Integrating Gait and Clinical Data with Explainable Artificial Intelligence for Parkinson's Prediction: The EDAM System

Nicoletta Balletti[1,2], Emanuela Guglielmi[2], Gennaro Laudato[2], Rocco Oliveto[2,3], Jonathan Simeone[3] and Roberto Zinni[4]

[1]*Center for Biotechnology, Institute of Biomedical Sciences of the Ministry of Defense, Rome, Italy*

[2]*University of Molise, Pesche (IS), Italy*

[3]*Datasound srl, Pesche (IS), Italy*

[4]*WordPower, San Salvo (CH), Italy*

Keywords: Gait Analysis, Clinical Biomarkers, Explainable Artificial Intelligence, Parkinson's Disease Detection.

Abstract: Several machine learning (ML) approaches have been introduced for gait and posture analysis, recognized as crucial for early diagnosing neurological disorders, particularly Parkinson's disease. However, these existing methods are often limited by their lack of integration with other clinical biomarkers and their inability to provide transparent, explainable predictions. To overcome these limitations, we introduce EDAM (Explainable Diagnosis Recommender), a system that leverages Explainable Artificial Intelligence (XAI) techniques to deliver both accurate predictions and clear, interpretable explanations of its diagnostic decisions. We evaluate the capabilities of EDAM in two main areas: distinguishing between healthy individuals and those with Parkinson's disease, and classifying abnormal gait patterns that may indicate early-stage Parkinson's disease. To ensure a comprehensive evaluation, we constructed one of the largest known dataset by merging and standardizing several existing datasets. This dataset includes 557 features and 7,303 labelled instances, covering a wide range of gait patterns and clinical features. Results show that EDAM achieves high accuracy in both tasks, demonstrating its potential for early detection of neurological disorders.

1 INTRODUCTION

The analysis of gait and posture (motion analysis) is crucial for the early diagnosis of several pathologies, especially neurological disorders, as well as for monitoring disease progression and evaluating a patient's therapeutic response (Buckley et al., 2019). Research has shown that by monitoring upper-body movements, it is possible to differentiate between healthy individuals and those with Parkinson's disease, while patients with ataxic symptoms, such as those with multiple sclerosis, exhibit deficits in postural control. Moreover, slow gait has been identified as a predictor of dementia, with early signs manifesting up to nine years before an official diagnosis is made (Buckley et al., 2019).

In the literature, several machine learning (ML) approaches have been proposed for monitoring and predicting specific diseases using gait and posture data acquired through motion analysis systems (*e.g.*, (Abdulhay et al., 2018; Costa et al., 2016; Cuzzolin et al., 2017; Mannini et al., 2016; Raknim and Lan, 2016)). However, these approaches often rely on univariate analyses, treating posture and gait data in isolation from other clinical biomarkers. While this simplification can streamline processing, it may overlook crucial insights that could emerge from a more comprehensive, multivariate analysis (Holzinger et al., 2017a). Additionally, the robustness of these systems is a recurring concern. Many studies are based on relatively small patient samples, which limits the generalizability of the findings (Buckley et al., 2019).

Another significant limitation of ML-based motion analysis systems is their lack of transparency. The predictions made by these models are frequently perceived as "black-box" decisions, leaving specialists unable to grasp the reasoning behind the outputs (Holzinger et al., 2017b). This opacity can reduce trust in the system, prompting specialists to dismiss even accurate predictions. To address this challenge, there is an increasing need for systems that offer ex-

Balletti, N., Guglielmi, E., Laudato, G., Oliveto, R., Simeone, J. and Zinni, R.
Integrating Gait and Clinical Data with Explainable Artificial Intelligence for Parkinson's Prediction: The EDAM System.
DOI: 10.5220/0013179400003911

In *Proceedings of the 18th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2025) - Volume 2: HEALTHINF*, pages 129-140
ISBN: 978-989-758-731-3; ISSN: 2184-4305

plainable and transparent predictions. Explainable Artificial Intelligence (XAI) techniques have emerged as a promising solution, enabling users to understand the rationale behind ML model outputs (Edwards and Veale, 2017). These techniques are particularly valuable in the medical field, where practitioners must interpret complex, heterogeneous data. AI methods, especially ML, are crucial for extracting actionable insights from such data, and XAI enhances this process by making predictions interpretable and useful for human experts.

To overcome the limitations of current approaches, in this paper, we present EDAM (Explainable Diagnosis Recommender), a system designed to support predictions related to neurological disorders by integrating posture and gait data with other clinical biomarkers. The Decision Support System (DSS) within EDAM—based on advanced ML techniques (Hastie et al., 2009)—assists specialists in two closely related tasks: distinguishing between healthy individuals and those with Parkinson's disease, and detecting abnormal gait patterns, thereby facilitating the early detection of neurological pathologies.

EDAM applies XAI techniques to generate transparent, interpretable predictions. By leveraging explainable machine learning algorithms, EDAM provides specialists with not only accurate predictions but also visual and natural-language explanations that clearly outline the factors influencing the diagnosis. Additionally, EDAM generates preliminary diagnostic reports (pre-reports) to help specialists analyze the results both quantitatively and qualitatively. These user-friendly predictions enhance the interpretation of data-driven insights, allowing specialists to seamlessly integrate their intuition, judgment, and experience into the decision-making process.

To evaluate the accuracy of EDAM, we conducted a study using a dataset that was constructed by homogenizing and merging several gait datasets from the literature (*i.e.*, datasets provided by Mehrizi et al., 2019, Schreiber and Moissenet, 2019, Jun et al., 2020, Kour et al., 2020). This process enabled us to create the largest dataset in the literature, consisting of 557 features and 7,303 labeled instances, making it a highly comprehensive resource for gait analysis and Parkinson's disease prediction. Besides evaluating the accuracy of EDAM DSS, the study allowed for an extensive evaluation of the importance of all gait-related features in both predicting Parkinson's disease and classifying gait types. The results demonstrated that EDAM achieved high accuracy in both tasks, with notable performance in detecting early signs of Parkinson's disease. Specifically, the classification model was able to distinguish between Parkinsonian and healthy subjects with a high degree of precision, and it accurately classified various gait patterns, including those associated with early neurological symptoms. These findings underscore the potential of EDAM in supporting early diagnosis through gait analysis.

Thus, the specific contributions of the paper can be summarized as follows:

- the introduction of EDAM, emphasizing its key features and the integration of explainable AI techniques for clinical decision support;
- the creation of the largest dataset (to our knowledge) by merging several existing datasets, enabling the most comprehensive validation of machine learning-based prediction models for Parkinson's disease detection and gait classification;
- an extensive evaluation of the effectiveness of EDAM in predicting Parkinson's disease and classifying various gait patterns. The evaluation leverages a wide range of features, such as 3D joint trajectories, rotations, step analysis, and energy images (Han and Bhanu, 2005)—more than any prior study has combined simultaneously—allowing for a comprehensive analysis of their collective impact on model performance.

