UNIVERSIDAD NACIONAL DE COLOMBIA

Cognitive impairment inference in Parkinson's disease patients from spatiotemporal gait assessments using machine learning

Jose Elkin Serna Soto

Universidad Nacional de Colombia

Facultad de Medicina, Departamento de Imágenes Diagnósticas

Maestría en Ingeniería Biomédica

Bogotá, Colombia

2025

Cognitive impairment inference in Parkinson's disease patients from spatiotemporal gait assessments using machine learning

Jose Elkin Serna Soto

Thesis or degree work presented as a partial requirement to apply for the title of:

Master in Biomedical Engineering

Director:

Ph.D., Eduardo Romero

Research Line:

Motion and Biosignal Analysis

Research Group:

CIM@LAB (Computer Imaging and Medical Applications Laboratory)

Universidad Nacional de Colombia

Facultad de Medicina, Departamento de Imágenes Diagnósticas

Maestría en Ingeniería Biomédica

Bogotá, Colombia

2025

Dedicatoria

A mi Esposa e Hija, quien impulsan mi camino un paso a la vez y me incentivan a continuar cuando las fuerzas se agotan. Gracias por apoyarme incondicionalmente durante este proceso que hemos recorrido juntos. A mis Padres que me enseñaron a avanzar con resiliencia enfocándome mas en el proceso que en el resultado. A todas las personas (familia, amigos, conocidos) que me han apoyado dando sustento y motivación cuando el camino muestra su lado mas oscuro. A este camino que llamamos vida que me permite recorrerlo y disfrutarlo sin juzgar sus direcciones, dándome la promesa que a medida que lo avance encontrare el sentido de mi existencia.

"Carpe diem, quam minimim credula postero."
"Aprovecha el día de hoy; confía lo menos posible en el mañana"

Horacio (65 - 8 a. de C.).

Acknowledgments

I would like to thank Professor Eduardo Romero for showing me a way of life in research. Thank you for teaching me to love knowledge with passion and dedication, to seek my limits and to expand them despite my resistance. Thank you for showing me that at the frontier of knowledge we are all ignorant. I would like to thank Santiago Toledo for his patience, his help, his kindness and above all the generosity with which he selflessly shared his knowledge with me. I would like to thank all my colleagues at CIM@LAB, whose ideas and conversations have enriched my vision of the world, thank you for courageously presenting your research and giving me the courage to support my own. I would like to thank my family and friends for their unconditional support, without which this collaborative achievement would not have been possible. Finally, I would like to thank the National University of Colombia for giving me the opportunity to enjoy and live in a space that beats with our ideas and grows with our dreams.

Resumen

Inferencia de deterioro cognitivo en pacientes con enfermedad de Parkinson a partir de evaluaciones espaciotemporales de la marcha mediante aprendizaje automático

El deterioro cognitivo es una de las complicaciones más relevantes en la enfermedad de Parkinson (PD), afectando significativamente la calidad de vida de los pacientes. La identificación temprana de estos déficits es crucial para mejorar la intervención clínica y el pronóstico de la enfermedad. Este estudio analiza la relación entre las características espaciotemporales de la marcha y el deterioro cognitivo en pacientes con PD en etapa temprana.

Se utilizaron datos de 48 pacientes de la base de datos de la *Parkinson's Progression Markers Initiative* (PPMI), dividiendo la muestra en dos grupos: pacientes sin deterioro cognitivo (PD) y aquellos con algún grado de deterioro cognitivo (PD-DC). Para el análisis, se implementaron siete algoritmos de aprendizaje automático, optimizando sus hiperparámetros y abordando el desequilibrio de clases. El desempeño de los modelos se evaluó mediante métricas como *Recall*, Precisión, *F1-score* y Exactitud (*Accuracy*), priorizando la sensibilidad para evaluar la capacidad de detección de la clase minoritaria.

Los resultados indicaron que las características espaciotemporales de la marcha permiten diferenciar significativamente entre los grupos PD y PD-DC. Entre los modelos evaluados, el Perceptrón Multicapa (MLP) y el Cognitive Assessment through Gait in Parkinson's Disease (CoGait-PD) presentaron el mejor desempeño, alcanzando valores de exactitud de 0.78 ± 0.08 y 0.77 ± 0.05, respectivamente. Estos modelos lograron un equilibrio entre sensibilidad y precisión, destacándose en la identificación de casos positivos y reduciendo falsos negativos.

Los hallazgos sugieren que las características de la marcha pueden servir como biomarcadores no invasivos para la detección temprana del deterioro cognitivo en

pacientes con PD. Además, los modelos de aprendizaje automático, particularmente CoGait-PD y MLP, muestran un alto potencial en la evaluación clínica. Se recomienda realizar estudios adicionales para validar estos hallazgos y explorar su aplicabilidad en entornos clínicos.

Palabras clave: Deterioro cognitivo, Evaluación de la marcha, Aprendizaje automático interpretable, Biomarcadores no invasivos, Enfermedad de Parkinson

.

Abstract

Cognitive impairment inference in Parkinson's disease patients from spatiotemporal gait assessments using machine learning

Cognitive decline is a significant complication in Parkinson's disease (PD), severely impacting patients' quality of life. Early identification of these deficits is crucial to improving clinical intervention and disease prognosis. This study investigates the relationship between spatiotemporal gait characteristics and cognitive decline in early-stage PD patients. Data from 48 patients were obtained from the *Parkinson's Progression Markers Initiative* (PPMI) database, categorised into two groups: patients without cognitive impairment (PD) and those exhibiting some degree of cognitive decline (PD-CD). Seven machine learning algorithms were implemented, optimising hyperparameters and addressing class imbalance. Model performance was evaluated using *Recall*, Precision, *F1-score*, and Accuracy, prioritising sensitivity to assess the classifiers' ability to detect the minority class.

The results indicate that spatiotemporal gait characteristics significantly differentiate PD and PD-CD groups. Among the evaluated models, the Multilayer Perceptron (MLP) and the Cognitive Assessment through Gait in Parkinson's Disease (CoGait-PD) demonstrated the highest performance, achieving accuracy scores of 0.78 ± 0.08 and 0.77 ± 0.05, respectively. Both models balanced sensitivity and precision, excelling in identifying positive cases and reducing false negatives.

These findings suggest that gait characteristics may serve as non-invasive biomarkers for early detection of cognitive decline in PD patients. Additionally, machine learning models, particularly CoGait-PD and MLP, show strong potential for clinical assessment. Further studies are recommended to validate these findings and explore their applicability in clinical settings.

Keywords: Cognitive impairment, Gait assessment, Interpretive machine learning, Non-invasive biomarkers, Parkinson's disease .

Contents

1. Abbreviations

Abbreviation	**Meaning**
PD	Parkinson's disease or control group
CI	Cognitive Impairment
$PD-CI$	Parkinson with some degree of cognitive impairment
$PD+CI$	PD Patient with cognitive impairment or study group
$PD-D$	Parkinson leading to progression to dementia
MCI	Mild Cognitive Impairment
$PD-MCI$	Parkinson leading to progression to Mild Cognitive impairment
$PPMI$	Parkinson Progression Markers Initiative
$JERK$	Spatiotemporal measures of smoothness
MDS	International Parkinson and Movement Disorders Society
$UPDRS$	Unified Parkinson's Disease Rating Scale
$UPDRSIII$	Unified Parkinson's Disease Rating Scale in its exploration of motor aspects
AD	Alzheimer's Disease
DT	Dual task gait conditions
$CIM@LAB$	The Computer Imaging and Medical Applications Laboratory

Abbreviation	Methods
ML	Machine Learning
DL	Deep Learning
$CoGait - PD$	Cognitive assessment through Gait in Parkinson's Disease
$KQMU$	Kernel Quantum Measurement Unit
$DQOR$	Deep Quantum Ordinal Regressor
QMR	Quantum Measurement Regression
$CNNs$	Convolutional Neural Networks
RFF	Random Fourier Features
MAE	Mean Absolute Error
$MANOVA$	Multivariate Analysis of Variance
TP	True positives
TN	True negatives
FP	False positives
FN	False negatives
RF	Random Forest Classifier
PCA	Principal Component Analysis
$RF10$	The 10 most relevant characteristics
KNN	K-Nearest Neighbor
SVM	Support Vector Machine
LR	Logistic Regression
GP	Gaussian Processes
MLP	Multilayer perceptron

List of Figures

List of Tables

2. Introduction

2.1. Parkinson's disease and Cognitive Impairment

Parkinson's disease (PD) is one of the most prevalent degenerative diseases worldwide. By 2024, the number of people living with Parkinson's disease is expected to exceed 10 million [1]. This represents a significant increase from the estimated 9.4 million people affected in 2020[2]. According to a Lancet editorial [3], the PD prevalence is expected to continue rising, reaching between 12 and 17 million by 2040, driven by increasing life expectancy and the reduction of other death causes. This growing prevalence highlights the increasing PD burden on global health systems, particularly in low- and middle-income countries, where access to treatment is often limited.

