

Self-Supervised Temporal Representation Learning for Parkinson's Disease Detection Using 3D Gait Data

Abdul Khaliq¹, Muhammad Asghar Khan², Muhammad Abbas³, Sana Alam^{4*}, Qurrat Ul Ain Naiyar⁵, Najmul Islam⁶

Abstract: Parkinson's disease (PD) is a progressive neurodegenerative disorder characterized by motor impairments, including tremor, rigidity, and postural instability. Early detection and continuous assessment of PD remain challenging, particularly when clinical datasets contain limited labeled samples. Recent advances in machine learning have enabled automated analysis of biomechanical data; however, most existing approaches rely on supervised learning and substantial annotated datasets. This study proposes a self-supervised temporal representation learning framework for PD detection using three-dimensional (3D) gait data. The framework employs a bidirectional Long Short-Term Memory (BiLSTM) network with a masked sequence reconstruction task to learn temporal motion representations from gait sequences. The pretrained encoder is subsequently fine-tuned to classify gait patterns associated with PD. The proposed framework was evaluated on the CARE-PD dataset containing 3D movement data from individuals with different clinical conditions. The model achieved an accuracy of 78.8%, a receiver operating characteristic area under the curve (ROC-AUC) of 0.81, and a reported recall of 0.90 for participants with PD. The proposed approach achieved the highest accuracy among the evaluated LSTM, BiLSTM, CNN-BiLSTM, and Transformer baseline models. The main contribution of this study is the use of masked sequence reconstruction to pretrain a temporal gait representation before supervised PD classification, thereby reducing the dependence on disease-specific labels. The findings provide initial evidence that self-supervised temporal representation learning can be useful for gait-based PD detection under limited labeled-data conditions.

Keywords: Parkinson's disease, self-supervised learning, 3D gait analysis, temporal representation learning, BiLSTM, masked sequence reconstruction, deep learning, gait-based detection.

I. INTRODUCTION

Parkinson's disease (PD) is a chronic progressive neurodegenerative disorder that primarily affects movement and is associated with symptoms including tremor, rigidity, bradykinesia, postural instability, and gait impairment [1], [2]. Recent global epidemiological evidence has documented the continuing burden and changing trends of PD worldwide [3]. The progressive nature of PD and the variability of its clinical manifestations make early identification and continuous assessment challenging. Gait abnormalities are particularly important because changes in walking pattern can reflect motor dysfunction and may provide objective information that complements conventional clinical assessment.

Traditionally, clinical assessments of Parkinson's disease have relied on standardized rating scales such as the Unified Parkinson's Disease Rating Scale (UPDRS), Movement Disorder Society Unified Parkinson's Disease Rating Scale (MDS-UPDRS), and Hoehn-Yahr staging. Although these measures are widely used, clinical assessment can involve subjectivity and inter-rater variability. Consequently, there has been increasing interest in objective gait-analysis technologies capable of quantifying movement characteristics.

Traditional gait-analysis systems include optical motion capture systems such as VICON, wearable inertial measurement units (IMUs), XSENS systems, and plantar-pressure systems [4], [5], [6]–[10]. These systems can provide detailed measurements of gait characteristics but may require specialized equipment, controlled environments, sensor placement, or participant compliance. Recent advances in computer vision and human pose estimation have therefore encouraged the development of markerless approaches for extracting gait and movement characteristics from video data [11], [12].

Li et al. [13] demonstrated that motion patterns associated with Parkinson's disease and levodopa-induced dyskinesia could be identified using vision-based pose estimation. Zeng et al. [14] subsequently investigated gait-score prediction using skeleton-contour fusion and dual-view imaging. Kondo et al. [15] demonstrated the potential of multi-view 3D pose estimation and target tracking for freezing-of-gait detection, while Guo et al. [16] investigated graph convolutional networks for five-category freezing-of-gait assessment.

Deep learning architectures have subsequently developed from convolutional and recurrent models toward graph-based and Transformer-based approaches. STC-Former and GLA-GCN demonstrated the ability to model spatiotemporal and skeletal relationships for 3D human pose estimation [17], [18]. Hybrid architectures such as MotionAGFormer combine Transformer and graph-convolutional representations for monocular 3D human pose estimation [19], while Motion2Vec combines Transformer-GCN representations with self-distillation and motion-aware pretraining [20]. These approaches have produced strong results on established 3D human pose benchmarks, although their direct application to clinical Parkinson's disease populations remains less established.