The remainder of the paper is organized as follows. In Section 2 we provide an overview of the EDAM system, while in Section 3 and Section 4 we present the study we conducted to evaluate the EDAM DSS and the achieved results, respectively. Section 6 concludes the paper, after a discussion of the related literature (Section 5).

2 EDAM OVERVIEW

EDAM (Explainable Diagnosis Recommender) is an advanced diagnosis support system integrating motion analysis with other clinical biomarkers, enhancing disease prediction and early diagnosis through XAI. The system addresses several limitations of traditional ML-based diagnostic approaches, which often focus narrowly on single data streams and deliver predictions without offering comprehensible explanations to specialists.

2.1 Architecture and Functionality

The EDAM system is built upon a modular architecture that enables it to process data from various motion analysis systems and clinical devices, providing

both diagnostic insights and explainability for medical professionals. The key components and functionalities of EDAM are depicted in Figure 1.

EDAM is designed to capture data from a wide range of devices, including motion-tracking systems such as Vicon and Azure Kinect DK, heart rate monitors like the Polar H10, and EEG/EMG sensors such as DSI-7 and Cometa Mini Wave. The system is device-agnostic, allowing seamless integration of new sensors with minimal effort through the use of dedicated drivers. These drivers ensure that data from different devices is formatted into a standardized structure, enabling consistent and uniform processing by the system. This flexibility allows EDAM to collect a comprehensive set of data, including gait parameters, heart rate, EEG, and EMG signals, providing a richer, multidimensional analysis that surpasses traditional univariate motion analysis approaches.

The acquired data are stored in a hybrid database system: A relational database for medical records and a NoSQL database for sensor-derived data like gait dynamics and clinical biomarkers. This architecture allows for efficient querying and flexible data management, facilitating real-time and historical analyses.

The core of EDAM is its Decision Support System (DSS), which leverages pre-trained machine learning models to automatically diagnose specific diseases based on collected data. EDAM currently includes models for detecting gait deviations, predicting Parkinson's disease, and calculating the Dynamic Gait Index (DGI) (Shumway-Cook and Woollacott, 1995) (Balletti et al., 2024). Its flexible design allows for the easy integration of additional predictive models through standardized APIs, ensuring scalability for future enhancements.

A standout feature of EDAM is its Explainable AI (XAI) module, which uses SHapley Additive exPlanations (SHAP) (Lundberg and Lee, 2017) to provide interpretable insights into the model's predictions. SHAP values highlight the individual and combined contributions of features such as gait speed or EMG signals to a diagnosis. This interpretability is crucial in clinical settings, where understanding the reasoning behind a diagnosis fosters trust and supports informed decision-making by specialists.

The XAI module also generates textual explanations, clearly outlining the key factors that influenced the prediction. These pre-reports are designed for easy interpretation by medical professionals, offering transparency and aiding early diagnosis. Additionally, the system incorporates visual aids like force plots and image analyses (*e.g.*, Gait Energy Images and Skeleton Energy Images (Han and Bhanu, 2005)) to further illustrate the influence of specific features on the predictions.

The continuous learning capability of EDAM ensures it stays current with evolving clinical knowledge. After each diagnosis, validated data from specialists are added to the knowledge base, and once enough new data is gathered, the ML models are retrained to enhance accuracy, keeping EDAM up-to-date with the latest advancements.

2.2 Data Acquisition and Analysis

The EDAM system facilitates gait data acquisition and analysis by integrating various sensor types. The process starts with the technical operator selecting a patient from a list. Once the patient is chosen, the operator initiates the data acquisition by selecting the "Acquisition" option, which opens a window to specify the examination type. The operator then selects the data sources, including posture (kinematics), EEG (brain activity), EMG (muscle activity), and heart rate. While posture data is mandatory, other sources are optional depending on the exam's focus. Appropriate devices are chosen for each source. If EMG data is included, the system provides a configuration page where the operator assigns sensors to muscles using a drag-and-drop interface (see Figure 2), with visual indicators confirming correct sensor-to-muscle pairing.

After configuration, the system presents both a visual and textual guide for the patient, including a video demonstration of the required movement and a detailed text description. This ensures the patient clearly understands the task to be performed.

During acquisition, the system captures real-time data from the selected sources and displays it for the operator. A 3D avatar visually represents the patient's movements, providing front, left, and right views of the gait (see Figure 3). The operator can zoom in on specific body parts for a closer examination of movement details. Device status is shown using color-coded icons: green for successful data transmission, red for errors, and yellow for connection attempts. After the exercise is completed, the operator selects "End Acquisition" to conclude the session, at which point a window appears for adding session notes, which can be saved or discarded based on the operator's evaluation.

After the data is acquired, it is analyzed through a comprehensive dashboard. The operator can select specific joints or muscles for detailed analysis, displayed in graphical form. If EMG data was collected, the system allows the operator to choose specific muscles for analysis; otherwise, it directly displays graphs of the available data (see Figure 4). The dashboard is

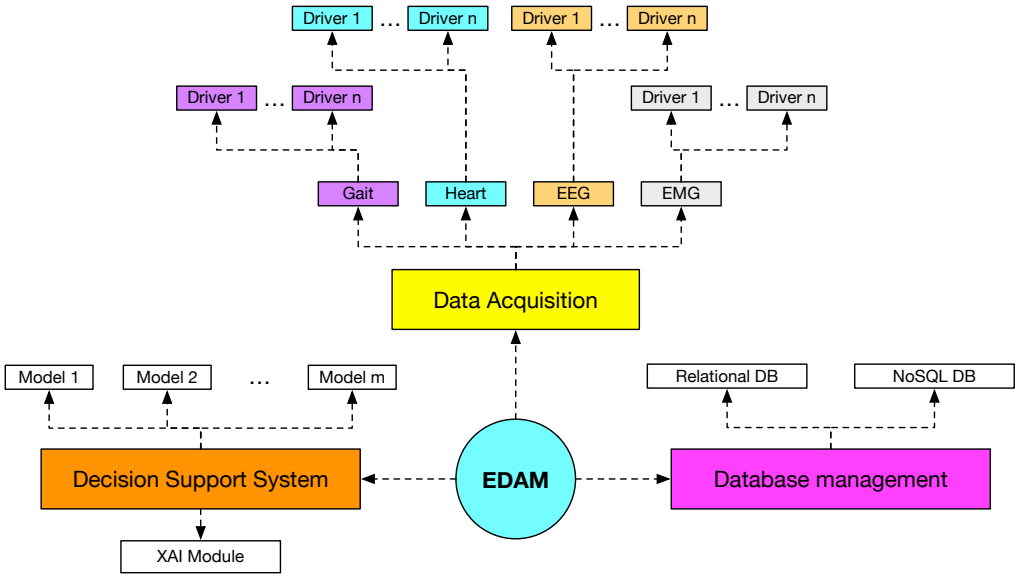

Figure 1: The EDAM architecture.