Although PD is described as a degenerative brain disease associated with motor symptoms, there are a wide variety of disease manifestations, among which cognitive impairment (CI)[1] has been observed as increasingly present [4], progressing to dementia (PD-D) in late stages of the disease [4]. Cognitive impairment of mild severity (MCI) is common in PD patients [5], and this condition predisposes development of dementia [6, 7, 8]. Overall, gait patterns are considered as PD biomarkers of the cognitive impairment and dementia [9] yet their characterization and quantification is still weak. Previous evidence shows gait changes precede and may predict cognitive decline in ageing and other neurodegenerative pathologies [10, 11, 12, 13]. However, despite documented relations between gait and cognition, correlation is still unclear [12, 14] and previous studies fail to capture the specific impairment since this is usually confused with other disorders [15]. In addition, cohorts in these

studies are small, the biological variability is too high and sources of noise are multiple. Traditional statistical methods, such as linear models or factor analyses, are constrained by assumptions of normality, independence, and linearity, which usually are not hold in complex clinical scenarios. [16, 17, 18]. These limitations emphasize the need for advanced techniques capable of handling complex, multidimensional data and capturing intricate interdependencies, which are essential for identifying early biomarkers of cognitive decline in PD.

2.2. Prediction of Parkinson Cognitive Decline: Gait Patterns

Research about the relationship between PD and MCI has increased in recent years. Biomarkers in magnetic resonance imaging [4, 19, 20] or genetic profiling [4, 19, 15, 21] have been documented to be helpful in classifying and predicting cognitive decline in different domains [21]. These approaches offer reliability and repeatability, but they are expensive and difficult to access.

Gait patterns are promising as an easily accessible and cost-effective biomarker. Morris et al. [9] evaluated gait patterns along with the cognitive function of 119 participants and developed a mixed model with independent gait predictors. Their research concludes that specific gait characteristics, such as pace (step velocity, step length), variability (swing time variability, step stance variability, step length variability), and postural control (step width), predict cognitive performance decline, particularly in fluctuating attention and visual memory. These findings highlight that gait-related predictors are not typically identified through specialised expert assessments. However, the study is limited by its inability to adequately capture non-linearities and complex interactions between variables, emphasising the need for advanced modelling techniques.

Pieruccini-Faria et al [22]. employed statistical methodologies to analyse the relationship between gait parameters and cognitive performance in neurodegenerative diseases. Using factor analysis, they grouped 11 gait parameters into four main

domains: rhythm, stride, variability, and postural control. They then adjusted factor scores for covariates such as age, gender, education, and gait speed. Subsequently, multivariable linear regression models were applied to explore associations between gait domains and cognitive performance, while classification analysis using ROC curves assessed the discriminative ability of gait parameters between clinical groups.

To evaluate differences in gait performance across clinical groups, multivariate analysis of variance (MANOVA) was applied. Specifically, Z-scores of residuals from linear regression models, adjusted for covariates, were entered into the MANOVA, employing Pillai's trace test to detect significant multivariate effects across gait domains. When a significant effect was identified, subsequent univariate tests examined between-group differences in each gait domain.

The authors conclude that gait could serve as a marker of cognitive-cortical dysfunction to identify Alzheimer's-type dementia. However, they acknowledge that the small sample size in certain subgroups may limit the generalisability of their findings. Additionally, they highlight that traditional statistical methods may not fully capture the complex interactions and non-linearities inherent in the gait-cognition relationship.

Recent research has shown an increase in the use of machine learning models such as support vector machine (SVM), random forest (RF) and principal component analysis (PCA) to analyse spatio-temporal gait characteristics and make predictions about cognitive impairment [23, 24, 21]. These studies suggest that cognitive impairment in neurodegenerative diseases such as Alzheimer's and dementia can be inferred from gait parameters using classification methods such as SVM when small data sets are available [25].

Chen et al. [23] used machine learning to develop gait-based classification models to assess patients with neurodegenerative diseases such as Parkinson's with different types of mild cognitive impairment (MCI). This study again demonstrated how kernel-based models achieved good classification performance for detecting different types of MCI based on gait information. Specifically, the PCA-SVM model

attained a classification accuracy of 91.67% and an area under the receiver operating characteristic curve (ROC AUC) of 0.9714 when distinguishing Parkinson's disease-related MCI (PD-MCI) from non-PD MCI. These quantitative measures highlight the model's effectiveness; however, the study does not provide details on the cross-validation procedure used for classification, which is critical for assessing the generalisability of the results. Furthermore, it is not exclusive to Parkinson's disease, as it includes different neurodegenerative pathologies, introducing potential heterogeneity in the dataset. Additionally, SVM performance should be compared with other classical machine learning models to establish its relative efficacy. A more comprehensive analysis incorporating multiple classification approaches and a clearer explanation of the validation process would strengthen the study's conclusions.

Recently, application of probabilistic models [26] naturally integrate the results of individual predictions into a single diagnosis. The decisions made by the model are based on a representative measurement among the samples of the data set, performed in a representation space constructed by a kernel function. The model employs density matrices to represent a joint probability space between input and output data, enabling it to capture all relationships between data samples and labels. This approach boosts the small, daily datasets typically encountered in clinical settings

Recently, the application of probabilistic models [26] has enabled the natural integration of individual predictions into a unified diagnosis. In particular, the Deep Quantum Ordinal Regressor (DQOR) combines deep convolutional neural networks (CNNs) with a probabilistic regression framework based on quantum measurement regression (QMR). This model employs density matrices to represent the joint probability distribution of inputs and outputs, allowing it to capture complex relationships between data samples and labels while quantifying uncertainty. Additionally, by leveraging a representation space constructed through random Fourier features (RFF), DQOR enhances the robustness of predictions.

This probabilistic approach contributes to optimising small clinical datasets by improving model generalisation and diagnostic reliability. Specifically, experimental

results show that DQOR reduces the mean absolute error (MAE) by approximately 8.7% compared to conventional deep learning models and improves classification accuracy by up to 5% in binary diagnostic tasks. Moreover, its uncertainty quantification capabilities allow better identification of misclassified cases, enhancing interpretability and reducing diagnostic errors

2.3. Contribution

A recurring bottleneck in biomedical science is the need to make inferences from small cohorts while ensuring that experimental groups meet certain conditions.
This investigation explores the problem of inferring cognitive decline in a small cohort of Parkinson's disease patients from spatiotemporal gait characteristics using different machine learning strategies. In this context, the small cohort refers specifically to the dataset used for training the models, highlighting the challenge of developing robust predictive frameworks with limited clinical data.
The contribution is developed on the basis of three strategies:

- A framework to infer the course of PD despite having a relatively small set of patients.
- A set of custom techniques that exploit the hidden relationships in the data of these patients at high risk of developing cognitive impairment.
- A clinically interpretable approach to ensure robust support for decision-making and enhance its applicability in the clinical context.

3. Experimental Setup

In this study, the prediction of cognitive impairment was approached as a binary classification task, utilising a dataset consisting of gait assessment records from 48 patients. Given the limited nature of the dataset, 20 random partitions were experimented with, with 70%-30% assigned to the training and test sets, respectively. This approach introduces variability into the training and test sets. The mean classification metrics across the 20 partitions provide a central measure of performance and avoid dependence on a single partition that may be atypical.

Despite the possibility of the reported standard deviation being lower than the actual value due to the non-complete independence of the partitions, the 20-partition approach was selected on the grounds of its simplicity and effectiveness in the maximisation of the use of limited data.The inherent limitations resulting from dependencies between partitions are acknowledged; nevertheless, this method provides a valuable initial approximation for the evaluation of model performance [27].

The following section provides details of the datasets, model architecture, and hyperparameter settings for the training and evaluation metrics.

3.1. Dataset

Data used in this study were extracted from the Parkinson Progression Markers Initiative (PPMI), an international multicenter study designed to identify biomarkers of Parkinson's disease (PD) progression [28]. A total of 48 patient records were included, including individuals at different early stages of PD (prodromal, stage 1 and stage 2). Of these, 33 were cognitively normal (control group), while 15

were diagnosed with cognitive impairment one year after motor assessment (disease group). Patients were recruited from the Tel Aviv Medical Centre in Israel and were part of the PPMI Motor Assessment Database. All subjects were assessed using a gait system consisting of three wireless, lightweight, wearable sensors with axial accelerometers, gyroscopes, and magnetometers (Opal, APDM Ltd). The sensors were placed on the participants' wrists and lower back during all gait measurements to quantify the spatio-temporal measurements.

To be included in the analysis, patient data had to meet the selection criteria, which included complete patient records with spatiotemporal gait assessments in two different protocols: the normal gait assessment (preferred gait speed) and the dual task gait assessment (walking while simultaneously subtracting 3 serially from a predefined 3-digit number).

A total of 36 gait characteristics were selected from the two protocols, reflecting spatio-temporal measures of consistency, rhythmicity, and smoothness (jerk) for the tasks performed during testing (see in appendex Table A.1)

In addition, the cognitive status of the 48 patients was recorded one year after the gait assessments described above. This information was extracted from the cognition data set in the PPMI non-motor assessments. To construct the cognitive categories, the PPMI

assessment group used a combination of responses from the subject or another informant, the investigator's judgment, and the results of six neuropsychological tests covering four cognitive domains (see Table 3.1).

As the cognitive category records showed four different outcomes and three of them represented at least one cognitive deficit in one of their domains, it was decided to combine the observations of the three cognitive labels into a new common outcome, referred to as Cognitive Impairment (CI). This decision was motivated by the small number of records within the four categories and the fact that the categories could provide more clinically relevant information on the occurrence of cognitive impairment when analysed together (see Table 3.2).

As a result, subjects were mainly divided into two categories: (i) patients with

CLINICAL LABELS	
Normal (PD)	No low score on any neuropsychological tests
Cognitive Complaint (PD+CC)	Low score on 1 neuropsychological test
Intermediate (PD+I)	Low score on 2 neuropsychological tests
Mild Cognitive Impairment (PD+MCI)	Marked Cognitive Decline (Expert Decision) + Low score on 2 neuropsychological tests

Table 3.1.: This table delineates the categorization of cognitive states used in the PPMI study, based on comprehensive evaluations involving neuropsychological testing across four cognitive domains, subjective reports from participants or their informants, and clinical judgments by investigators. Each category is defined by specific criteria related to the performance on these tests.

