diagnosis, monitoring, and evaluation of disease progression [5, 9]. Quantitative gait analysis has therefore become an important tool for studying motor dysfunction in PD and for supporting objective assessment beyond visual clinical observation [24, 12]. Traditional gait analysis is based on motion-capture laboratories, force platforms, or wearable sensors. These technologies can provide accurate spatiotemporal and kinematic measurements, but often require specialized equipment, controlled acquisition environments, careful calibration, and trained operators [18, 7]. These requirements limit their scalability and make routine or home-based use difficult.

Recent progress in computer vision has made markerless gait analysis a practical alternative: pose-estimation models can extract anatomical keypoints directly from videos, enabling joint trajectories and gait descriptors to be computed without attaching markers or sensors to the subject [11, 25]. Most markerless gait-classification pipelines transform keypoint trajectories into a set of handcrafted statistical descriptors. In the case of kinematic features, these descriptors are often computed from joint angles such as the knees, hips, shoulders, and elbows. Because of the complexity of the problem, one of the main approaches in markerless gait classification includes summarizing the distribution over time of each angle by standard summary measures such as their arithmetic means, quartiles, standard deviations, and ranges (see e.g. [19]). However, such an approach may undervalue the specificity of angular signals. Angles are directional quantities, and lie on a nonlinear manifold. That is, on the circumference of the unit circle. Using standard measures designed for data lying on a line in place of specific measures can misrepresent the distribution of angular signals and lead to a loss of information. For such a reason, following some preliminary results made available in [1], this work aims at studying the role of circular statistical descriptors in markerless gait analysis for PD classification. In that respect, it differs from previous work that established a markerless pipeline based on pose estimation, feature extraction, and explainable machine learning [19]. The focus here is on the feature representation itself.

As a consequence, the main contribution of this paper is in the comparison of classical statistical descriptors against circular descriptors extracted from the same angular gait signals, evaluating their effect on supervised classification. As a further contribution, it is observed that pose-estimation might undergo some relevant measurement errors. Henceforth, robust descriptors of gait distributions are also evaluated. It is expected they are particularly effective for classification when the angular distributions contain irregular observations.

Circular descriptors are well described in many relevant works (e.g., in [21]). Circular location (means) and dispersion (standard deviation) measures are available. Robust circular measures (median, median absolute deviation [8]) have also been designed to respect the geometry of directional observations. The use of dispersion measures is potentially relevant for gait analysis because pathological movement may appear not only as a shift in the typical joint configuration, but also as increased irregularity or dispersion of angular trajectories. In PD, variability of movement is especially meaningful, since gait impairment may involve shortened steps, reduced joint excursion, altered coordination, and increased inconsistency across the gait cycle [26, 6].

Finally, this work uses a public vision-based dataset as a case study. This supports reproducibility and allows the resulting comparison to be used as a benchmark for future research [15].

The remainder of the paper is organized as follows. Section 2 reviews related work. Section 3 describes the dataset and gait-signal extraction process. Section 4 presents the machine-learning methodology. Section 5 reports the experimental results. Finally, Section 6 discusses conclusions and future directions.

2 Related Work

Gait analysis for PD has been investigated using marker-based systems, wearable sensors, and video-based markerless methods. These approaches differ in accuracy, accessibility, cost, and clinical practicality.

2.1 Marker-Based Approaches

Motion-capture systems are often considered the reference technology for detailed gait assessment because they provide highly accurate kinematic measurements. Park et al. used kinematic features and a Random Forest classifier to discriminate PD subjects from healthy controls [20], while Ricciardi et al. used gait parameters acquired with optoelectronic cameras to distinguish PD from atypical Parkinsonism [22]. Despite their accuracy, motion-capture systems require markers, multiple cameras, and controlled laboratory environments, limiting their scalability.

2.2 Markerless Video-Based Approaches

Markerless methods exploit computer vision to infer body motion directly from video, making gait analysis more accessible and less intrusive. Previous studies have used skeleton-based representations, deep-learning models, 3D pose estimation, smartphone videos, silhouette analysis, and graph-based approaches for PD assessment [2, 10, 17].

Recent pose-estimation architectures, including ViTPose [25], have further improved the reliability of markerless keypoint extraction. Pace et al. demonstrated the feasibility of public-dataset markerless PD classification using ViTPose and gait descriptors derived from joint motion [19]. The present work builds on this line of research by focusing on how angular gait trajectories should be summarized.