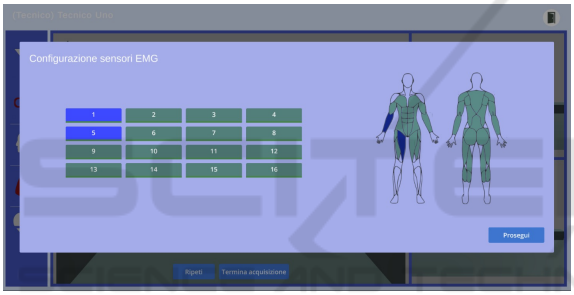

Figure 2: Configuration page for EMG data acquisition.

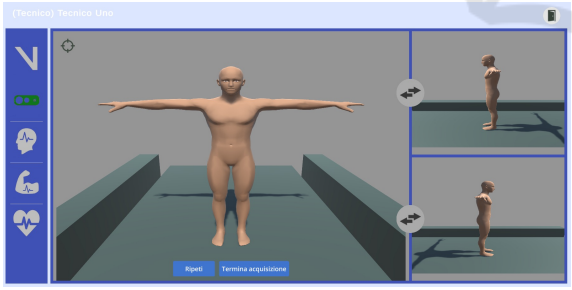

Figure 3: Real-time data acquisition interface of EDAM.

divided into two sections: on the left, a video of the gait cycle is shown alongside predictions generated by machine learning models and the Dynamic Gait Index (DGI); on the right, synchronized graphs display data from posture, EMG, EEG, and heart rate sensors. These synchronized visuals provide a holistic view of the data captured during the session.

The Explainable AI (XAI) module of EDAM provides detailed insights into how predictions are made, offering visual aids such as force plots and image-based analyses like Gait Energy Images (GEI) and Skeleton Energy Images (SEI) to highlight the features influencing predictions (see Figure 5). For instance, the system can identify specific areas of the body contributing to the classification of walking patterns or the diagnosis of conditions such as Parkinson's disease (see Figure 5). This transparency fosters trust by helping medical professionals understand the reasoning behind the system's predictions.

Additionally, the operator can generate a pre-report that summarizes the collected data and analysis. This report, which includes system-generated predictions and explanations (see Figure 5), as well as any observations made by the operator, can be reviewed and refined by specialists, such as physiatrists.

3 EVALUATION OF EDAM PARKINSON'S DETECTION CAPABILITIES

We conducted an empirical evaluation to assess the effectiveness of the EDAM DSS in two key areas: (i) automatically distinguishing between individuals with Parkinson's disease and healthy subjects, and (ii) classifying gait patterns to support the early detection of Parkinson's disease. This second study focuses specifically on identifying subtle gait deviations that may indicate the early stages of Parkinson's disease.

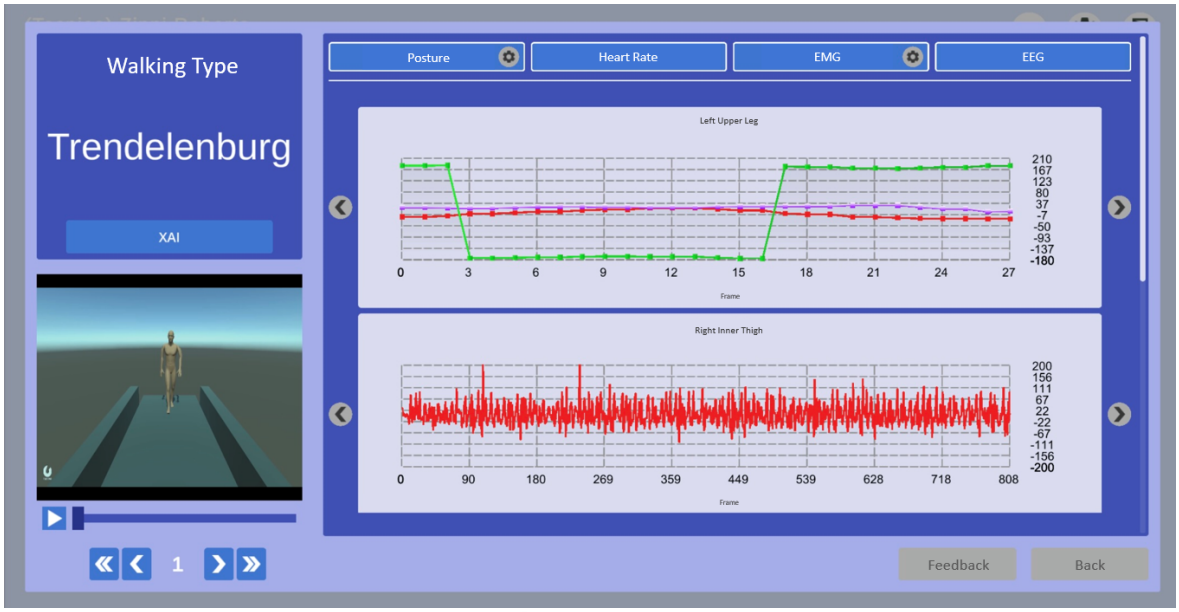

Figure 4: EDAM dashboard displaying gait cycle video, predictions, and synchronized sensor data.

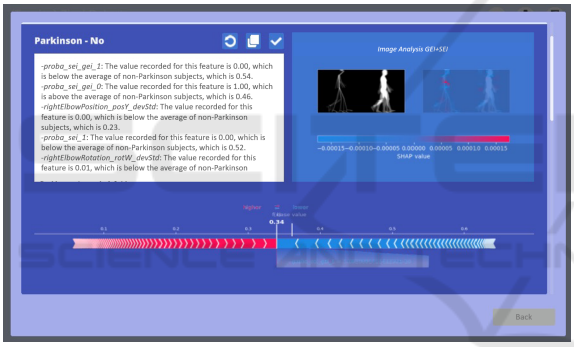

Figure 5: Explainable AI (XAI) Interface Displaying Gait Energy Images, Force Plots, and System-Generated Predictions for Transparent Analysis.