PD who were cognitively normal (PD or control) and (ii) patients with cognitive impairment (PD+CI or disease). The demographic data and the characteristics of the PD of all participants are summarised in Table 3.3, together with their p-values. To compare the two groups, hypothesis tests were conducted. Specifically, independent samples t-tests were applied to continuous variables such as age and education, while chi-square tests were used for categorical variables. The null hypothesis (H_0) assumed no significant differences between groups in terms of demographic and clinical characteristics, whereas the alternative hypothesis (H_1) proposed that at least one characteristic differed significantly between groups.

Significant differences were observed for age and education characteristics, while, more importantly, no significant differences were found for the specific motor impairment characteristic of MDS-UPDRS UPDRSIII. This suggests that traditional assessments designed to measure motor deficits, including UPDRS, do not detect subtle differences in individuals at risk of cognitive impairment in patients with PD.

Clinical Label	Binary Label	Samples
PD	0 (PD)	33
PD+CC	1 (PD+CI)	7
PD+I	1 (PD+CI)	4
PD+MCI	1 (PD+CI)	4

Table 3.2.: To improve clinical relevance, the three cognitive deficit labels were combined into a single category, Cognitive Impairment (CI), given the limited number of records in the four categories. This decision was taken to provide a more comprehensive understanding of the presence of cognitive impairment when analysed together.

3.2. Training and Validation

The experiments were conducted to evaluate different machine learning models for predicting cognitive decline in Parkinson's disease. The models were implemented using the *scikit-learn* library [29] in Python. The evaluated methods include K-Nearest Neighbours (KNN), Support Vector Machine (SVM), Logistic Regression (LR), Random Forest (RF), Gaussian Process Classifier (GPC), and Multilayer Perceptron (MLP). Additionally, CoGait-PD, a probabilistic model integrating the Kernel Density Matrix (KDM), was assessed.

The key hyperparameters for each method were selected based on standard practices and prior research. A detailed description of the models is provided in Section 3.2.1. Hyperparameter tuning procedures are outlined in Section 3.2.2.

3.2.1. Methods

Various methods were evaluated for inferring cognitive decline in Parkinson's disease patients. The subsequent section provides a concise overview of the aforementioned methodologies.

	PD (Control Patients)	PD + CI (Disease Patients)	p-value	Test Used
Number of Samples	33	15	-	-
Gender (F/M)	17/16	7/8	1.0	χ^2
Hispanic or Latino (Y/N)	7/26	7/8	0.15	χ^2
Age (Yr)	63 ± 6.6	68 ± 8.9	0.006	t-test
Education (Yr)	13.1 ± 3.7	15.8 ± 3.6	0.02	t-test
UPDRSIII (Total Score)	8.21 ± 8.32	13.6 ± 13.1	0.28	t-test
Parkinson Diagnosis (Yr)	2.5 ± 3.3	3.2 ± 3.2	0.46	t-test

Table 3.3.: Subjects were categorised into PD (control) and PD+CI (disease) groups. Statistical tests (t-test or χ^2-test) were used to compare variables between groups. Significant differences were found for age and education.

K-Nearest Neighbours (KNN)

A non-parametric classification algorithm that assigns a new sample to the majority class among its K nearest neighbours in the feature space. It is widely used due to its simplicity and effectiveness, particularly in low-dimensional datasets. The main hyperparameters include the number of neighbours K, the distance metric (e.g., Euclidean, Manhattan), and the weighting scheme (uniform or distance-based). Further implementation details can be found in [29].

Support Vector Machine (SVM)

A supervised learning algorithm that identifies the optimal hyperplane maximising the margin between different classes. It is effective in high-dimensional spaces and robust to overfitting when properly regularised. The primary hyperparameters include the choice of kernel (linear, polynomial, or radial basis function), the regularisation parameter C, and kernel-specific coefficients. Further details on implementation are available in [29].

Logistic Regression (LR)

A linear model that estimates the probability of a categorical outcome using a logistic function. It is commonly applied in binary and multinomial classification problems. The main hyperparameters include the regularisation strength C, the penalty type (L1 or L2), and the solver algorithm. A more detailed description is provided in [29].

Random Forest (RF)

An ensemble learning technique based on multiple decision trees, where the final classification is determined by majority voting. It is known for its robustness and ability to handle non-linear relationships. The key hyperparameters include the number of trees, the maximum depth of each tree, and the minimum number of samples required for node splits. Further details can be found in [29].

Gaussian Process Classifier (GPC)

A probabilistic classification model that extends the Gaussian process framework to classification tasks. It provides flexible decision boundaries and quantifies prediction uncertainty. The model's performance is influenced by the choice of covariance function (e.g., radial basis function kernel) and noise level parameters. Further implementation details are available in [29].

Multilayer Perceptron (MLP)

A fully connected artificial neural network that consists of multiple layers of neurons, typically trained using backpropagation. It is capable of learning non-linear patterns from data. The key hyperparameters include the number of hidden layers, the number of neurons per layer, and the activation function (ReLU, tanh, sigmoid). More details on the implementation are provided in [29].

CoGait-PD

A probabilistic machine learning model designed to analyse spatiotemporal gait characteristics for predicting cognitive decline in Parkinson's disease. It constructs a representation of gait patterns that captures clinically relevant features using kernel-based density matrices. Unlike traditional classifiers, it does not assume predefined data distributions but instead learns an empirical representation from a small dataset. The training process employs a probabilistic framework that encodes gait information into density matrices, allowing for interpretable and robust decision-making in clinical applications[30].

Kernel Density Matrix (KDM) As the probabilistic core of CoGait-PD, KDM models the joint probability distribution between gait features and cognitive impairment labels. It operates in a kernel-induced Hilbert space, leveraging density matrices to capture uncertainty and latent structures within the data.
Through Schmidt decomposition, KDM factorises gait-related information, optimising computational efficiency while preserving probabilistic properties. Feature vectors undergo kernel transformations to construct the density matrix ρ_{XY}, ensuring an optimal mapping of data points. Similarity measurements between new samples and learned prototypes enable probabilistic classification, enhancing model generalisation despite limited dataset sizes [30].

3.2.2. Baseline Experiments

To ensure robust model evaluation, twenty different *training-test* partitions of the dataset were created for the experimental setup. In each case, 70% of the data was used for training and 30% for testing. To maintain consistency across experiments, L2 normalization was applied to all features. In addition, to address class imbalance in the dataset, class weight compensation was systematically applied. This technique assigns higher weights to less frequent classes during model training, helping to mitigate bias [31](see Figure 3.1).

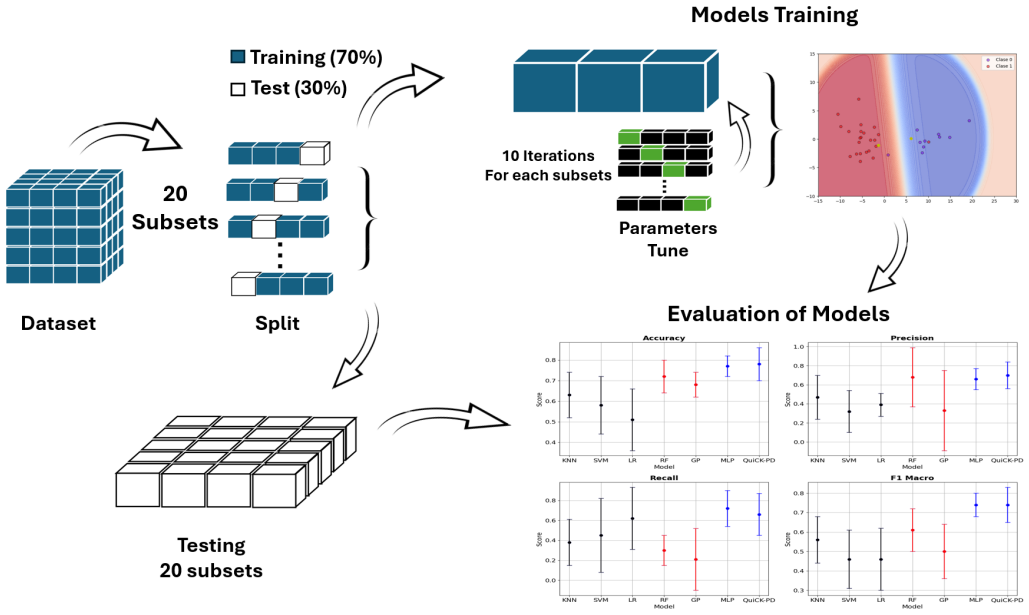

Figure 3.1.: Data processing with a 70-30 split and K-fold (10) cross-validation. Includes L2 normalization.

Hyperparameters Optimisation

For each data partition, hyperparameter optimization was performed using 5-fold cross-validation. This approach ensured robust tuning of the model parameters while minimizing the risk of overfitting, even with the limited dataset size. The hyperparameter tuning process followed a structured approach, and Table B.1 (Appendix B) summarises the key parameters explored for each method.

For K-Nearest Neighbours (KNN), different numbers of neighbours (K) were tested, ranging from 1 to 30. Distance metrics included Euclidean, Manhattan, and Minkowski, with weighting schemes set to uniform or distance-based.

For logistic regression, it explored key hyperparameters, including the regularization parameter C, across a logarithmic scale from 10^{-3} to 10^{3}, along with different penalty norms (L1, L2) and solvers (`lbfgs`, `liblinear`, and `sag`).