2.3 Circular Statistics for Movement and Gait

Circular statistics are designed for variables measured on a periodic domain, such as directions, phases, and angles [21]. Their use is well established when the topology of the data makes linear statistics inappropriate. In gait and biomechanics, circular methods have been used to study phase relationships and coordination variability. Stock et al. discussed circular-statistical issues in vector-coding variability and proposed alternatives for bivariate coordination analysis [23]. More recently, Jeong and van Emmerik compared circular and ellipse-based methods for assessing coordination variability [14]. Robust circular dispersion measures have also been proposed for handling anomalous directional observations [8].

Despite these developments, circular descriptors remain underused in supervised markerless gait classification for PD. Many pipelines treat joint angles as ordinary scalar time series and summarize them with linear statistics. This may be adequate when angular trajectories remain far from the wrap-around boundary and contain few outliers, but it can become fragile when measurements are dispersed, noisy, or affected by pose-estimation errors. A systematic comparison between standard and circular descriptors is therefore needed to understand whether circular summaries improve classification and robustness in markerless PD gait analysis.

3 Material

This section describes the data and input representations used to compare standard and circular statistical features. The material consists of a public video-based gait dataset, markerless pose estimates obtained from walking sequences, and two families of descriptors extracted from joint-angle trajectories.

3.1 Dataset

The experiments use the public vision-based gait dataset introduced by Kour et al. [15]. The dataset was designed to support gait-analysis research from video recordings and includes subjects with PD, healthy controls (HC), and knee osteoarthritis. Videos were acquired in the sagittal plane using a DSLR camera at 1920×1080 resolution and 50 frames per second. Participants walked along a marked path in front of a controlled background, and each subject was recorded in two directions, corresponding to two trials.

Following the previous markerless PD study on this dataset [19], this work considers the binary PD-versus-HC problem. The original PD group includes early, moderate, and severe disease stages. Severe-stage recordings are excluded because subjects are commonly assisted while walking, which changes their natural gait pattern and can introduce ambiguity for pose-estimation algorithms when an accompanying person appears in the scene. After this exclusion, the analyzed cohort contains 13 PD subjects from the early and moderate stages and 30 HC subjects. Each subject contributes two walking trials, one for each acquisition direction.

3.2 Markerless Pose and Angular Gait Signals

Markerless body motion is extracted from the videos using ViTPose, a transformer-based human pose-estimation architecture [25]. ViTPose follows a top-down strategy: a person detector first identifies the subject in each frame, and the pose-estimation model then predicts anatomical keypoints using the MS COCO keypoint convention. Tracking these keypoints over time provides a sequence of body-joint trajectories for each walking trial.

To make the two acquisition directions comparable, right-to-left videos are horizontally normalized before feature extraction. From the detected keypoints, joint angles are computed for the main upper- and lower-limb articulations used in gait description: shoulders, elbows, hips, and knees. These angles form temporal signals that describe how each joint evolves during walking. Since these signals are angular by construction, they are the central object of the comparison between linear and circular feature representations.

Pose-estimation outputs may contain occasional errors due to occlusions, blurred frames, imperfect detections, or confusion with other visible body parts. Such errors can produce angular observations that are physically implausible or far from the dominant movement pattern. This motivates the inclusion of robust descriptors, especially for the circular feature family, where dispersion and median-based quantities can reduce the influence of anomalous angular samples.

3.3 Standard and Circular Feature Families

Each joint-angle trajectory is summarized into a compact feature vector before classification. Two feature families are compared.

The standard family treats angle trajectories as ordinary scalar signals. It includes classical descriptive statistics such as arithmetic mean, median, first and third quartiles, standard deviation, median absolute deviation, and range of motion. These descriptors are simple, widely used in gait-analysis pipelines, and provide a baseline representation of location, dispersion, and joint excursion.

The circular family treats the same trajectories as angular observations on a periodic domain. It includes circular mean, circular median, circular hinges, circular standard deviation, circular median absolute deviation, and circular range of motion. These descriptors account for the wrap-around structure of angles and are expected to better represent distributions that are close to the $0°/360°$ boundary or affected by directional outliers.