More recent research has further demonstrated the growing importance of computer vision and artificial intelligence in Parkinson's disease assessment. A systematic review by Franco et al. [12] identified 25 studies published between 2018 and 2023 and reported that deep learning approaches using wearable sensors and human pose estimation were increasingly being

^{1,2,3,4,5,6} Department of Computer Science, College of Computer Science and Information Systems, Institute of Business Management, Karachi
Country: Pakistan
Email: *sana.alam@iobm.edu.pk

applied to PD gait analysis. A 2025 systematic review of computer-vision approaches identified 45 eligible studies and reported gait as the most frequently investigated motor task, while also emphasizing the need for standardized acquisition procedures and external validation for clinical translation [21]. A 2025 review of machine-learning and deep-learning approaches similarly highlighted gait analysis as an important source of information for automated PD detection and monitoring [22].

Despite the substantial progress in 3D gait analysis and deep learning-based approaches for Parkinson's disease, an important challenge remains the dependence of many existing methods on labeled training data. Clinical gait datasets are comparatively difficult and expensive to annotate because diagnostic and clinical assessments require expert knowledge. This limitation can be particularly important when developing deep learning models, which generally require substantial quantities of labeled data.

Self-supervised learning provides an alternative strategy because a model can first learn representations from the structure of the input data itself and subsequently adapt those representations to a downstream supervised task. Recent biomechanical research has demonstrated that masked reconstruction-based self-supervised learning can improve data efficiency when only limited labeled gait-related data are available [23]. Similarly, recent work on Parkinsonian gait has investigated self-supervised representation learning to reduce dependence on conventional supervised learning and large labeled datasets [24].

Motivated by these limitations, this study proposes a self-supervised temporal representation learning framework for Parkinson's disease detection using 3D gait data. The framework employs a BiLSTM encoder that is pretrained using a masked sequence reconstruction task. The pretrained encoder is then fine-tuned using labeled gait data for PD classification.

The significance of this study lies in addressing the labeled-data limitation from a temporal representation-learning perspective rather than relying solely on architectural complexity. The proposed framework first learns general temporal characteristics of gait sequences and subsequently adapts these representations to the PD detection task. This provides a direct connection between self-supervised representation learning and clinically relevant gait classification.

The main contributions of this study are as follows:

1. A self-supervised temporal representation learning framework is developed for PD detection using 3D gait sequences.
2. A masked sequence reconstruction task is used to pretrain a BiLSTM encoder without requiring disease-specific labels during the representation-learning phase.
3. The pretrained temporal representation is subsequently fine-tuned using labeled gait data for PD classification.
4. The proposed approach is compared with LSTM, BiLSTM, CNN-BiLSTM, and Transformer baseline models.
5. The framework is evaluated under a limited-labeled-data setting using the CARE-PD dataset.

II. LITERATURE REVIEW

A. Clinical Assessment of Parkinson's Disease

Traditionally, clinical assessments of Parkinson's disease (PD) were done using standardized rating scales. Two of the best-known tools, the Unified Parkinson's Disease Rating Scale (UPDRS) and the newer MDS-UPDRS, are also the most well-validated methods to assess both motor and non-motor symptoms of all stages of PD. Additional tools, such as the Hoehn-Yahr staging and the Berg Balance Scale, can also be used to look at fall risk and functional ability. Even though these scales are commonly used, all of them have one major limitation: they are based on the opinion of a neurologist who has been trained to do so, and there is often a lot of error between different people who rate PD using these scales. Therefore, there is increasing interest in developing objective methods to assess walking using technology.

B. Wearable Sensor-Based Gait Analysis

Inertial measurement units (IMUs), which are worn on the body, have been researched significantly in terms of their use as objective measures of gait in individuals with PD. IMUs provide numerous advantages, such as portability, continuous monitoring, and relatively low cost compared with other forms of gait-analysis equipment. IMUs can reliably extract common gait parameters such as stride length, cadence, double support time, and indices of asymmetry from IMU signals in both laboratory settings and participants' free lives outside of an experimental setting [4], [5].