For Random Forest, the tuned hyperparameters included the number of estimators (100, 200), the maximum tree depth (None, 10), the minimum number of samples required to split an internal node, and the minimum number of samples required at a leaf node (minimum samples per split: 2, 5; minimum samples per leaf: 1, 2).

For the Gaussian Process Classifier, the posterior distribution of the latent function is non-Gaussian due to the discrete nature of the class labels, making a Gaussian likelihood inappropriate. Instead, a non-Gaussian likelihood based on the logistic link function (logit) it used. To approximate this non-Gaussian posterior, GPC uses the Laplace approximation. It initialized the kernel with a Gaussian (RBF) kernel and explored different values for the σ parameter (values for the length-scale parameter σ explored in $\{0.05, 0.1, 0.2\}$.).

Similarly, for the Support Vector Machine (SVM), both cosine and Gaussian (RBF) kernels were evaluated. For the RBF kernel, the kernel coefficient γ was explored logarithmically between 10^{-3} and 10^{1}, and the regularisation parameter C was tuned between 10^{-3} and 10^{3}.

For the Multilayer Perceptron (MLP), It experimented with the number of hidden layers (1 or 2), the number of neurons per hidden layer (32, 64, 128), and the activation functions (ReLU or tanh). Learning rates were tested between 10^{-4} and 10^{-1}. These hyperparameters were matched to those used in the encoder of CoGait-PD, allowing a direct performance comparison between the models.

CoGait-PD, requires the exploration of multiple hyperparameters. For the encoder, which is an MLP, the number of hidden layers (up to a maximum of 2) and the size of these layers were explored. The size of the layers determines the dimension of the embedding, which is fed into the Kernel Density Matrix (KDM). Within the

KDM, the following hyperparameters were optimized: the number of components, the kernel type, the σ parameter of the (Gaussian) kernel (which is learned during training), and the learning rate. The kernel type was selected from the Cosine and Gaussian options. Given the normalized data, the Gaussian kernel consistently performed well when σ was initialized to 0.1.

The number of components determines the size of the density matrix decomposition, as described in Section 3.2, and defines the maximum number of prototypes the model can learn. This value was explored in the range of 32, 64, and 128. These prototypes were initialized with randomly sampled instances from the training data set. The learning rate was explored logarithmically between 10^{-6} and 0.1. Adam [32] was used as the gradient descent optimization algorithm. Training was performed for a maximum of 200 epochs, using callbacks to allow early termination if no performance improvement was observed in the validation set after 20 epochs.

Hyperparameter tuning and training was done in Python. The kernel density matrix (KDM) and the CoGait-PD model were implemented using *TensorFlow* [33]. The code for the KDM implementation and the CoGait-PD framework is available at `https://github.com/fagonzalezo/kdm` and `https://github.com/santiagotoco/CoGait-PD`, respectively.

4. Results

Model predictions were made on the test set data. This was done by comparing the test set labels according to the result of each classifier with the test set label corresponding to one year of evolution of the patient's cognitive impairment, evaluating the overall performance of the models.

Performance is evaluated in terms of the agreement between predicted and actual labels using a confusion matrix. The confusion matrix quantifies how many of the positive samples (PD+CI) were detected positive by the model (true positives, TP), how many negative samples (PD) were indeed flagged as negative (true negatives, TN), how many samples that were predicted as positive were actually negative (PD) (false positives, FP), and how many samples that were predicted as negative were actually positive (false negatives, FN). From these results, we calculated key performance metrics such as accuracy, precision, recall (or sensitivity), and macro F1 score for each model in each test partition as follows:

$$\text{accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{4.1}$$

$$\text{precision} = \frac{TP}{TP + FP} \tag{4.2}$$

$$\text{recall} = \frac{TP}{TP + FN} \tag{4.3}$$

$$F1 = \frac{2 \cdot \text{recall} \cdot \text{precision}}{\text{recall} + \text{precision}} \tag{4.4}$$

It compared the performance of different machine learning methods such as K-Nearest Neighbor(KNN), Support Vector Machine (SVM), Logistic Regression (LR), Random Forest Classifier (RF), Gaussian Processes (GP), Multilayer perceptron (MLP) and Cognitive assessment through Gait in Parkinson's Disease (CoGait-PD) It report the average result and standard deviation of the 20 experiments obtained by classifying the test subsets (Table 4.1).

Model	**Accuracy**	**Precision**	**Recall**	**F1 Macro**
KNN	0,63 ± 0.11	0,47 ± 0,23	0,38± 0,23	0,56 ±0,12
SVM	0,58 ± 0.14	0,32 ± 0,22	0,45± 0,37	0,46 ±0,15
LR	0,51 ± 0,15	0,39 ± 0,12	0,62 ± 0,31	0,46 ± 0,16
RF	0,72 ± 0,08	0,68 ± 0,31	0,30 ± 0,15	0,61 ± 0,11
GP	0,68 ± 0,06	0,33 ± 0,42	0,21 ± 0,31	0,50 ± 0,14
MLP	0,77 ± 0,06	0,66 ± 0,12	**0,72** ± 0,19	**0,74** ± 0,06
CoGait-PD	**0,78** ± 0.08	**0,70** ± 0,15	0,64± 0,21	**0,74** ±0,09

Table 4.1.: Results of the different classifiers expressed as experimental mean and standard deviation of the 20 test subsets.

Analysis of the performance of different machine learning models for predicting cognitive decline in Parkinson's patients based on spatio-temporal gait characteristics showed significant variation in the ability of the models to identify hidden relationships within this data. The effectiveness of the models depended on their ability to uncover complex patterns that were not immediately apparent, which directly affected their performance in correctly classifying patients.

KNN showed limited performance, with a precision of 0.47 ± 0.23 and a recall of 0.38 ± 0.23. This poor performance reflected KNN's inability to detect hidden patterns that clearly distinguished patients with cognitive decline from those without. The high variability in results suggested that the model did not generalise well in this

dataset, suggesting that it missed key information in the internal relationships of the gait data, affecting both sensitivity and specificity. KNN failed to capture the complexity required to make consistent predictions.

The SVM performed poorly, with a precision of 0.32 ± 0.22 and a recall of 0.45 ± 0.37, suggesting that it did not adequately capture important relationships in the gait data. The low precision reflected a high rate of false positives, while the variability in recall suggested inconsistency in identifying cases of cognitive decline. The SVM struggled to generalise from hidden patterns in the data, making it ineffective in this particular context.

LR had a relatively high recall (0.62 ± 0.31), meaning that it detected a higher proportion of positive cases than other low performing models. However, its low precision (0.39 ± 0.12) and F1 score of 0.46 ± 0.16 suggested that the model misclassified many negative cases as positive. Although this linear model was able to detect some cases of cognitive decline, it failed to capture the non-linear interactions between quantitative gait variables that may be indicative of the patient's cognitive state. This resulted in a high rate of false positives and limited the usefulness of the model in clinical practice.

As previously shown, RF achieved a precision of 0.68 ± 0.31 for all gait characteristics but had a low recall (0.30 ± 0.15). While it could handle some complexity in the data, its low recall indicated that it failed to correctly identify a significant proportion of patients with cognitive decline. This suggests that RF missed some important hidden relationships in the data, leading to a high rate of false negatives. RF was able to detect some relationships in the data, but its inconsistency in detecting positive cases limited its applicability.

GP showed high precision (0.83 ± 0.24), suggesting that it effectively classified patients without cognitive decline. However, its low recall (0.21 ± 0.31) and F1 score of 0.50 ± 0.14 indicated that it missed many positive cases. GP did not appear to capture all the hidden relationships needed to correctly identify patients at risk of cognitive decline, resulting in a high rate of false negatives. Although it was effective in minimising false positives, it did not adequately detect the cases

that mattered most in this clinical context.

MLP demonstrated an F1 score of 0.74 ± 0.06 and an accuracy of 0.77 ± 0.06, reflecting its ability to correctly identify both positive and negative cases. MLP effectively captured hidden relationships in the gait data, allowing it to balance precision and recall. Its consistent performance suggested that it could uncover complex patterns in the data that other models missed, making it a reliable option for predicting cognitive decline.

CoGait-PD was the best performing model overall, with an F1 score of 0.74 ± 0.09 and an accuracy of 0.78 ± 0.08. This indicated that CoGait-PD was highly effective at detecting the hidden relationships between gait data and risk of cognitive decline. Its ability to detect both positive and negative cases with high accuracy and consistency showed that this model could capture the underlying complexity in the data. CoGait-PD excelled in balancing sensitivity and specificity, making it the most reliable model for clinical prediction of cognitive decline.