The comparison between these two families is designed to isolate the effect of the statistical representation. Both families are extracted from the same markerless joint-angle signals and evaluated under the same supervised-learning protocol. In the subsequent experimental section, classification performance and feature-selection stability will be analyzed separately for the two walking trials.

4 Methods

This section describes the supervised-learning protocol used to compare the two feature families introduced in Section 3.3: the standard (linear) descriptors and

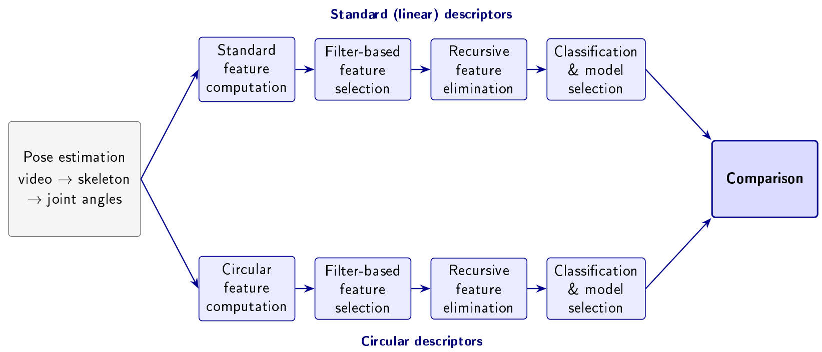

Fig. 1: Overview of the supervised-learning protocol. Markerless pose estimation turns each video frame into a skeleton, from which the joint-angle trajectories are computed. The same trajectories are summarised by two feature families, standard (linear) and circular, which are then compared under one identical pipeline of feature computation, filter-based feature selection, recursive feature elimination, and classifier training with model selection. Both families are processed in exactly the same way, and their classification results are contrasted at the final comparison stage.

the circular descriptors. The objective is not to identify the best-performing classifier, but to assess whether circular descriptors, which respect the angular nature of joint signals, can match or improve upon the standard descriptors commonly used in gait analysis.

Both feature families are extracted from the same markerless joint-angle trajectories and evaluated under an identical pipeline consisting of feature computation, filter-based selection, recursive feature elimination (RFE), classifier training, and performance evaluation. Therefore, any difference in classification performance or feature-selection patterns can be attributed to the statistical representation of the angular signals rather than to changes in the experimental workflow. The overall procedure is summarized in Figure 1.

The protocol follows and extends previous markerless gait-analysis studies based on standard descriptors [19] and preliminary circular-statistics investigations on the same dataset [1]. The remainder of this section describes the extracted descriptors (Section 4.1), the robust measures considered (Section 4.2), the feature-selection procedure (Sections 4.3.1 and 4.3.2), the evaluated classifiers (Section 4.4), and the performance measures (Section 4.5).

4.1 Standard and Circular Features

The supervised-learning experiments were conducted separately for the standard and circular feature families described in Section 3.3. For each walking trial, descriptors extracted from the eight joint-angle trajectories were concatenated into fixed-length feature vectors before classification.

The standard linear feature family included the mean, first quartile, median, third quartile, standard deviation, and range of motion. The median is a robust measure of central tendency being less sensitive to outliers than the mean, while the first and third quartiles are 25% robust as they remain unaffected by the extreme observations in the outer 25% of the data. The circular feature family included the circular mean, the circular hinge 1, the circular median, the circular hinge 3, the circular standard deviation, and the circular range of motion. In contrast to linear statistics, circular descriptors explicitly account for angular wrap-around effects, ensuring that angles close to $0°$ and $360°$ are treated as neighboring values rather than distant observations.

A key measure in circular statistics is the mean resultant length, which quantifies angular concentration. Values close to 1 indicate tightly clustered angles, whereas values near 0 indicate high angular variability. Circular variability can also be quantified using the circular standard deviation, which reflects dispersion on the circular manifold rather than along a linear axis. Circular quartile analogs, referred to as circular hinges, represent lower and upper quartile estimates adapted to directional data.

Each descriptor family was evaluated independently under the same supervised-learning protocol in order to isolate the effect of the statistical representation.