There is also significant work utilizing IMU systems for freezing-of-gait episode identification, including deep learning approaches combining CNN, BiLSTM, and attention mechanisms. While there are these advantages to using IMUs, there are also limitations to their use in clinical situations. Issues such as signal drift over long periods, especially in gyroscope-based orientation and position estimates, can result in cumulative error in long-duration recordings. The accuracy of estimated joint angles can be lower than optical motion capture, and the need to wear sensors places a burden on the individual and may create compliance issues. These limitations have encouraged interest in vision-based and markerless alternatives [11], [25]

C. Vision-Based Pose Estimation for Parkinson's Disease

As a result of advancements in convolutional neural networks (CNNs) and graph convolutional networks (GCNs), vision-based markerless pose estimation has progressed rapidly. Earlier clinical uses of pose estimation for Parkinson's disease provided analysis with disease-discriminative characteristics through 2D pose estimation from video sources. Li et al. [13] demonstrated that motion patterns unique to PD, in addition to features of levodopa-induced dyskinesia, could be successfully identified using visual pose estimation techniques.

Zeng et al. [14] generated gait-score predictions for the MDS-UPDRS using their skeleton-contour fusion network combined with dual-view imaging and reported 71.25% prediction accuracy within an early-stage patient population.

With the use of multiple cameras, researchers have been able to achieve more accurate three-dimensional (3D) kinematic

analysis. Kondo et al. [15] reported FoG detection accuracy of 93.2% through multi-view 3D pose estimation and object tracking. Guo et al. [16] proposed a GCN-based classifier with an accuracy of 91.32% for five-category FoG classification based on multi-camera imaging.

Franco et al. [12] authored a systematic literature review of deep learning and gait analysis in PD. The review included 25 studies published between 2018 and 2023, with both wearable and non-wearable sensing approaches represented. The authors identified CNN-based signal analysis and human pose-estimation-based video analysis as important methodological paradigms and highlighted challenges related to dataset heterogeneity and clinical validation [12].

More recently, Kim et al. [26] performed a direct clinical comparison of markerless HPE against 3D gold-standard motion capture for PD gait analysis, evaluating near-frontal and lateral camera perspectives. Their findings confirmed acceptable agreement for spatiotemporal parameters but revealed limitations in kinematic angle estimation accuracy.

The recent 2025 systematic review by di Biase et al. further indicates that gait remains the most frequently investigated motor task in computer-vision-based PD research and emphasizes standardized acquisition and external validation as important requirements for clinical translation [21].

D. 3D Human Pose Estimation: Deep Learning Architectures

3D human pose estimation research has followed a number of successive architectural developments. Earlier methods used temporal convolutional networks (TCNs) that processed sequences of 2D poses, from which they estimated 3D coordinates for the joints of a human skeleton. These methods were followed by graph convolutional networks (GCNs), which provide an effective way of modelling the 3D human skeleton and local joint dependencies [27].

The introduction of Transformer networks and global self-attention mechanisms enabled the modelling of long-range spatiotemporal dependencies across joint trajectories. Through STC-Former, Tang et al. [17] incorporated spatiotemporal cross-attention and positional encoding. Yu et al. [18] introduced GLA-GCN, which combines global spatiotemporal information with local joint information.

Hybrid architectures that combine Transformers and GCNs have also emerged. Mehraban et al. [19] presented MotionAGFormer, which computes joint sequences through parallel Transformer and GCN streams before adaptively merging the representations. Ye et al. [20] expanded this design principle through

Motion2Vec, using Transformer-GCN dual streams and self-distillation.

Li et al. [28] incorporated skeletal structural priors and uncertainty into a self-attention framework. Hu et al. [29] investigated a high-resolution GCN-MLP hybrid model for 3D pose estimation using smartphone-based measurements and VICON reference measurements.

Self-supervised learning has also increasingly been applied to human-motion representation learning. GaitForeMer, for example, used self-supervised motion forecasting to pretrain a Transformer on 3D human motion before adapting it to gait-impairment estimation, demonstrating the value of learning general motion representations before using limited clinical labels [30]. More recent biomechanical work has similarly used masked reconstruction of gait kinematics for self-supervised pretraining and subsequent fine-tuning with limited labeled data [23].