5. Interpretability and Conclusions

5.1. Feature Importance and Associations Between Clinical Biomarkers

A preliminary approach to determining the associations between gait characteristics and cognitive decline can be performed using the Random Forest (RF) classifier. RF facilitates the identification of the most relevant features by learning to partition the feature space based on information gain. During training, it evaluates how much each feature contributes to discriminating between classes, thereby providing insights into the relationship between specific gait characteristics and cognitive decline. This process yields valuable information regarding which gait features are most strongly associated with cognitive impairments. In this way, the 10 most relevant gait characteristics(RF10) were selected (See Figure 5.1) and grouped into four domains: consistency, rhythmicity, smoothness, and asymmetry, as documented during data collection as described below:

Consistency (Arm Movement)

- **LA_STD_DT**: Variability of the left arm in dual-task walking
- **ASYM_IND_DT**: Calculated arm asymmetry index in dual-task walking
- **ASA_DT**: Asymmetry in the magnitude of arm swing in dual-task walking

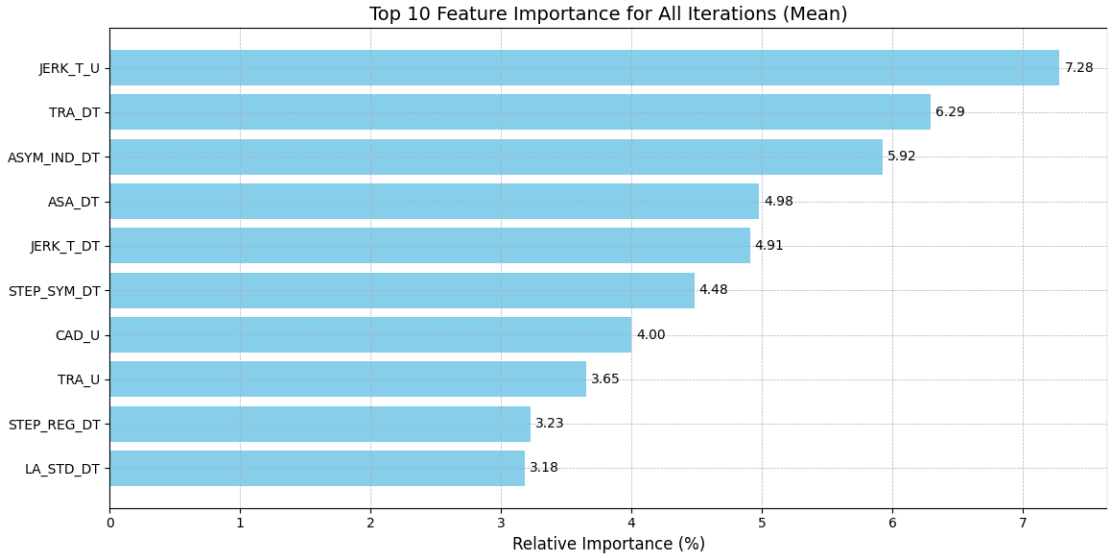

Figure 5.1.: The top 10 most relevant Characteristics for classification in the Random Forest classifier (RF10). The identification of these features underscores the intricate relationship between gait characteristics and cognitive impairment in Parkinson's disease patients.

Rhythmicity

- **STEP_SYM_DT**: Step symmetry in dual-task walking
- **CAD_U**: Cadence in usual walking
- **STEP_REG_DT**: Step regularity in dual-task walking

Smoothness (Jerk)

- **JERK_T_U**: Trunk smoothness (Jerk) in usual walking
- **JERK_T_DT**: Trunk smoothness (Jerk) in dual-task walking

Asymmetry (Trunk Rotation)

- **TRA_DT**: Trunk rotation asymmetry index in dual-task walking
- **TRA_U**: Trunk rotation asymmetry index in usual walking

Subsequently, the identified characteristics were compared with those reported by different authors, providing information on the association scales between gait characteristics and cognitive decline in different neurodegenerative disease.

Consistency

Li et al. in their study of gait for the recognition of Alzheimer's Disease (AD) and cognitive impairment (CI) reported a decrease in maximum arm speed in patients with AD and MCI, expressing that arm movements may be influenced by cognitive load possibly due to the shared cortical networks involved in gait and cognition control [34]. These reports are based on studies reported by Krasovsky et al. [35] and Cohen & Verghese [36].

Baron et al. show a significant decrease in arm swing under different dual task gait conditions (DT), suggesting that this measure is affected when cognitive resources are sufficiently overloaded[37], supporting our hypothesis of spatiotemporal assessments as predictive biomarkers of cognitive impairment in patients with Parkinson's disease.

Rhythmicity

Poole et al. determined in a study of 438 older adults that gait speed is associated with temporal declines in global cognition, executive functioning, and memory abilities. They also found that several gait-related brain structures, such as regional brain volumes in the ventricles and temporal lobe, were associated with declines in cognition.[38]
Hoang et al. concluded in a study evaluating cognitive workload in young adults during walking conditions varying in complexity that the 24 participants showed greater gait variability during both simple and complex dual task walking conditions compared to habitual walking. This suggests that dual-tasking, regardless of its nature, influences gait stability. As the complexity of the walking task increased, the speed, cadence, and length of stride decreased in the population, indicating that

individuals needed to recruit a greater amount of cognitive resources to perform cognitive and walking tasks simultaneously.[39]

Smoothness

The connection between JERK and cognitive decline is evident in several recent studies, in which increases in JERK have been associated with axial rigidity in the trunk not detected by conventional methods[40, 41, 42]. This is particularly relevant in conditions such as Parkinson's disease and in frail older adults, where variability in trunk acceleration and reduced stride regularity correlate with neural control difficulties and postural stability.

Duncan et al. describe how unilateral and violent muscle jerks can indicate rapid cognitive decline, reinforcing the idea that visible motor alterations can be an early reflection of underlying neurological problems.[42]

Furthermore, Mirelman et al. indicate that changes in motor control, such as jerks during simple transitions, can be sensitive predictors of cognitive decline, providing a useful tool for the early detection of conditions such as dementia or mild cognitive impairment.[43]

Pantall et al. demonstrate that postural dynamics, including jerk, decline with disease progression in PD, indicating that changes in postural control reflect cognitive decline. This aligns with the idea that higher jerk levels during quiet standing or dual tasks can indicate impaired neural control and postural coordination.[44]

Chen et al.provide additional support by showing that patients with early PD exhibit higher jerk levels in postural sway, especially during cognitive tasks. This suggests that increased jerk is not only a marker of physical instability but also correlates with cognitive and motor impairment.[45]

Our findings that are consistent with mencionated studies underscore the importance of monitoring the jerk trunk as a potential early indicator of cognitive decline, offering a useful tool for early detection and management of conditions such as PD.

Asymmetry

Lamoth et al. in a study of 26 elderly individuals, 13 of whom had dementia, found differences in movement execution during gait in dual-task conditions. The study highlights that dual-tasking reduces trunk stability, noting that elderly individuals with cognitive impairment exhibited less correlated trunk acceleration patterns. These measures, provided by the acceleration in the anterior-posterior and mediolateral directions, suggest greater variability and reduced stability compared to the cognitively intact group [46].

Doi et al. investigated the relationship between trunk stability and cognitive decline in a cohort of 110 older adults, aged 65 to 94 years, during dual-task walking—a type of walking that imposes higher cognitive demands. The authors found that trunk stability significantly decreases under dual-task walking conditions, and this reduction is independently associated with brain atrophy. Specifically, they observed that changes in the harmonic ratio of trunk movement in the vertical direction, a measure of stability, correlate with brain atrophy, suggesting that trunk instability may serve as a biomechanical marker of cognitive decline in this population.[47]

5.2. Variability in classifier performance and Analysis of prototypes

A comparative analysis of the different models for predicting cognitive decline in Parkinson's disease (PD) reveals differences in performance and consistency. This analysis takes into account changes in gait and their correlation with future cognitive decline, as well as the ability to characterise specific relationships with reproducibility in small observation samples. Figure 5.2 shows the visible differences that allow the classifiers to be divided into three groups:

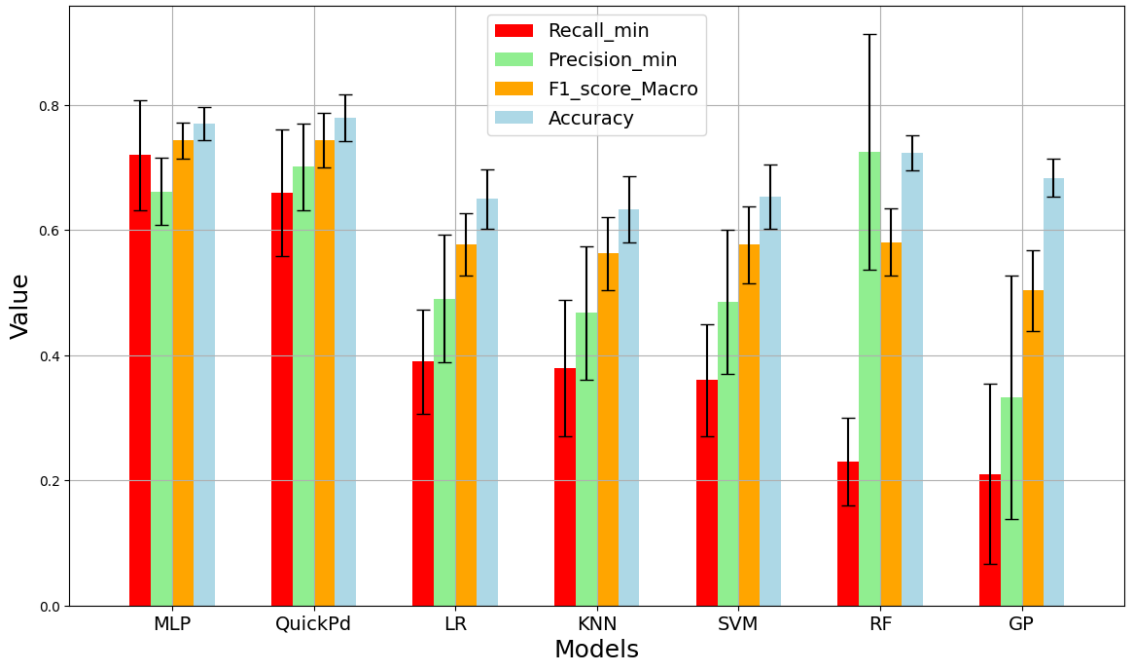

Figure 5.2.: Comparison of the performance of different classification models in predicting cognitive decline in Parkinson's patients, focusing on their confidence intervals for key metrics. The models were evaluated based on accuracy, precision, recall and F1-macro, with each bar representing the mean and the error bars representing the 95% confidence intervals. CoGait-PD and MLP have the narrowest confidence intervals for all metrics, especially for accuracy and F1-macro, reflecting the highest level of consistency and reliability in their predictions.