3.1 Study Definition

The study is guided by the following interconnected research questions:

- RQ_1: *To what extent can EDAM distinguish a healthy subject from a subject with Parkinson's disease?*
- RQ_2: *To what extent can EDAM automatically classify a subject's gait to support early Parkinson's detection?*

The first research question assesses EDAM accuracy in distinguishing between healthy individuals and those with Parkinson's disease. Building on this, the second research question evaluates EDAM ability to classify a subject's gait into one of six categories: *Antalgic*, *Lurch*, *Normal*, *Steppage*, *Stiff-Legged*, and *Trendelenburg*. This link between the two research questions is crucial, as certain gait types, such as *Steppage* or *Stiff-Legged*, are known to be early indicators of Parkinson's disease. By accurately classifying these gait patterns, EDAM can detect subtle deviations in movement that may suggest the early onset of Parkinson's, thereby facilitating early diagnosis. Thus, while the first question focuses on distinguishing known cases of Parkinson's, the second question explores EDAM potential for early detection through gait analysis.

3.2 Context of the Study

To evaluate EDAM's capabilities in detecting Parkinson's disease and classifying gait types, we compiled a comprehensive and standardized dataset by integrating several well-established datasets from the literature, including those referenced in the following studies:

- Mehrizi et al., 2019: This dataset includes gait recordings from 23 patients with Parkinson's disease, 22 with postural stroke, 25 with orthopedic issues, and 25 healthy controls. Participants walked on a treadmill for about a minute, while two digital cameras captured their movement, and a motion capture system tracked reflective markers placed on key body joints. The dataset includes 24 time series representing the 3D position of 8 body joints in three directions (x, y, z), collected at a 100 Hz sampling rate. The goal was to detect health problems related to gait using deep

neural networks for pose estimation.

- Schreiber and Moissenet, 2019: This dataset contains gait data from 50 participants (24 women and 26 men), acquired during a single session where participants walked naturally on a 10-meter walkway. Five walking speed conditions were recorded, from slow to fast, using a 10-camera optoelectronic system sampling at 100 Hz. The system tracked 3D trajectories of 52 reflective skin markers placed on anatomical landmarks. The dataset is designed for analyzing human gait at different walking speeds.
- Jun et al., 2020: This dataset consists of gait data from 10 healthy participants, each simulating five different pathological gaits (antalgic, stiff-legged, lurch, steppage, and Trendelenburg). The data was captured using six Microsoft Kinect v2 sensors, providing 3D coordinates for 25 body joints. Each participant walked 20 times per gait type, generating 3D skeletal data for the classification of different gait patterns. The simulation guidelines focused on replicating mechanical limitations, such as restricted joint movement.
- Kour et al., 2020: This dataset includes gait recordings from 96 subjects: 50 with knee osteoarthritis, 16 with Parkinson's disease, and 30 healthy individuals. Each participant performed two gait sequences (left to right and right to left) in the frontal/sagittal plane. The data was collected using a NIKON DSLR camera positioned 8 meters away from the walking path, and six passive reflective markers attached to the subjects' joints. The dataset is in video format (.mov), designed for analyzing gait differences between healthy individuals and those with musculoskeletal or neurological conditions.

These datasets were initially developed for training machine learning (ML) models in gait analysis and Parkinson's disease prediction. Still, they varied in terms of the features they provided, leading to challenges in ensuring reliable and reproducible model performance across different clinical and commercial settings. To address this issue, we merged instances from various datasets and enhanced the dataset by calculating missing features, creating a unified resource that covers all relevant domains.

The analysis of these datasets indicated that the most commonly used gait features fall into three domains: (i) 3D trajectories of body joints, (ii) step-related features such as swing and stance phases, and (iii) Gait Energy Image (GEI) and Skeleton Energy Image (SEI) data. However, none of the original datasets contained a complete set of features from all domains. Therefore, we employed an extensive feature engineering process to create a cohesive dataset.

For datasets that contained only 3D body joint trajectories, we derived additional step-related features (*e.g.*, swing and stance) using custom-built feature engineering strategies (Amboni et al., 2021). Furthermore, using the 3D trajectory data, we animated a mannequin and reconstructed GEI and SEI data, providing a more comprehensive feature set. In cases where only exercise execution videos were available, we utilized the Plask tool[1] to extract 3D trajectories, allowing us to compute all necessary features using the same procedures applied to the other datasets.

Through this process, we constructed the largest dataset ever used in the literature for gait classification and Parkinson's disease prediction, featuring a total of 557 distinct features. These include 304 features related to joint rotations, 228 features representing body trajectories, 17 features capturing pitch, and 6 features derived from GEI and SEI data. The dataset encompasses 7,303 labelled instances, categorized as follows: 20 instances labelled as Parkinson's disease, 34 as healthy controls, and the remaining instances divided among six different gait types (*e.g.*, antalgic, lurch).

3.3 Experimental Procedure

To address RQ_1 and RQ_2, we adopted a Leave-One-Subject-Out (L1SO) cross-validation approach (Hastie et al., 2009). In this method, the dataset is divided into *n* folds, each corresponding to a different patient. For each iteration, one fold is used as the test set, while the remaining *n-1* folds are used for training. This ensures that the data of each patient are included in the training set *n-1* times and in the test set only once, preventing the model from being tested on data from the same patient. This design simulates a real-world scenario where predictions are made for a patient being tested for the first time.

As previously mentioned, the dataset used for the experiments consists of 7,303 labeled instances. For RQ_1, we used the 20 instances labelled as Parkinson's disease and the 34 instances labelled as healthy, totalling 54 instances. For RQ_2, the remaining 7,093 instances, corresponding to the six different gait types, were used to evaluate EDAM gait classification accuracy.

We evaluated the proposed approach in two different scenarios: *Lower body*, where only features from the lower part of the body were used, and *Full body*, where features from both the lower and upper body were provided as input.

[1] https://plask.ai/en-US

During the experimentation, we applied two feature engineering techniques: (i) correlation analysis to discard features with a correlation index higher than 0.95 (Guyon and Elisseeff, 2003), and (ii) automatic feature selection to identify the most relevant descriptors for classification (Li et al., 2017).

We tested 13 different machine learning models (Hastie et al., 2009): Random Forest (RF), Multi-Layer Perceptron (MLP), Logistic Regression (LR), K-Nearest Neighbors (KNN), Gaussian Naive Bayes (GNB), Stochastic Gradient Descent (SGD), Decision Tree (DT), Bagging Classifier (BC), Gradient Boosting Classifier (GBC), AdaBoost (AB), Passive Aggressive Classifier (PAC), Extra Trees Classifier (ETC), and Support Vector Machine (SVM).

3.4 Evaluation Metrics

The following metrics (Hastie et al., 2009) were used to evaluate the performance of the EDAM DSS and address our research questions:

- **Accuracy:** The ratio of correctly classified instances to the total number of instances.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}$$

- **Precision:** The ratio of correctly classified positive instances to the total number of instances classified as positive.

$$\text{Precision} = \frac{TP}{TP+FP}$$

- **Recall:** The ratio of correctly classified positive instances to the sum of correctly classified positive instances and those incorrectly classified as negative.

$$\text{Recall} = \frac{TP}{TP+FN}$$

- **F1-Score:** The harmonic mean of precision and recall.