E. Research Gaps and Motivation

A thorough analysis of the existing body of work demonstrates that substantial gaps remain in gait-analysis research. First, many PD-specific vision-based gait-analysis systems rely on multiple cameras, introducing additional operational requirements and infrastructure. Second, generic 3D pose-estimation models have demonstrated strong performance on benchmark datasets, but their direct application to clinical populations, particularly individuals with PD, remains less extensively validated.

Third, supervised learning approaches continue to depend on labeled data, whereas clinical gait datasets can contain relatively limited disease-specific annotations. This creates an important opportunity for self-supervised learning methods that can exploit the temporal structure of gait sequences without requiring clinical labels during pretraining.

The principal research gap addressed by this study is therefore not simply 3D pose estimation, but the learning of temporal gait representations before supervised PD classification. Existing studies have demonstrated self-supervised motion learning and supervised PD gait analysis separately, while this study combines masked sequence reconstruction with a BiLSTM temporal encoder and subsequent PD classification using 3D gait sequences.

The proposed approach therefore focuses on learning intrinsic temporal characteristics of gait before adapting the learned representation to the PD classification task.

F. Summary Comparison of Related Works

TABLE 1. Comparison of Representative Related Works in PD Gait Analysis and 3D Pose Estimation

Ref.	Author / Year	Method	Modality	Task / Application	Dataset Subjects /	Performance	Limitations
[12]	Franco et al., 2024	Deep learning (CNN, HPE)	Wearable + camera	PD classification, monitoring	25 studies, 2018–2023	N/A	Heterogeneous datasets, limited 3D validation
[13]	Li et al., 2018	Vision-based pose estimation (CNN)	Monocular camera	PD & LID feature extraction	PD patients	Feature extraction validated	No 3D, limited scalability
[14]	Zeng et al., 2023	Skeleton-contour CNN	Dual-camera	MDS-UPDRS gait prediction	80 early PD + healthy	Accuracy = 71.25%	Dual-camera, moderate accuracy

[15]	Kondo et al., 2024	3D pose + target tracking	Multi-view camera	FoG vs. walking-stop detection	Clinical PD video data	Accuracy = 93.2%	Multi-camera, clinical-only setting
[16]	Guo et al., 2024	GCN	Multi-camera	5-class FoG assessment	PD patients	Accuracy = 91.32%	High system complexity, multi-camera required
[19]	Mehraban et al., 2024	Transformer-GCNFormer	Monocular video	3D pose estimation	Human3.6M, MPI-INF-3DHP	MPJPE = 42.5 mm	Not clinically validated
[20]	Ye et al., 2025	Transformer-GCN + self-distillation	Monocular video	3D pose estimation	Human3.6M, MPI-INF-3DHP	MPJPE = 38.0 mm	No demonstrated PD application
[24]	Rangel & Martínez, 2024/2025	Self-supervised generative representation	Gait video	Parkinsonian gait modelling	PD and control datasets	ROC-AUC reported	Limited cohort size
[25]	Al-Adhaileh et al., 2025	CNN + BiLSTM + Attention	Wearable IMU	FoG detection	Daily living + multimodal datasets	AUC = 0.96, F1 = 0.94	Wearable-dependent
[26]	Kim et al., 2025	Markerless HPE	Single camera	Spatiotemporal gait extraction	PD patients	High agreement with gold standard	Limited joint-angle accuracy
[29]	Hu et al., 2024	HGcnMLP	Smartphone monocular camera	3D pose estimation	Musculoskeletal patients	High ICC vs. VICON	Not PD-specific
[31]	Gholami et al., 2023	Weakly supervised learning	Wearable IMU	Automatic gait labeling	PD patients	Accuracy = 89%	IMU drift, limited pose detail
[32]	Özel / 2026	Self-supervised masked autoencoder (MAE); masked gait modeling; downstream classification; explainability analysis	VGRF gait signals	PD vs. healthy-control detection under limited labeled data	Public VGRF gait data from PD patients and healthy controls	Strong PD classification under low-resource/limited-label conditions; masking-ratio ablation performed	Uses 1D VGRF rather than 3D skeletal gait; anatomical joint relationships are not explicitly modeled; transfer to 3D motion remains untested.
[33]	Brand et al. / 2026	ElderNet; self-supervised deep representation learning for gait-metric estimation	Wrist-worn accelerometer	Continuous gait speed, cadence and stride-length estimation; mobility-disability classification	819 older adults (Rush Memory and Aging Project) + 85 participants with gait impairments (Mobilise-D)	Gait-speed MAE: 8.82 cm/s; ICC: 0.87; mobility-disability AUC: 0.80	Not specifically designed for PD diagnosis; wrist accelerometry does not explicitly capture 3D skeletal geometry; primary task is gait assessment.
[34]	Lai et al. / 2026	Skeletal Gait Energy Image (SkeGEI); data augmentation; CNN-LSTM; RF/SHAP-assisted feature analysis	3D Kinect V2 skeletal/movement data	Prediction/classification of PD motor phenotypes from straight and turning walking	Clinical PD cohort assessed using Kinect V2 3D motion acquisition	CNN-LSTM reported as highest-performing approach for straight and turning walking; RF/SHAP supported feature analysis	Supervised rather than SSL; dependence on Kinect acquisition conditions; limited evidence of label-efficient pretraining and cross-system generalization.
[35]	Mahali et al. / 2026	Hybrid ConvTransLSTM: CNN + Transformer + BiLSTM; sliding-window augmentation; voting; cross-subject CV	Upper- and lower-limb IMU accelerometer/gyroscope signals	Three-class classification: healthy/no-PD, early PD, advanced PD	102 subjects: training set 48; test set 54	Accuracy: 0.90; F1: 0.90; confidence: 0.97 (reported optimal setting)	Supervised; only a subset has explicit H&Y stage labels; broad labels may obscure progression; larger windows increase computation.
[36]	X. Wang et al. / 2026	Multi-Scale Directed Graph Neural Network for dynamic gait representation	VGRF signals	PD detection from gait	VGRF gait recordings from PD and control participants	Reported improved PD recognition through graph-based modeling of	VGRF rather than 3D skeletal joint trajectories; no SSL; sensor-specific graph