4.2 Robust Features

Gait datasets frequently contain noise, atypical steps, sensor inaccuracies, and occasional tracking failures. Such irregular observations may strongly influence classical statistical measures, particularly the mean and standard deviation, which are highly sensitive to extreme values. To address this limitation, robust statistical descriptors were additionally considered in order to obtain feature representations that remain stable in the presence of anomalous observations.

In contrast to classical measures, which can be substantially shifted by a small number of outliers, robust measures preserve representative information about the central tendency and dispersion of the majority of the data. In particular, the median is a robust measure of location, as it is minimally affected by extreme observations and possesses a high breakdown point of 50%, meaning that up to half of the observations can be contaminated before the estimator becomes arbitrarily distorted. Similarly, robust measures of dispersion characterize variability without being unduly influenced by outliers. This property is particularly important in gait analysis, where variability may arise from both natural stride-to-stride fluctuations and measurement artifacts.

For linear data, we consider the median absolute deviation [16] (MAD) as a robust alternative to the standard deviation, defined as $MAD = \text{median}(|x_i - \tilde{x}|)$, where $\tilde{x}$ denotes the sample median. Unlike the standard deviation, which relies on squared deviations and therefore amplifies extreme observations, the MAD uses absolute deviations around the median and consequently provides a more resistant estimate of variability. Due to its high breakdown point (50%), the MAD remains stable even when a substantial proportion of the observations are contaminated by outliers. In addition, the MAD has a bounded influence function, which means that the effect of a single extreme observation remains limited.

Robust circular dispersion was quantified using the circular median absolute deviation (CMAD) [8], which corresponds to the linear median of the distribution of the shortest arc distances from the circular median. The CMAD is thus defined as $CMAD = \text{median}\{d(\theta_i, \tilde{\theta})\}$, where $d(\theta_i, \tilde{\theta})$ denotes the shortest angular distance between the observation θ_i and the circular median $\tilde{\theta}$, i.e., $d(\theta_i, \tilde{\theta}) = \pi - |\pi - |\theta_i - \tilde{\theta}||$.

Similar to the MAD in the linear setting, the CMAD is a robust measure of angular variability that is resistant to outliers. Moreover, the CMAD has a bounded influence function and a high breakdown point around 50%. These properties make the CMAD particularly suitable for characterizing variability in circular gait features, where occasional atypical measurements may distort classical measures of dispersion.

4.3 Feature Selection

Feature selection was carried out in two sequential stages. Both stages operated exclusively on the training data. The test set was never used during feature selection.

4.3.1 Filter Selection Each feature was evaluated using a two-sample t-test comparing PD and HC groups on the training split. The resulting p-values were adjusted for multiple comparisons using False Discovery Rate correction [4]. Features with an adjusted p-value below 0.05 were retained. If fewer than five features passed this threshold, the top-ranked features according to adjusted p-value were retained as a fallback. This ensured that at least a minimal set of features was always passed to the next stage, even when group differences were weak.

4.3.2 Recursive Feature Elimination RFE was then applied to the subset of features retained after filtering. A Linear SVM served as the internal estimator during the RFE procedure. Several subset sizes were evaluated using 5-fold cross-validation on the training data, and the subset with the best cross-validated accuracy was selected. The filter stage reduced the number of candidate features, while RFE identified a smaller subset of informative descriptors. The same feature selection procedure was applied independently to the standard and circular descriptor families.

4.4 Classifiers

Four classifiers were evaluated: Logistic Regression (LR), Linear Discriminant Analysis (LDA), Support Vector Machine (SVM), and Naive Bayes (NB) (for a general description see e.g., [13]). These classifiers were selected because they represent different classification strategies, including probabilistic, discriminative, margin-based, and generative approaches. The objective of the study was not to identify the best performing classifier, but rather to compare the effect of standard and circular feature representations under the same supervised learning protocol. Results that remain relatively consistent across different classification approaches provide more confidence than findings depending on a single model choice.

Logistic Regression estimates class probabilities through a logistic model whose parameters are obtained by maximum likelihood estimation. Linear Discriminant Analysis classifies observations by modeling the distribution of predictors within each class and assuming a common covariance matrix across groups. Support Vector Machine is a margin-based classifier that identifies the hyperplane that maximizes the separation between classes. Naive Bayes is a probabilistic classifier based on Bayes' theorem that assumes conditional independence among predictors given the class label.