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Before evaluating the machine learning models in relation to our research questions, we performed a feature analysis using Principal Component Analysis (PCA) (Wold et al., 1987) to identify the most relevant features for (i) predicting Parkinson's disease and (ii) classifying gait. This analysis was carried out for both the *Lower body* and *Full body* scenarios.

4 ANALYSIS OF THE RESULTS

In this section, we present the analysis of results for the two research questions (RQs) of our study.

4.1 RQ$_1$: Parkinson's Prediction

The PCA analysis conducted for both the *Lower Body* and *Full Body* scenarios revealed important insights regarding the relevance of various feature domains in predicting Parkinson's disease. In the *Lower Body* scenario, PCA showed that 32 features explained 97.8% of the variance (see Figure 6), including 20 from the rotation domain, 7 from the pitch domain, 4 from trajectories, and 1 from the GEI/SEI domain. This result is significant as it highlights the contribution of each feature domain, particularly the rotation domain, which was introduced in EDAM and has not been extensively explored in prior studies.

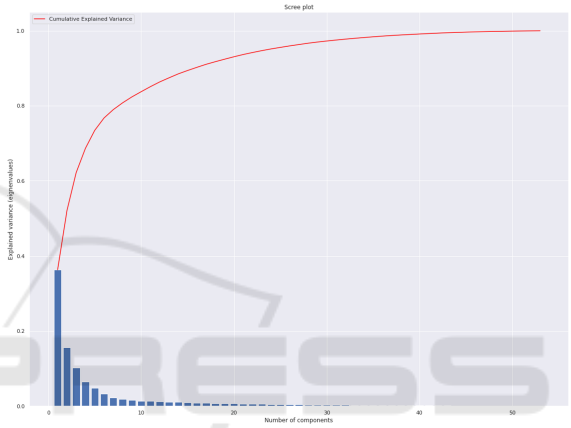

Figure 6: Parkinson (Lower body): Variance Analysis.

In the *Full Body* scenario, PCA revealed that 33 features explained 97.5% of the variance (see Figure 7). Notably, none of the features from the GEI/SEI domain were included in this set; the features were instead distributed across the rotation (25 features), trajectory (6 features), and pitch (2 features) domains. This suggests that GEI/SEI features may be more valuable when only lower body information is available. Importantly, the rotation features once again proved to be crucial in identifying Parkinson's, reinforcing their relevance across both scenarios.

Focusing on the accuracy of detecting subjects with Parkinson's disease, experimental results indicate that in the *Lower Body* scenario, 170 out of the 384 tested machine learning pipelines achieved 100% accuracy. Similarly, in the *Full Body* scenario, 168 pipelines reached perfect accuracy. While these results are highly encouraging, further experimentation is necessary to validate the generalizability of the findings across more diverse datasets.

Nevertheless, it is worth noting that the combination of numerical features, such as 3D joint trajectories and rotational data, with graphical features like GEI and SEI, clearly enhances the overall predic-

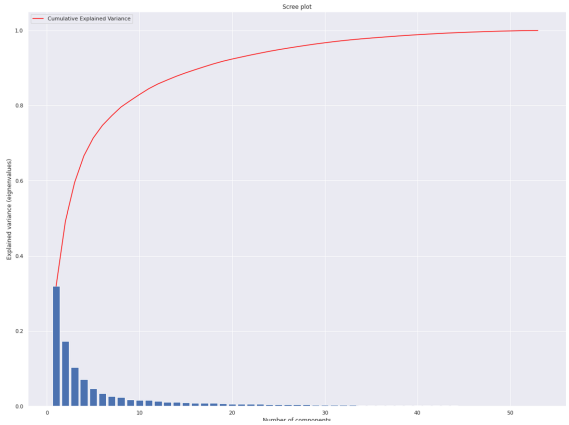

Figure 7: Parkinson (Full body): Variance Analysis.

tion accuracy. As shown in Table 1, a system relying solely on GEI and SEI features achieved an accuracy of approximately 94%, underscoring their valuable contribution to the prediction process.

Table 1: Performance of EDAM Parkinson's prediction model based exclusively on GEI and SEI features.

	Accuracy	Precision	Recall	F1-Score
Mean	0.94	1.00	0.94	0.95
Median	1.00	1.00	1.00	1.00
Std.dev	0.21	0.00	0.21	0.18

To further analyze the contribution of individual features, we examined one of the models with the highest accuracy, which was based on a Decision Tree. This model did not use synthetic oversampling or feature correlation analysis but applied an automatic feature selection algorithm based on Random Forest. Decision Trees were chosen not only for their accuracy but also for their interpretability, making them well-suited for the EDAM system, which aims to generate understandable preliminary reports (pre-reports) based on predictions.

As seen in Table 2, the distribution of selected features across the four domains (Rotations, Trajectories, Step, and GEI/SEI) for both the *Lower Body* and *Full Body* scenarios reinforces the findings from the PCA analysis. This highlights the importance of considering all feature domains in knowledge base of EDAM, with rotation and trajectory features proving especially influential.

Answer to RQ$_1$. The evaluation of the EDAM DSS demonstrates that the selected machine learning models, in both the *Lower Body* and *Full Body* scenarios, achieved high accuracy in predicting Parkinson's disease. The distribution of features across the domains of rotations, trajectories, step, and GEI/SEI underscores the importance of each domain in the knowledge base of the system, further supporting the insights gained from the PCA analysis.

Table 2: Feature distribution across *Lower Body* and *Full Body* scenarios.

Domain	Lower Body	Full Body
Rotations	208	304
Trajectories	132	228
Step	13	17
GEI/SEI	6	6
Total Features	359	557

4.2 RQ$_2$: Gait Classification

In the *Lower Body* scenario, PCA revealed that 113 features explained 98.4% of the variance (see Figure 8). Of these, 76 were from the rotation domain, 18 from the step domain, 18 from trajectories, and 1 from the GEI/SEI domain. This finding underscores the relevance of all feature domains in the context of gait classification, particularly highlighting the rotation domain, which, although rarely explored in the literature, proves to be especially important in this context.