						dynamic VGRF relationships	representation may constrain transferability.
[37]	D. Wang et al. / 2026	CNN-based EEG representation + Mutual Cross-Attention multimodal fusion + unsupervised clustering	Resting-state EEG + dual-task gait features	Data-driven subtyping of early-stage PD	40 idiopathic early-stage PD patients	Three clinically distinct PD subtypes identified; multimodal fusion improved clinical separation	Small cohort (n=40); focused on PD subtyping rather than PD-vs-control detection; multimodal acquisition complexity; not 3D skeletal SSL.
[38]	Ji et al. / 2026	Improved Particle Swarm Optimization + SVM (IPSO-SVM); ZUPT trajectory estimation	Plantar-pressure sensors + IMU	PD gait analysis and UPDRS-based severity prediction	Multisource gait data from PD patients and healthy controls	RBF IPSO-SVM average accuracy: 95.68%	Depends on engineered features and sensor fusion; supervised ML; no self-supervised representation learning; no direct 3D skeletal representation.
[39]	Carter & Wang / 2026	Seven ML classifiers; RFE/SFS feature selection; bilateral/asymmetric gait features	In-shoe VGRF/load sensors	PD vs. healthy-control diagnosis using gait asymmetry	279 gait recordings from three public datasets; 26 spatial, temporal and kinetic features	Random Forest balanced accuracy: 84.3%	Hand-engineered gait features; conventional supervised ML; no direct temporal representation learning from raw sequences; VGRF rather than 3D skeleton.
[40]	Michaud et al. / 2026	Systematic review of real-world gait-monitoring studies	Wearable sensors	Synthesis of real-world gait assessment methods and outcomes in PD	Multiple published studies involving PD gait monitoring with wearable sensors	N/A - systematic review rather than a single predictive model	Substantial heterogeneity in sensors, protocols, populations, gait metrics and analytical approaches; does not propose a 3D gait SSL diagnostic model.
Proposed	Proposed work	Self-supervised BiLSTM with masked 3D gait reconstruction + supervised fine-tuning	3D gait sequences	PD detection under limited labeled data	CARE-PD	Accuracy = 78.8%, AUC = 0.81, reported recall = 0.90	Moderate overall accuracy; requires larger external validation

Note: The comparison in Table 1 highlights a methodological distinction between the proposed approach and previous studies. Existing approaches have investigated wearable sensing, multi-camera gait analysis, monocular pose estimation, graph-based modelling, Transformer-GCN architectures, and supervised deep learning. The proposed work instead focuses on learning temporal representations directly from 3D gait sequences through masked sequence reconstruction before applying supervised PD classification

III. DATASET AND PREPROCESSING

A. Dataset Description

This study uses the CARE-PD dataset, which is a public resource for research into Parkinson's disease [41]. The CARE-PD dataset differs from earlier datasets that relied exclusively on limited sensor readings or smaller cohorts because it provides 3D gait data from different experimental conditions and populations.