Low-Performance Models

Models such as RF and GP had narrower confidence intervals, but still showed variability in certain metrics. RF stood out in terms of precision, but its confidence interval remained significantly wider compared to the high performance models, indicating that while it could provide accurate predictions in some cases, it lacked consistency across all scenarios. GP, on the other hand, showed considerable variability in its recall metrics, reinforcing the idea that although it outperformed low-performing models, it still struggled to achieve sufficient stability in its predictions.

Moderate-performing Models

The models such as KNN, SVM and LR showed high variability in their confidence intervals, indicating reduced stability in their predictions. In particular, KNN

showed a wide range of precision and recall, reflecting significant inconsistency in its ability to predict cognitive decline. This reinforced previous observations that these models struggled to capture the complex relationships between gait changes and cognitive decline. The wide variability in their confidence intervals suggested that their predictions could vary significantly, limiting their applicability in clinical settings where consistent predictions are required.

High-Performance Models

Finally, models such as MLP and CoGait-PD had the narrowest and most consistent confidence intervals. This suggests that their predictions were more reliable and consistent across all metrics evaluated. MLP showed particularly strong performance with reduced confidence intervals for precision, recall and F1-macro, further validating its suitability for clinical applications where stable predictions are required. Similarly, CoGait-PD not only demonstrated narrow confidence intervals for all metrics, but also low variability even for the more challenging metrics such as recall and precision. This demonstrates that CoGait-PD was able to maintain an unusually high level of consistency, a critical factor in clinical contexts, unlike other models which showed significant variability. This further underlines the model's ability to capture the underlying relationships between gait changes and cognitive decline with precision and stability, making it the most reliable and robust option for long-term prediction in medical settings.

Representations of space during training of high-performance models: a novel clinical interpretability

Additional insights emerged from the analysis of latent space representations, providing a more integrated and nuanced understanding of the relationship between gait patterns and cognitive decline. Among the models evaluated, only the MLP and CoGait-PD learn intermediate representations that allow them tomap the original feature space into a latent space that better captures class distinctions. Visualization plays a critical role in interpreting these learned spaces and assessing the ability of these models to separate classes.

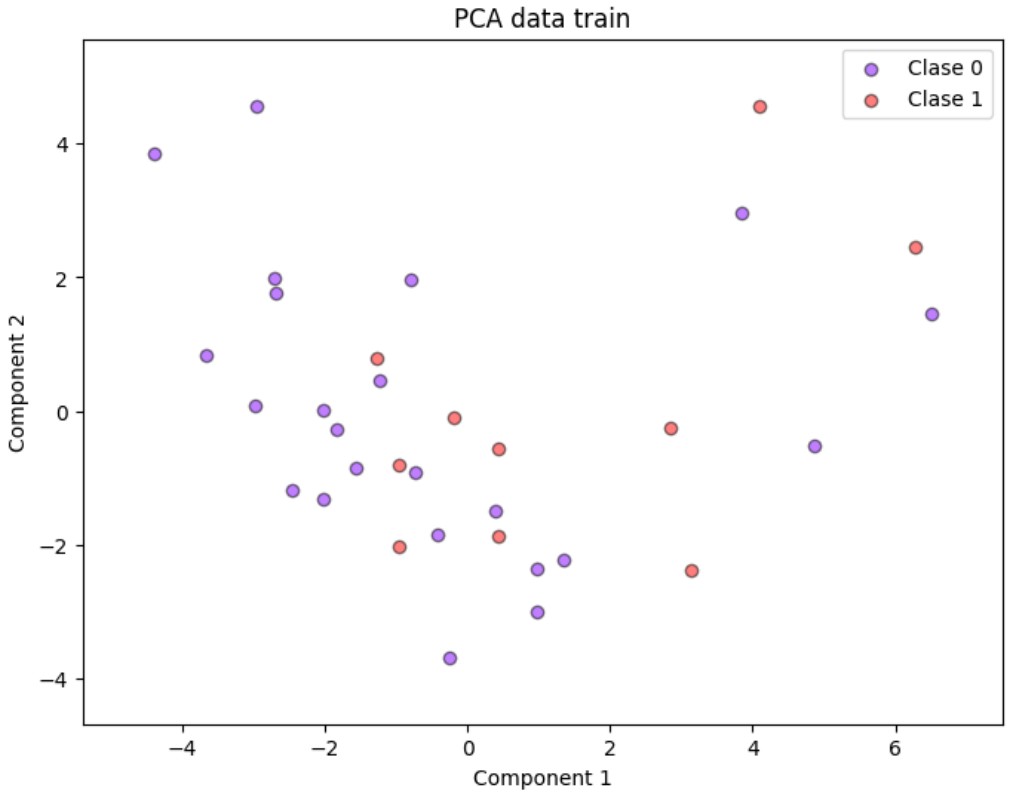

Figure 5.3.: Analysis of the latent space representation using principal component analysis (PCA) of spatiotemporal gait characteristics in the training data.

A principal component analysis (PCA) visualisation of the representation learned from a subset of spatiotemporal gait characteristics in the training data (Figure 5.3) facilitates the assessment of differences between models.

In this case, the original representation space reveals the absence of a simple or linear decision boundary between classes, making the detection of complex patterns less apparent.

A comparison with the MLP representation, which employs dense layers and non-linear activation functions, shows a smoother and more gradual class separation (Figure 5.4). This demonstrates the model's ability to capture subtle variations in gait data, enabling it to detect a greater number of positive cases through more inclusive approaches.

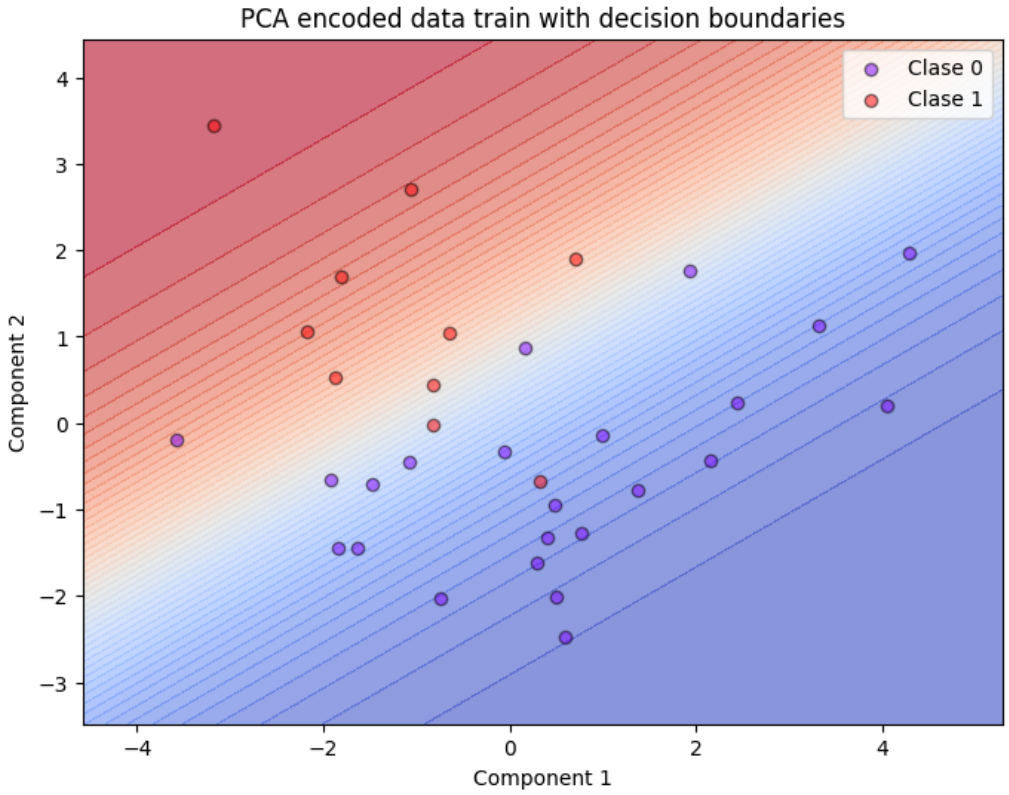

Figure 5.4.: MLP representation, which learns feature distributions using dense layers and non-linear activation functions, exhibits a smoother and more gradual class separation.

This separation indicates that MLP is the most appropriate choice when the primary clinical objective is to ensure the detection of all cases of cognitive impairment while minimising false negatives (i.e., prioritising sensitivity). This model is particularly useful in large-scale screenings, where a significant number of patients must be evaluated and critical cases cannot be overlooked.

In contrast, the CoGait-PD representation, with its prototyping approach (Figure 5.5), establishes sharper classification boundaries and enhances interpretability, allowing for a deeper understanding of how data is classified. This model benefits from prototype consolidation and learning, with prototypes illustrated as yellow dots in Figure 5.5.