4.5 Performance measures

Classification performance was evaluated using Accuracy, Sensitivity, Specificity, Matthews Correlation Coefficient (MCC), and Area Under the Receiver Operating Characteristic Curve (AUC). These measures provide complementary information on overall classification performance, the identification of Parkinson's disease (PD) subjects, the recognition of healthy controls (HC), and the overall discriminative ability of the classifiers.

Accuracy measures the proportion of correctly classified observations. Sensitivity quantifies the ability to correctly identify PD subjects, whereas Specificity measures the ability to correctly recognize HC subjects. MCC provides a balanced assessment by considering all entries of the confusion matrix simultaneously and remains informative in the presence of class imbalance. AUC summarizes the discriminative capability of a classifier across all decision thresholds.

Since the dataset exhibits a moderate imbalance between PD and HC subjects, particular attention was given to MCC, Sensitivity, and AUC, which provide a more informative evaluation than Accuracy alone.

5 Experimental Results

This section presents the classification results obtained from the two descriptor families across the two gait trials. Before discussing the results, we briefly describe the evaluation protocol used throughout the study. The performance of the standard and circular descriptors is then analyzed separately for each trial, followed by an overall comparison between the two descriptor families.

5.1 Experimental Setup

All experiments were repeated 50 times using stratified subject-level train–test splits. In each repetition, 70% of the subjects were assigned to the training set and 30% to the test set while preserving class proportions. Subject-level splitting ensured that recordings from the same individual never appeared in both sets.

Features were normalized using MinMax scaling fitted on the training data and subsequently applied to the test data. Feature selection and model tuning were performed exclusively on the training set according to the procedure described in Section 4.3, using five-fold cross-validation when required.

Performance was evaluated on the independent test set using Accuracy, Sensitivity, Specificity, MCC, and AUC. Reported results correspond to mean values over the 50 repetitions.

5.2 Classification Performance

Tables 1–4 summarize the classification results obtained for the standard and circular descriptor families across the two gait trials. Results are reported separately for each trial in order to assess the consistency of the findings across repeated recordings of the same subjects.

Table 1: Standard descriptors for Trial 1. Mean values over 50 repeated splits. *Feat.*: mean number of features selected after RFE. **Best result by performance measure.**

Classifier	Accuracy	Sensitivity	Specificity	MCC	AUC	Feat.
LDA	0.89	**0.79**	0.93	**0.74**	**0.90**	7
LR	0.89	0.67	**0.96**	0.70	0.85	7
Naive Bayes	0.87	0.72	0.92	0.68	0.85	7
SVM (Linear)	**0.90**	0.70	**0.96**	0.73	0.87	7

Table 2: Circular descriptors for Trial 1. Mean values over 50 repeated splits. *Feat.*: mean number of features selected after RFE. **Best result by performance measure.**

Classifier	Accuracy	Sensitivity	Specificity	MCC	AUC	Feat.
LDA	**0.92**	**0.80**	0.96	**0.79**	**0.96**	6
LR	0.90	0.69	**0.97**	0.74	0.90	6
Naive Bayes	0.89	0.72	0.94	0.72	0.85	6
SVM (Linear)	0.91	0.74	**0.97**	0.77	0.92	6

Table 3: Standard descriptors for Trial 2. Mean values over 50 repeated splits. *Feat.*: mean number of features selected after RFE. **Best result by performance measure.**

Classifier	Accuracy	Sensitivity	Specificity	MCC	AUC	Feat.
LDA	0.78	0.55	0.86	0.46	0.76	11
LR	0.83	0.46	**0.97**	0.56	0.82	11
Naive Bayes	**0.84**	**0.61**	0.93	0.59	**0.85**	11
SVM (Linear)	**0.84**	0.55	0.95	**0.61**	0.83	11

Table 4: Circular descriptors for Trial 2. Mean values over 50 repeated splits. *Feat.*: mean number of features selected after RFE. **Best result by performance measure.**

Classifier	Accuracy	Sensitivity	Specificity	MCC	AUC	Feat.
LDA	0.82	**0.68**	0.87	0.60	0.85	13
LR	0.87	0.54	**1.00**	0.71	0.91	13
Naive Bayes	0.87	0.64	0.95	0.65	0.86	13
SVM (Linear)	**0.89**	0.66	0.97	**0.73**	**0.92**	13

Tables 1 and 2 show that both descriptor families produced good classification performance in Trial 1. The differences between standard and circular descriptors were relatively small, and both approaches distinguished PD and HC subjects with a similar degree of accuracy. With standard descriptors, the best accuracy was obtained by SVM (Linear), while LDA produced the best sensitivity, MCC, and AUC. With circular descriptors, LDA produced the strongest overall result. Circular descriptors also used fewer selected features on average, with 6 features compared to 7 for the standard descriptors.