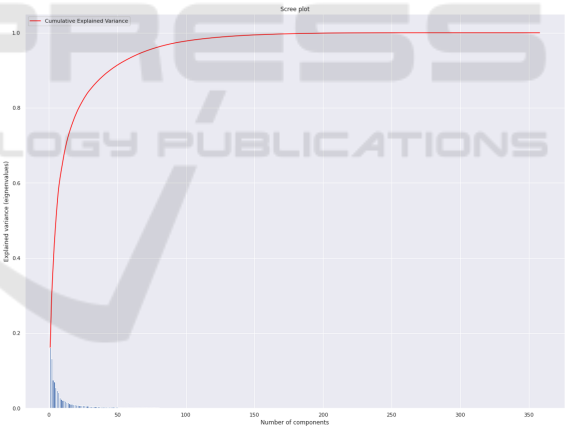

Figure 8: Gait Classification (Lower body): Variance Analysis.

A similar result was observed in the *Full Body* scenario (see Figure 9), where PCA identified 174 features that explained 98.8% of the variance. In contrast to the *Lower Body* scenario, no GEI/SEI features were included in this set. Specifically, the 174 features came from the rotation domain (124), the trajectory domain (26), and the step domain (21). This aligns with the findings from the Parkinson's prediction analysis, suggesting that GEI/SEI features may be particularly useful when only lower body motion data is available, while the rotation domain consistently plays a crucial role in gait classification.

Table 3: Confusion matrix for the EDAM gait classification model in the *Lower Body* scenario.

Actual	**Antalgic**	**Lurch**	**Normal**	**Steppage**	**Stiff-Legged**	**Trendelenburg**
Antalgic	2,122	16	32	38	62	108
Lurch	0	2,308	0	2	0	20
Normal	4	2	2,346	0	0	40
Steppage	44	28	8	2,246	2	42
Stiff-Legged	84	2	36	0	2,208	42
Trendelenburg	206	98	386	20	14	1,636

Table 4: Results obtained from the gait classification model in the *Lower body* scenario.

Class	**Precision**	**Recall**	**F1-score**
Antalgic	0.86	0.89	0.88
Lurch	0.94	0.99	0.96
Normal	0.84	0.98	0.90
Steppage	0.97	0.95	0.96
Stiff-Legged	0.97	0.93	0.95
Trendelenburg	0.87	0.69	0.77
Global Accuracy			**0.91**

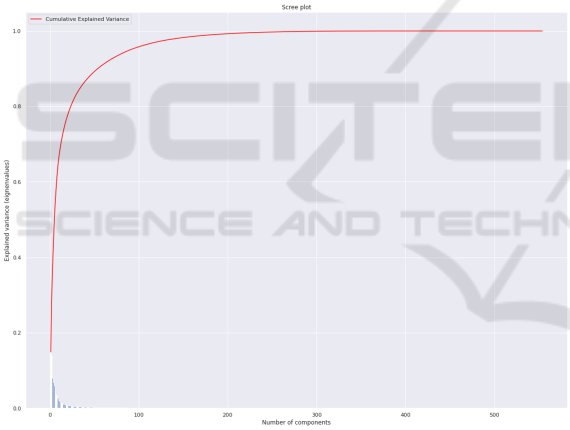

Figure 9: Gait Classification (Full body): Variance Analysis.

For gait classification, the results in the *Lower Body* scenario demonstrated that the best performance was achieved using a machine learning pipeline based on a linear Support Vector Machine (SVM) classifier. This pipeline did not use synthetic oversampling or feature correlation analysis, but included automatic feature selection via Logistic Regression. The model was built using 124 features: 12 from the rotation domain, 94 from the trajectory domain, 15 from the step domain, and 3 from the GEI/SEI domain. This outcome aligns with the findings of the PCA analysis, reinforcing the critical importance of trajectory features in the gait classification process for the lower body scenario. The model achieved an accuracy of 91%, indicating strong performance in classifying gait types.

Table 3 presents the confusion matrix for the gait classification model in the *Lower Body* scenario, while Table 4 reports the results obtained in terms of precision, recall and F1-score for each class.

The analysis of the results shows that both the *Steppage* and *Stiff-legged* gait classes, which are associated with early neurological symptoms, performed very well. The *Steppage* class achieved metrics exceeding 95%, while the *Stiff-legged* class, often linked to early signs of Parkinson's disease, reached an accuracy of 97% and a recall of 93%. These findings underscore the potential of EDAM in the early detection of Parkinson's disease, demonstrating its ability to effectively identify key gait deviations associated with the onset of the condition.

In the *Full Body* scenario, similar to the *Lower Body* scenario, the best performance was achieved using a linear Support Vector Machine (SVM) classification algorithm. This pipeline also excluded synthetic oversampling and feature correlation analysis, but employed an automatic feature selection algorithm based on Extra Trees. The model was constructed with 250 features, including 131 from the rotation domain, 63 from trajectories, 5 from the step domain, and 6 from the GEI/SEI domain. Unlike the *Lower Body* scenario, the most influential features for gait classification in this case were those from the rotation domain, aligning with the findings from Parkinson's disease prediction.

Regarding the accuracy metrics, Table 5 and Table 6 present the precision, recall, and F1-score values for each class. In this scenario, the highest accuracy was achieved for the *Lurch* class. However, the *Steppage* and *Stiff-legged* classes continued to perform well, with recall values of 94% and 95%, respectively, and precision values of 96% and 97%, respectively. These results reinforce the robustness of EDAM in classifying early symptoms of neurological conditions such as Parkinson's disease.

Answer to RQ$_2$. The gait classification models in the *Lower Body* scenario achieved 91% accuracy,

Table 5: Confusion matrix of the EDAM gait classification model in the *Full body* scenario.

Actual	Antalgic	Lurch	Normal	Steppage	Stiff-Legged	Trendelenburg
Antalgic	2,042	4	48	102	12	170
Lurch	8	2,262	2	12	28	18
Normal	48	0	2,312	10	0	22
Steppage	52	22	10	2,278	2	6
Stiff-Legged	30	6	10	8	2,308	16
Trendelenburg	134	18	186	2	10	2,010

Table 6: Results obtained from the gait classification model in the *Full body* scenario.

Class	Precision	Recall	F1-score
Antalgic	0.88	0.86	0.87
Lurch	0.98	0.97	0.97
Normal	0.90	0.97	0.93
Steppage	0.94	0.96	0.95
Stiff-Legged	0.95	0.97	0.97
Trendelenburg	0.90	0.85	0.87
Global Accuracy			**0.93**

with trajectory features playing a key role, especially for the *Steppage* class, which is associated with early neurological symptoms. Similarly, in the *Full Body* scenario, the models reached 93% accuracy, where rotation features proved most influential. Notably, in both scenarios, the *Stiff-Legged* class, which is closely linked to the early signs of Parkinson's disease, showed strong performance, achieving high precision and recall. These results underscore the potential of EDAM to effectively identify gait patterns that could be early indicators of Parkinson's disease, demonstrating its capability for early diagnosis.