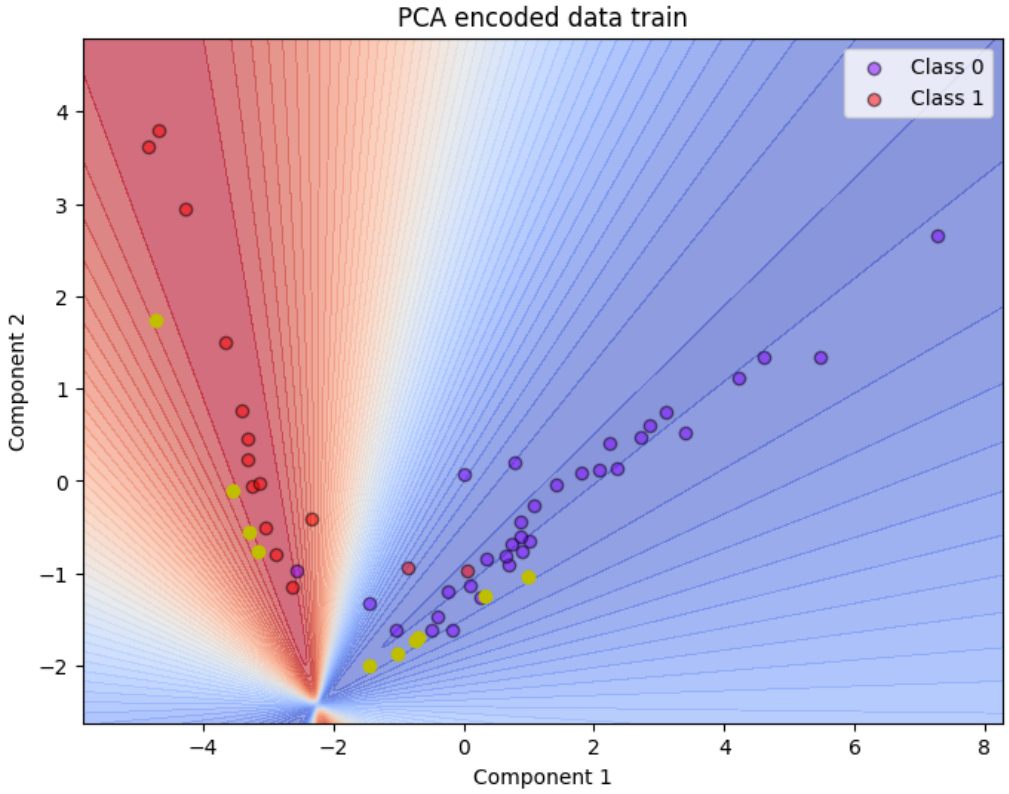

Figure 5.5.: PCA visualization of the feature space learned by CoGait-PD. The background colors define the decision region of the model. CoGait-PD identifies patient prototypes (shown as yellow dots) for each class, along with a kernel similarity measure. These components are used for inference on new samples and are critical for clinical interpretability, enabling clinicians to understand and trace patient-level predictions.

These prototypes, referred to as 'key patients', serve as reference points in the latent space, significantly improving the interpretability of the model by enabling clinicians to track the progression of cognitive impairment.

New patient samples can be compared against these prototypes, offering unique insights into disease progression. Unlike distance-based methods, which rely on explicit neighbourhood relationships for classification, this approach leverages learned

probabilistic representations that encode disease progression patterns. This consolidation enables a clearer interpretation of the key features that distinguish cases of cognitive impairment.

This functionality provides a powerful visualization and understanding a patient'sdisease evolution, with significant potential for personalized healthcare interventions, while it builds confidence in the model's predictions, providing clinicians with the necessary evidence to support any decision.

5.3. Conclusions

The relationship between Parkinson's disease (PD) and cognitive impairment represents a significant medical and social challenge, reflected in its impact on the quality of life of those affected. In this context, this research offers an innovative perspective by exploring the potential of gait patterns as non-invasive clinical biomarkers for predicting cognitive impairment in patients with PD.

Through the instrumental analysis of gait data and the application of advanced machine learning techniques, the findings show remarkable performance in predicting cognitive impairment up to a year before its onset. In particular, the deep learning methods leveraging high-performance models, such as the Cognitive assessment through Gait in Parkinson's Disease (CoGait-PD), demonstrated significant predictive power, outperforming traditional models such as random forests and support vector machines. These advances enable the generation of clinically interpretable and relevant results for daily clinical practice, highlighting the value of modern technological tools in the field of personalised medicine.

The use of 20 random partitions in the experimental design provided a robust framework to evaluate the model's performance, yielding reliable metrics despite the limited dataset. This methodological approach emphasises the importance of using non-parametric models that do not rely on assumptions about data distribution, facilitating the identification of hidden relationships between gait characteristics and cognitive impairment. Moreover, the model's robust performance with small cohorts highlights its potential for use in clinical contexts where large datasets are limited.

However, it is essential to acknowledge the limitations of this study, such as the small sample size and the need for further validation in larger and more diverse cohorts. This study presents an underexplored avenue of research, and its clinical implementation will require additional deterministic studies, as well as careful consideration of the ethical and practical aspects that this approach entails.

In conclusion, this study paves the way for more personalised and effective interventions in the management of PD. By using gait patterns as a predictive biomarker,

clinicians will be able to more accurately monitor the cognitive status of people with PD, allowing for early intervention in the disease's progression. These results emphasise the need for continued research in this field to validate the findings in broader clinical settings, ensuring that the proposed approach can be applied reliably and ethically in practice.

Bibliography

[1] W. H. Organization, *Parkinson disease: a public health approach. Technical Brief*. World Health Organization, 2022.

[2] N. Maserejian, L. Vinikoor-Imler, and A. Dilley, "Estimation of the 2020 global population of parkinson's disease (pd)," *MDS Abstracts*, 2020. Accessed: 2024-08-28.

[3] Lancet, "What next in parkinson's disease?," *The Lancet*, vol. 403, no. 10370, p. 219, 2024. Accessed: 2024-08-28.

[4] D. Aarsland, B. Creese, M. Politis, K. R. Chaudhuri, D. Weintraub, C. Ballard, *et al.*, "Cognitive decline in parkinson disease," *Nature Reviews Neurology*, vol. 13, no. 4, pp. 217–231, 2017.

[5] D. Aarsland, K. Bronnick, C. Williams-Gray, D. Weintraub, K. Marder, J. Kulisevsky, D. Burn, P. Barone, J. Pagonabarraga, L. Allcock, *et al.*, "Mild cognitive impairment in parkinson disease: a multicenter pooled analysis," *Neurology*, vol. 75, no. 12, pp. 1062–1069, 2010.

[6] C. C. Janvin, J. P. Larsen, D. Aarsland, and K. Hugdahl, "Subtypes of mild cognitive impairment in parkinson's disease: progression to dementia," *Movement disorders: official journal of the Movement Disorder Society*, vol. 21, no. 9, pp. 1343–1349, 2006.

[7] C. Williams-Gray, T. Foltynie, C. Brayne, T. Robbins, and R. Barker, "Evolution of cognitive dysfunction in an incident parkinson's disease cohort," *Brain*, vol. 130, no. 7, pp. 1787–1798, 2007.

[8] C. H. Williams-Gray, J. R. Evans, A. Goris, T. Foltynie, M. Ban, T. W. Robbins, C. Brayne, B. S. Kolachana, D. R. Weinberger, S. J. Sawcer, *et al.*, "The distinct cognitive syndromes of parkinson's disease: 5 year follow-up of the campaign cohort," *Brain*, vol. 132, no. 11, pp. 2958–2969, 2009.

[9] R. Morris, S. Lord, R. A. Lawson, S. Coleman, B. Galna, G. W. Duncan, T. K. Khoo, A. J. Yarnall, D. J. Burn, and L. Rochester, "Gait rather than cognition predicts decline in specific cognitive domains in early parkinson's disease," *Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences*, vol. 72, no. 12, pp. 1656–1662, 2017.

[10] J. Verghese, C. Wang, R. B. Lipton, R. Holtzer, and X. Xue, "Quantitative gait dysfunction and risk of cognitive decline and dementia," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 78, no. 9, pp. 929–935, 2007.

[11] M. M. Mielke, R. O. Roberts, R. Savica, R. Cha, D. I. Drubach, T. Christianson, V. S. Pankratz, Y. E. Geda, M. M. Machulda, R. J. Ivnik, *et al.*, "Assessing the temporal relationship between cognition and gait: slow gait predicts cognitive decline in the mayo clinic study of aging," *Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences*, vol. 68, no. 8, pp. 929–937, 2013.

[12] S. Lord, B. Galna, S. Coleman, A. Yarnall, D. Burn, and L. Rochester, "Cognition and gait show a selective pattern of association dominated by phenotype in incident parkinson's disease," *Frontiers in aging neuroscience*, vol. 6, p. 249, 2014.

[13] A.-M. De Cock, E. Fransen, S. Perkisas, V. Verhoeven, O. Beauchet, M. Vandewoude, and R. Remmen, "Comprehensive quantitative spatiotemporal gait analysis identifies gait characteristics for early dementia subtyping in community dwelling older adults," *Frontiers in neurology*, vol. 10, p. 313, 2019.

[14] R. Morris, S. Lord, J. Bunce, D. Burn, and L. Rochester, "Gait and cognition: mapping the global and discrete relationships in ageing and neurodegenerative disease," *Neuroscience & Biobehavioral Reviews*, vol. 64, pp. 326–345, 2016.

[15] C. Baiano, P. Barone, L. Trojano, and G. Santangelo, "Prevalence and clinical aspects of mild cognitive impairment in parkinson's disease: A meta-analysis," *Movement Disorders*, vol. 35, no. 1, pp. 45–54, 2020.