Tables 3 and 4 show a different pattern for Trial 2. Classification performance decreased when standard descriptors were used, especially for sensitivity and MCC. In contrast, circular descriptors produced higher values for both metrics across all classifiers. The gap between the two descriptor families was therefore more visible in Trial 2 than in Trial 1. Although circular descriptors required more selected features on average, they gave a more stable performance across classifiers.

The sensitivity values reported in Tables 3 and 4 indicate that the circular descriptors were often more effective at identifying PD subjects, particularly in the second trial. This observation suggests that circular representations may capture aspects of angular gait data that are not fully described by standard descriptors.

Across both trials, SVM (Linear) and LDA generally produced the strongest results. Although the magnitude of the differences varied among classifiers, circular descriptors consistently achieved performance that was comparable to, and in several cases better than, that obtained with the standard descriptors. Taken together, these findings indicate that circular statistical descriptors represent a viable alternative to the standard descriptors traditionally used for angular gait data.

5.3 Feature Selection Analysis

To better understand which variables contributed most to the classification task, feature selection frequencies obtained after the RFE procedure were examined. Selection rates were calculated across the 50 repeated train-test splits and provide an indication of feature stability throughout the classification process. A selection rate of 100% indicates that a feature was selected in every repetition. Tables 5–6 summarize the most frequently selected features for the standard and circular descriptor families across the two gait trials. Features that were retained repeatedly across different splits can be considered more stable contributors to the classification task, since their selection was less dependent on the specific composition of the training data.

Table 5: Most frequently selected features in Trial 1. Selection rates are reported across the 50 repeated splits. Robust measures are shown in italics.

Standard Features

Joint	**Metric**	**Selection Rate**
Left Knee	Standard deviation	96%
Right Knee	Standard deviation	96%
Left Shoulder	*Median*	*62%*
Left Knee	First Quartile	54%
Left Shoulder	First Quartile	48%
Left Shoulder	Mean	46%
Right Shoulder	Third Quartile	46%

Circular Features

Joint	**Metric**	**Selection Rate**
Right Knee	Circular standard deviation	100%
Left Knee	Circular standard deviation	98%
Left Shoulder	*Circular Median*	*44%*
Left Knee	First Circular Hinge	38%
Left Shoulder	Circular Mean	32%
Right Shoulder	*Circular Median*	*30%*

Table 6: Most frequently selected features in Trial 2. Selection rates are reported across the 50 repeated splits. Robust measures are shown in italics.

Standard Features

Joint	**Metric**	**Selection Rate**
Right Shoulder	*Median*	*90%*
Left Knee	First Quartile	84%
Right Shoulder	First Quartile	82%
Right Shoulder	Mean	82%
Left Knee	*Median absolute deviation*	*72%*
Right Shoulder	*Median absolute deviation*	*62%*
Left Shoulder	*Median absolute deviation*	*60%*
Left Elbow	*Median absolute deviation*	*58%*
Right Shoulder	Third Quartile	58%

Circular Features

Joint	**Metric**	**Selection Rate**
Left Knee	Circular standard deviation	100%
Right Knee	Circular standard deviation	98%
Right Shoulder	*Circular Median*	*94%*
Right Shoulder	First Circular Hinge	86%
Right Shoulder	Circular Mean	86%
Left Knee	First Circular Hinge	84%
Left Knee	*Circular median absolute deviation*	*72%*
Right Shoulder	*Circular median absolute deviation*	*68%*
Left Shoulder	*Circular median absolute deviation*	*64%*
Left Elbow	*Circular median absolute deviation*	*62%*
Left Shoulder	Circular standard deviation	62%

In Trial 1 (Table 5), the standard descriptor family was characterized by knee related measures of dispersion. Shoulder based descriptors were also represented among the highest ranked variables, including measures of central tendency and distribution. The retained features were therefore distributed across both lower- and upper-limb joints, suggesting that information from multiple body regions contributed to the classification task.