5 RELATED WORK

Numerous studies in the literature have focused on evaluating the effectiveness of machine learning (ML) techniques in monitoring and predicting specific diseases based on gait and posture data acquired through motion analysis systems (*e.g.,* (Abdulhay et al., 2018; Costa et al., 2016; Cuzzolin et al., 2017; Mannini et al., 2016; Raknim and Lan, 2016)). Daliri, 2012 proposed a system that combines time series data from foot signals with Support Vector Machine (SVM) algorithms to predict diseases, including ALS. Additionally, Ajay et al., 2018 introduced a system for analyzing and classifying parkinsonian gait using videos captured by pervasive devices (*e.g.,* smartphones, webcams, and surveillance cameras) through a skeleton extraction model that directly detects joint information from video frames. Vilas-Boas et al., 2021 evaluated the use of ML techniques to build a model capable of identifying the Val30Met mutation based on gait characteristics. The study utilized the Kinect v2 sensor to capture 24 gait parameters while individuals walked toward the camera. Multiple machine learning algorithms were tested, including k-nearest neighbors (KNN), decision trees, random forest, SVM, and Multilayer Perceptron. The authors constructed a model with an average accuracy of 92% in distinguishing healthy individuals from mutation carriers (with or without symptoms), and 98% accuracy in distinguishing between asymptomatic and symptomatic carriers (both using SVM). Zhang and Ma, 2019 investigated the application of supervised machine learning algorithms in classifying sagittal gait patterns in children with spastic diplegia. Gait parameters were extracted from data obtained from 200 children, and the results demonstrated that an artificial neural network (ANN) achieved an accuracy of 93.5%, proving to be a promising tool for automatic interpretation of gait data. The literature highlights the benefits of gait analysis in assessing motor deficits, as gait is a fundamental, physiological, and unforced form of locomotion with direct clinical relevance. However, current systems that focus on videographic gait analysis often produce variable and non-repeatable results (Abada et al., 2013; Guillot et al., 2008; Hampton and Amende, 2009; Mead et al., 2011; Vinsant et al., 2013; Wooley et al., 2005). This variability is not surprising, as many key changes in limb positioning and movement dynamics are only visible from the lateral plane. Consequently, recent efforts have focused on developing systems that incorporate machine learning algorithms (*e.g.,* NeuroCube) and lateral view analysis (*e.g.,* MotoRater, Locomouse (Machado et al., 2015)) to analyze gait more comprehensively (Alexandrov et al., 2015; Bellardita and Kiehn, 2015; de Bruin et al., 2016; Talpalar et al., 2013). However, the full potential of lateral plane videography has yet to be realized, as current analyses are often limited to a few functional aspects and a small number of gait parameters

(Preisig et al., 2016). A multi-camera system could potentially provide better results by enabling detailed joint analysis from different angles.

6 CONCLUSION AND FUTURE WORK

In this paper, we addressed the challenges of gait and posture analysis for the early diagnosis of neurological disorders, particularly Parkinson's disease, through machine learning (ML)-based approaches. Traditional methods often lack integration with clinical biomarkers and fail to provide transparent, explainable predictions, limiting their clinical utility. To address these limitations, we introduced EDAM (Explainable Diagnosis Recommender), a decision support system that integrates posture and gait data with clinical biomarkers using Explainable AI (XAI) techniques.

EDAM not only predicts the likelihood of neurological disorders like Parkinson's disease but also explains its diagnostic decisions through visual and natural-language outputs. This combination enhances the trustworthiness and usability of predictions, supporting specialists in making data-informed decisions that incorporate their intuition, judgment, and experience. Furthermore, EDAM generates pre-reports that assist clinicians in both qualitative and quantitative evaluations of patient conditions.

To validate the effectiveness of EDAM, we constructed one of the largest dataset known in the literature by merging several established gait datasets. This dataset contains 557 features and 7,303 labelled instances, making it the most comprehensive resource for evaluating machine learning models in the context of Parkinson's prediction and gait classification. EDAM achieved high accuracy in distinguishing between healthy individuals and those with Parkinson's disease, as well as in classifying abnormal gait patterns linked to early-stage neurological disorders.

Future works will focus on expanding the range of pathologies covered by the system, improving model generalization across diverse populations, and further refining the interpretability of its predictions through advanced XAI techniques. We also plan to conduct studies to assess the contribution of additional biomarkers in identifying neurological disorders when combined with posture data. Furthermore, we intend to perform experiments with specialists to evaluate the acceptability of the predictions of EDAM and the clarity of their explanations.

ACKNOWLEDGEMENT

This research was supported by Italian General Inspectorate of Military Medical Services (IGESAN) under Military Health Research Plan - Project n. 9908115328 (CIG).

REFERENCES

Abada, Y.-s. K., Nguyen, H. P., Schreiber, R., and Ellenbroek, B. (2013). Assessment of motor function, sensory motor gating and recognition memory in a novel bachd transgenic rat model for huntington disease. *PloS one*, 8(7):e68584.

Abdulhay, E., Arunkumar, N., Narasimhan, K., Vellaiappan, E., and Venkatraman, V. (2018). Gait and tremor investigation using machine learning techniques for the diagnosis of parkinson disease. *Future Generation Computer Systems*, 83:366–373.

Ajay, J., Song, C., Wang, A., Langan, J., Li, Z., and Xu, W. (2018). A pervasive and sensor-free deep learning system for parkinsonian gait analysis. In *2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*, pages 108–111. IEEE.

Alexandrov, V., Brunner, D., Hanania, T., and Leahy, E. (2015). High-throughput analysis of behavior for drug discovery. *European journal of pharmacology*, 750:82–89.

Amboni, M. et al. (2021). Gait analysis may distinguish progressive supranuclear palsy and parkinson disease since the earliest stages. *Scientific Reports*, 11(1):1–9.

Balletti, N., Zinni, R., Russodivito, M., Laudato, G., Scalabrino, S., and Oliveto, R. (2024). A machine learning model for the qualitative assessment of human gait based on video features. In *Proceedings of the 17th International Joint Conference on Biomedical Engineering Systems and Technologies – HEALTHINF*, Rome, Italy.

Bellardita, C. and Kiehn, O. (2015). Phenotypic characterization of speed-associated gait changes in mice reveals modular organization of locomotor networks. *Current Biology*, 25(11):1426–1436.

Buckley, C., Alcock, L., McArdle, R., Rehman, R. Z. U., Del Din, S., Mazzà, C., Yarnall, A. J., and Rochester, L. (2019). The role of movement analysis in diagnosing and monitoring neurodegenerative conditions: Insights from gait and postural control. *Brain sciences*, 9(2):34.