[16] A. Moncada-Torres, M. C. van Maaren, M. P. Hendriks, S. Siesling, and G. Geleijnse, "Explainable machine learning can outperform cox regression predictions and provide insights in breast cancer survival," *Scientific reports*, vol. 11, no. 1, p. 6968, 2021.

[17] G. Álvarez Hernández, "Limitaciones metodológicas de la epidemiología moderna y una alternativa para superarlas: la epidemiología sociocultural," *Región y sociedad*, vol. 20, no. SPE2, pp. 51–75, 2008.

[18] I. Martínez-Baz, I. Delfrade, and J. Etxeberria, "Characteristics and statistical methods used in original articles published in public health journals in spain," in *Anales Del Sistema Sanitario de Navarra*, vol. 41, pp. 347–354, 2018.

[19] D. A. Morales, Y. Vives-Gilabert, B. Gómez-Ansón, E. Bengoetxea, P. Larrañaga, C. Bielza, J. Pagonabarraga, J. Kulisevsky, I. Corcuera-Solano, and M. Delfino, "Predicting dementia development in parkinson's disease using bayesian network classifiers," *Psychiatry Research: NeuroImaging*, vol. 213, no. 2, pp. 92–98, 2013.

[20] S. N. Gomperts, J. J. Locascio, D. Rentz, A. Santarlasci, M. Marquie, K. A. Johnson, and J. H. Growdon, "Amyloid is linked to cognitive decline in patients with parkinson disease without dementia," *Neurology*, vol. 80, no. 1, pp. 85–91, 2013.

[21] P.-H. Chen, T.-Y. Hou, F.-Y. Cheng, and J.-S. Shaw, "Prediction of cognitive degeneration in parkinson's disease patients using a machine learning method," *Brain Sciences*, vol. 12, no. 8, p. 1048, 2022.

[22] F. Pieruccini-Faria, S. E. Black, M. Masellis, E. E. Smith, Q. J. Almeida, K. Z. Li, L. Bherer, R. Camicioli, and M. Montero-Odasso, "Gait variability

across neurodegenerative and cognitive disorders: Results from the canadian consortium of neurodegeneration in aging (ccna) and the gait and brain study," *Alzheimer's & Dementia*, vol. 17, no. 8, pp. 1317–1328, 2021.

[23] P.-H. Chen, C.-W. Lien, W.-C. Wu, L.-S. Lee, and J.-S. Shaw, "Gait-based machine learning for classifying patients with different types of mild cognitive impairment," *Journal of medical systems*, vol. 44, pp. 1–6, 2020.

[24] D. Trabassi, M. Serrao, T. Varrecchia, A. Ranavolo, G. Coppola, R. De Icco, C. Tassorelli, and S. F. Castiglia, "Machine learning approach to support the detection of parkinson's disease in imu-based gait analysis," *Sensors*, vol. 22, no. 10, p. 3700, 2022.

[25] A. Shahzad, A. Dadlani, H. Lee, and K. Kim, "Automated prescreening of mild cognitive impairment using shank-mounted inertial sensors based gait biomarkers," *IEEE Access*, vol. 10, pp. 15835–15844, 2022.

[26] S. Toledo-Cortés, D. H. Useche, H. Müller, and F. A. González, "Grading diabetic retinopathy and prostate cancer diagnostic images with deep quantum ordinal regression," *Computers in biology and medicine*, vol. 145, p. 105472, 2022.

[27] I. H. Sarker, "Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions," *SN Computer Science*, vol. 2, no. 6, p. 420, 2021.

[28] K. Marek, D. Jennings, S. Lasch, A. Siderowf, C. Tanner, T. Simuni, C. Coffey, K. Kieburtz, E. Flagg, S. Chowdhury, *et al.*, "The parkinson progression marker initiative (ppmi)," *Progress in neurobiology*, vol. 95, no. 4, pp. 629–635, 2011.

[29] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn:

Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[30] S. Toledo-Cortés, J. E. Serna, F. A. González, and E. Romero, "Predicting cognitive decline in parkinson's disease by gait analysis: An interpretable machine learning perspective," *Computers in Biology and Medicine*, February 2025. Preprint, under review.

[31] S. Lerman and C. Manski, "On the use of simulated frequencies to approximate choice probabilities," *Structural analysis of discrete data with econometric applications*, vol. 10, pp. 305–319, 1981.

[32] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2017.

[33] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, "TensorFlow: Large-scale machine learning on heterogeneous systems," 2015. Software available from tensorflow.org.

[34] Z. Li, J. Zhu, J. Liu, M. Shi, P. Liu, J. Guo, Z. Hu, S. Liu, and D. Yang, "Using dual-task gait to recognize alzheimer's disease and mild cognitive impairment: a cross-sectional study," *Frontiers in Human Neuroscience*, vol. 17, 2023.

[35] T. Krasovsky, A. Lamontagne, A. G. Feldman, and M. F. Levin, "Effects of walking speed on gait stability and interlimb coordination in younger and older adults," *Gait & posture*, vol. 39, no. 1, pp. 378–385, 2014.

[36] J. A. Cohen and J. Verghese, "Gait and dementia," *Handbook of clinical neurology*, vol. 167, pp. 419–427, 2019.

[37] E. I. Baron, M. M. Koop, M. C. Streicher, A. B. Rosenfeldt, and J. L. Alberts, "Altered kinematics of arm swing in parkinson's disease patients indicates declines in gait under dual-task conditions," *Parkinsonism & related disorders*, vol. 48, pp. 61–67, 2018.

[38] V. N. Poole, S. Oveisgharan, L. Yu, R. J. Dawe, S. E. Leurgans, S. Zhang, K. Arfanakis, A. S. Buchman, and D. A. Bennett, "Volumetric brain correlates of gait associated with cognitive decline in community-dwelling older adults," *Frontiers in Aging Neuroscience*, vol. 15, 2023.

[39] I. Hoang, M. Ranchet, R. Derollepot, F. Moreau, and L. Paire-Ficout, "Measuring the cognitive workload during dual-task walking in young adults: a combination of neurophysiological and subjective measures," *Frontiers in human neuroscience*, vol. 14, p. 592532, 2020.

[40] W. Wright, V. Gurfinkel, J. Nutt, F. Horak, and P. Cordo, "Axial hypertonicity in parkinson's disease: direct measurements of trunk and hip torque," *Experimental neurology*, vol. 208, no. 1, pp. 38–46, 2007.

[41] M. Mancini, F. B. Horak, C. Zampieri, P. Carlson-Kuhta, J. G. Nutt, and L. Chiari, "Trunk accelerometry reveals postural instability in untreated parkinson's disease," *Parkinsonism & related disorders*, vol. 17, no. 7, pp. 557–562, 2011.

[42] M. Duncan, J. Cholfin, and L. Restrepo, "Clinical reasoning: a 72-year-old man with rapid cognitive decline and unilateral muscle jerks," *Neurology*, vol. 82, no. 22, pp. e194–e197, 2014.

[43] A. Mirelman, A. Weiss, A. S. Buchman, D. A. Bennett, N. Giladi, and J. M. Hausdorff, "Association between performance on timed up and go subtasks and mild cognitive impairment: further insights into the links between cognitive and motor function," *Journal of the American Geriatrics Society*, vol. 62, no. 4, pp. 673–678, 2014.

[44] A. Pantall, P. Suresparan, L. Kapa, R. Morris, A. Yarnall, S. Del Din, and L. Rochester, "Postural dynamics are associated with cognitive decline in parkinson's disease," *Frontiers in neurology*, vol. 9, p. 1044, 2018.

[45] T. Chen, Y. Fan, X. Zhuang, D. Feng, Y. Chen, P. Chan, and Y. Du, "Postural sway in patients with early parkinson's disease performing cognitive tasks while standing," *Neurological research*, vol. 40, no. 6, pp. 491–498, 2018.

[46] C. J. Lamoth, F. J. van Deudekom, J. P. van Campen, B. A. Appels, O. J. de Vries, and M. Pijnappels, "Gait stability and variability measures show effects of impaired cognition and dual tasking in frail people," *Journal of neuroengineering and rehabilitation*, vol. 8, pp. 1–9, 2011.

[47] T. Doi, H. Makizako, H. Shimada, D. Yoshida, K. Ito, T. Kato, H. Ando, and T. Suzuki, "Brain atrophy and trunk stability during dual-task walking among older adults," *Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences*, vol. 67, no. 7, pp. 790–795, 2012.

A. Appendix A: Spatiotemporal Gait Characteristics

Table A.1.: Spatiotemporal gait characteristics

Assessments	**Description**
SP U, SP DT	Speed base walking and Dual task walking (m/sec)
RA AMP U, RA AMP DT	Right arm amplitude base and Dual task walking (deg)
LA AMP U, LA AMP DT	Left arm amplitude base and Dual task walking (deg)
RA STD U, RA STD DT	Right arm variability (std deviation) in base and Dual task (%)
LA STD U, LA STD DT	Left arm standard deviation in base and Dual task (%)
SYM U, SYM DT	Between arm symmetry in base and Dual task walking (%)

B. Appendix B: Summary of Hyperparameters Explored

Table B.1.: Summary of hyperparameters explored

Model	**Hyperparameters**	**Explored Values**
K-NN	Number of neighbours (K)	1 – 30
	Distance metric	Euclidean, Manhattan, Minkowski
	Weighting scheme	Uniform, Distance-based
SVM	Kernel type	Cosine, Gaussian (RBF)
	C (Regularisation)	10^{-3} to 10^{3}
	γ (Kernel coefficient)	10^{-3} to 10^{1}