Within the circular descriptor family, the highest ranked variables were the circular standard deviations of the knees. Shoulder based circular descriptors were also present, particularly measures describing angular location. As a result, the circular feature set incorporated information related to both angular position and angular dispersion.

In Trial 2 (Table 6), robust measures occupied a more prominent role within the standard descriptor family. Median absolute deviation appeared for several joints, while shoulder related descriptors accounted for many of the highest ranked variables. This observation is consistent with the presence of anomalies in the second trial, for which robust statistics are generally less sensitive than mean and standard deviation based measures.

The circular descriptor family also included several measures of angular dispersion. Circular standard deviation remained among the highest ranked variables, while circular median absolute deviation appeared across multiple joints. In addition, shoulder based measures such as the circular median, circular mean, and circular hinge contributed substantially to the selected feature set.

Across both trials, measures of dispersion occupied many of the highest ranked positions regardless of feature family. The greater representation of robust measures in the second trial is consistent with the increased influence of anomalous observations. Overall, the feature selection results indicate that descriptors related to the dispersion of joint motion were repeatedly identified by the classification pipeline as informative variables for distinguishing between PD and HC subjects.

6 Conclusions

Markerless gait analysis is a promising field for the diagnosis of diseases that affect the ability to move, and many research works investigate how information gathered throughout pose-estimation techniques can be exploited. Within this growing field of research, this work aimed at giving a contribution in the direction of which features to extract from pose-estimation patterns. Particularly, with the final aim of improving the performance of supervised classifiers, the focus was on if choosing circular distribution summary measures may improve over the use of what we called standard measures (mean, standard deviation of angles, range of motions, ...).

As a first preliminary work, this study was based on a single publicly available dataset. Results are encouraging. A fair comparison between the use of standard features and the corresponding circular measures showed that the latter basically provide outperforming results, at least within the adopted framework.

As a secondary hint, this work recalls the importance of considering statistical methods which are robust with respect to the presence of outliers. Pose estimation cannot generally avoid the presence of anomalous values. The use of standard (median, median absolute deviation) and circular (circular median, circular median absolute deviation) robust measures adds learning power in such a setting.

The use of a single dataset is clearly a limitation of this work, as well as the use of a limited number of classifiers that have been compared. On the other hand, the main goal was to pave the way for further studies. As a matter of fact, this contribution suggests it is worth to go steps further and enrich pose estimation markerless gait analysis with the plethora of tools available within the circular statistics realm.

References

1. Abdul Rahman, A.R., Fontanella, F., Molinara, M., Pace, C.D., Porzio, G.C.: Gait analysis and circular statistics: Improving parkinson's disease diagnostics. In: EUt+ Dats and Stats 2026 Conference Proceedings. pp. 75–80. In press (2026)
2. Ajay, J., Song, C., Wang, A., Langan, J., Li, Z.: A pervasive and sensor-free deep learning system for parkinsonian gait analysis. In: 2018 IEEE EMBS International Conference on Biomedical & Health Informatics. pp. 108–111. IEEE (2018)
3. Armstrong, M.J., Okun, M.S.: Diagnosis and treatment of parkinson disease: a review. JAMA **323**(6), 548–560 (2020)
4. Benjamini, Y., Hochberg, Y.: Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society: Series B **57**(1), 289–300 (1995)
5. di Biase, L., Di Santo, A., Caminiti, M.L., De Liso, A., Shah, S.A., Ricci, L., Di Lazzaro, V.: Gait analysis in parkinson's disease: An overview of the most accurate markers for diagnosis and symptoms monitoring. Sensors **20**(12) (2020). https://doi.org/10.3390/s20123529
6. Bouca-Machado, R., Jalles, C., Guerreiro, D., Pona-Ferreira, F., Branco, D., Guerreiro, T., Matias, R., Ferreira, J.J.: Gait kinematic parameters in parkinson's disease: a systematic review. Journal of Parkinson's Disease **10**(3), 843–853 (2020)
7. Chen, S., Lach, J., Lo, B., Yang, G.Z.: Toward pervasive gait analysis with wearable sensors: A systematic review. IEEE Journal of Biomedical and Health Informatics **20**(6), 1521–1537 (2016)
8. Demni, H., Hubert, M., Porzio, G.C., Rousseeuw, P.J.: Robust measures of dispersion for circular data with an anomaly detection rule. arXiv preprint arXiv:2603.01237 (2026)
9. Di Biase, L., et al.: Gait analysis in parkinson's disease: An overview of the most accurate markers for diagnosis and symptoms monitoring. Sensors **20**(12), 3529 (2020)
10. Guayacan, L.C., et al.: Visualising and quantifying parkinsonian gait using 3d convolutional neural networks. IEEE Access **9**, 150156–150166 (2021)
11. Guo, Y., Yang, J., Liu, Y., Chen, X., Yang, G.Z.: Detection and assessment of parkinson's disease based on gait analysis: A survey. Frontiers in Aging Neuroscience **14** (2022). https://doi.org/10.3389/fnagi.2022.916971