Costa, L., Gago, M. F., Yelshyna, D., Ferreira, J., David Silva, H., Rocha, L., Sousa, N., and Bicho, E. (2016). Application of machine learning in postural control kinematics for the diagnosis of alzheimer's disease. *Computational intelligence and neuroscience*, 2016(1):3891253.

Cuzzolin, F., Sapienza, M., Esser, P., Saha, S., Franssen, M. M., Collett, J., and Dawes, H. (2017). Metric learning for parkinsonian identification from imu gait measurements. *Gait & posture*, 54:127–132.

Daliri, M. R. (2012). Automatic diagnosis of neurodegenerative diseases using gait dynamics. *Measurement*, 45(7):1729–1734.

de Bruin, N., Schmitz, K., Schiffmann, S., Tafferner, N., Schmidt, M., Jordan, H., Häußler, A., Tegeder, I., Geisslinger, G., and Parnham, M. (2016). Multiple rodent models and behavioral measures reveal unexpected responses to fty720 and dmf in experimental autoimmune encephalomyelitis. *Behavioural brain research*, 300:160–174.

Edwards, L. and Veale, M. (2017). Slave to the algorithm? why a'right to an explanation'is probably not the remedy you are looking for. *Duke L. & Tech. Rev.*, 16:18.

Guillot, T. S., Asress, S. A., Richardson, J. R., Glass, J. D., and Miller, G. W. (2008). Treadmill gait analysis does not detect motor deficits in animal models of parkinson's disease or amyotrophic lateral sclerosis. *Journal of motor behavior*, 40(6):568–577.

Guyon, I. and Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of machine learning research*, 3:1157–1182.

Hampton, T. G. and Amende, I. (2009). Treadmill gait analysis characterizes gait alterations in parkinson's disease and amyotrophic lateral sclerosis mouse models. *Journal of motor behavior*, 42(1):1–4.

Han, J. and Bhanu, B. (2005). Individual recognition using gait energy image. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, volume 2, pages i–i. IEEE.

Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer Science & Business Media.

Holzinger, A., Biemann, C., Pattichis, C. S., and Kell, D. B. (2017a). What do we need to build explainable ai systems for the medical domain? *arXiv preprint arXiv:1712.09923*.

Holzinger, A., Malle, B., Kieseberg, P., Roth, P. M., Müller, H., Reihs, R., and Zatloukal, K. (2017b). Towards the augmented pathologist: Challenges of explainable-ai in digital pathology. *arXiv preprint arXiv:1712.06657*.

Jun, K., Lee, Y., Lee, S., Lee, D. W., and Kim, M. S. (2020). Pathological gait classification using kinect v2 and gated recurrent neural networks. *IEEE Access*, 8:139881–139891.

Kour, N., Gupta, S., and Arora, S. (2020). Gait dataset for knee osteoarthritis and parkinson's disease analysis with severity levels.

Li, J., Cheng, K., Wang, S., Morstatter, F., Trevino, R. P., Tang, J., and Liu, H. (2017). Feature selection: A data perspective. *ACM Computing Surveys (CSUR)*, 50(6):1–45.

Lundberg, S. M. and Lee, S.-I. (2017). A unified approach to interpreting model predictions. In *Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS)*, pages 4765–4774. Curran Associates Inc.

Machado, A. S., Darmohray, D. M., Fayad, J., Marques, H. G., and Carey, M. R. (2015). A quantitative framework for whole-body coordination reveals specific deficits in freely walking ataxic mice. *elife*, 4:e07892.

Mannini, A., Trojaniello, D., Cereatti, A., and Sabatini, A. M. (2016). A machine learning framework for gait classification using inertial sensors: Application to elderly, post-stroke and huntington's disease patients. *Sensors*, 16(1):134.

Mead, R. J., Bennett, E. J., Kennerley, A. J., Sharp, P., Sunyach, C., Kasher, P., Berwick, J., Pettmann, B., Battaglia, G., Azzouz, M., et al. (2011). Optimised and rapid pre-clinical screening in the sod1g93a transgenic mouse model of amyotrophic lateral sclerosis (als). *PloS one*, 6(8):e23244.

Mehrizi, R., Peng, X., Zhang, S., Liao, R., and Li, K. (2019). Automatic health problem detection from gait videos using deep neural networks. *arXiv preprint arXiv:1906.01480*.

Preisig, D. F., Kulic, L., Krüger, M., Wirth, F., McAfoose, J., Späni, C., Gantenbein, P., Derungs, R., Nitsch, R. M., and Welt, T. (2016). High-speed video gait analysis reveals early and characteristic locomotor phenotypes in mouse models of neurodegenerative movement disorders. *Behavioural brain research*, 311:340–353.

Raknim, P. and Lan, K.-c. (2016). Gait monitoring for early neurological disorder detection using sensors in a smartphone: Validation and a case study of parkinsonism. *Telemedicine and e-Health*, 22(1):75–81.

Schreiber, C. and Moissenet, F. (2019). A multimodal dataset of human gait at different walking speeds established on injury-free adult participants. *Scientific Data*, 6(1):1–7.

Shumway-Cook, A. and Woollacott, M. H. (1995). The dynamic gait index to quantify gait ability in patients with vestibular and balance disorders. *Physical therapy*, 75(6):538–548.

Talpalar, A. E., Bouvier, J., Borgius, L., Fortin, G., Pierani, A., and Kiehn, O. (2013). Dual-mode operation of neuronal networks involved in left–right alternation. *Nature*, 500(7460):85–88.

Vilas-Boas, M. D. C., Rocha, A. P., Cardoso, M. N., Fernandes, J. M., Coelho, T., and Cunha, J. P. S. (2021). Supporting the assessment of hereditary transthyretin amyloidosis patients based on 3-d gait analysis and machine learning. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 29:1350–1362.

Vinsant, S., Mansfield, C., Jimenez-Moreno, R., Moore, V. D. G., Yoshikawa, M., Hampton, T. G., Prevette, D., Caress, J., Oppenheim, R. W., and Milligan, C. (2013). Characterization of early pathogenesis in the sod1 g93a mouse model of als: part i, background and methods. *Brain and behavior*, 3(4):335–350.

Wold, S., Esbensen, K., and Geladi, P. (1987). Principal component analysis. *Chemometrics and intelligent laboratory systems*, 2(1-3):37–52.

Wooley, C. M., Sher, R. B., Kale, A., Frankel, W. N., Cox, G. A., and Seburn, K. L. (2005). Gait analysis detects early changes in transgenic sod1 (g93a) mice. *Muscle & Nerve: Official Journal of the American Association of Electrodiagnostic Medicine*, 32(1):43–50.

Zhang, Y. and Ma, Y. (2019). Application of supervised machine learning algorithms in the classification of sagittal gait patterns of cerebral palsy children with spastic diplegia. *Computers in biology and medicine*, 106:33–39.