12. Hausdorff, J.M.: Gait dynamics in parkinson's disease: Common and distinct behavior among stride length, gait variability, and fractal-like scaling. Chaos **19**(2), 026113 (2009)
13. James, G., Witten, D., Hastie, T., Tibshirani, R.: An Introduction to Statistical Learning: With Applications in R. Springer, 2 edn. (2021)
14. Jeong, H., van Emmerik, R.: Comparison of circular and ellipse-based methods for assessing coordination variability. Journal of Applied Biomechanics (2026). https://doi.org/10.1123/jab.2025-0270, ahead of print
15. Kour, N., Arora, S.: Vision-based gait dataset for gait analysis. In: 2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering. pp. 2163–2168. IEEE (2022)
16. Law, J.: Robust statistics—the approach based on influence functions (1986)
17. Lu, M., Poston, K., Pfefferbaum, A., Sullivan, E.V., Fei-Fei, L., Pohl, K.M., Niebles, J.C., Adeli, E.: Vision-based estimation of mds-updrs gait scores for assessing parkinson's disease motor severity. In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2020. pp. 637–647. Springer (2020)
18. Moore, S.T., MacDougall, H.G., Gracies, J.M., Cohen, H.S., Ondo, W.G.: Long-term monitoring of gait in parkinson's disease. Gait & Posture **26**(2), 200–207 (2007). https://doi.org/10.1016/j.gaitpost.2006.09.011
19. Pace, C.D., De Nunzio, A.M., De Stefano, C., Fontanella, F., Molinara, M.: Markerless gait analysis for parkinson's disease diagnosis: A study on machine learning integration and features explainability. Engineering Applications of Artificial Intelligence **161**, 112031 (2025)
20. Park, H., Shin, S., Youm, C., Cheon, S.M., Lee, M., Noh, B.: Classification of parkinson's disease with freezing of gait based on 360 turning analysis using 36 kinematic features. Journal of NeuroEngineering and Rehabilitation **18**, 1–18 (2021)
21. Pewsey, A., Neuhauser, M., Ruxton, G.D.: Circular Statistics in R. Oxford University Press (2013)
22. Ricciardi, C., Amboni, M., De Santis, C., Improta, G., Volpe, G., Iuppariello, L., Ricciardelli, G., D'Addio, G., Vitale, C., Barone, P., Cesarelli, M.: Using gait analysis' parameters to classify parkinsonism: A data mining approach. Computer Methods and Programs in Biomedicine **180**, 105033 (2019). https://doi.org/10.1016/j.cmpb.2019.105033
23. Stock, H., van Emmerik, R., Wilson, C., Preatoni, E.: Applying circular statistics can cause artefacts in the calculation of vector coding variability: A bivariate solution. Gait & Posture **65**, 51–56 (2018)
24. Whittle, M.W.: Gait Analysis: An Introduction. Butterworth-Heinemann, 5th edn. (2014)
25. Xu, Y., Zhang, J., Zhang, Q., Tao, D.: Vitpose: Simple vision transformer baselines for human pose estimation (2022)
26. Zanardi, A., Da Silva, E., Costa, R., Passos-Monteiro, E., Dos Santos, I., Kruel, L., Peyre-Tartaruga, L.: Gait parameters of parkinson's disease compared with healthy controls: a systematic review and meta-analysis. sci rep 11: 752 (2021)