The dataset contains three-dimensional gait data recorded using 3D motion capture systems. The captured gait data provide information about the movement of different joints over time and therefore contain information about both local and global movement characteristics.

Another significant advantage of the CARE-PD dataset is the number of different experimental scenarios. Each subject is recorded in a number of different experimental conditions, including before and after medication. The availability of

multiple experimental conditions provides an opportunity to investigate variations in gait under different circumstances.

For every sample, there are three main elements found in the dataset. The first element is the actual 3D movement data (pose) for each of the joints at a point in time. Secondly, the data contain a time component, meaning that each sample represents a sequence of recorded frames. Thirdly, where applicable, clinical annotations and contextual information are available.

The dataset includes individuals diagnosed with PD and individuals who did not have a neurological condition. The presence of multiple individuals is important for reducing the possibility of learning subject-specific movement characteristics instead of disease-related characteristics.

The scale and structure of the CARE-PD study data are important for understanding its potential value for machine learning in Parkinson's disease research. The availability of temporally rich 3D movement data makes it particularly suitable for investigating representation-learning approaches in which the model can learn from the structure of gait sequences before supervised classification.

B. Data Structure

For creating an effective modeling pipeline it is essential to understand the structure of the dataset. The data are organized in time-series format, with human body pose and positioning tracked over time.

The data sample is represented as a $(T \times 72)$ matrix. In this case, T refers to the number of time points and 72 represents the pose features. The value of T varies between samples because recordings may have different durations.

Some samples may contain only a few time steps, while others may contain thousands of time steps. Therefore, preprocessing is needed to create a standard sequence length. Along with pose data, additional metadata may exist in each sample, including subject identification, trial type, medication status, and experimental conditions.

The hierarchical structure of the data also raises concerns regarding data contamination or leakage. Samples obtained from the same subject must not appear in both training and testing datasets because this could allow the model to learn subject-specific characteristics rather than disease-related characteristics.

C. Preprocessing

Due to the variability and complexity of the raw dataset, a preprocessing pipeline is necessary to prepare the data for deep learning. The objectives of preprocessing were to standardize the format of input data, reduce redundancy, and improve the ability of the model to learn meaningful gait patterns.

The first step was breaking the continuous recordings of gait into fixed-length segments. Since there is variability in the lengths of the original recordings, a fixed window of 500 time steps was selected.

Temporal down sampling was then applied to reduce redundant consecutive frames:

$$X' = X[:, ::5]$$

Downsampling by selecting every fifth frame reduced the number of frames from 500 to 100. A second downsampling operation reduced the sequence to 20 frames. This temporal compression reduced highly correlated information while retaining representative motion patterns.

Normalization was also applied to ensure that the input features were represented within a suitable range for neural-network training. Missing or invalid data were handled during preprocessing.

The preprocessing and labeling procedure was applied before model training, and the same preprocessing principles were maintained across the training and testing stages.

D. Train-Test Split

In a fair and consistent way, the data were divided into training and testing subsets. Stratified sampling was used to create an 80% training and 20% testing split.

The training set was used for self-supervised pre-training and supervised fine-tuning, whereas the testing set was kept separate for evaluating generalization.

Avoiding data leakage is important because the dataset can contain multiple samples from the same subject. In the final implementation, the 80%/20% split was stratified at the sample level using `random_state = 42`; participant-wise group splitting was not implemented. Consequently, repeated trials from the same participant may occur across the training and testing subsets, and the reported results should be interpreted as sample-level rather than unseen-subject generalization.

For reproducibility, the same random seed was used throughout the experiments.

IV. METHODOLOGY

A. Overview

This proposed framework allows for the effective learning of discriminative patterns in high-dimensional time-series gait data with limited labeled samples using a 2-stage learning approach that incorporates self-supervised representation learning for pre-training followed by supervised classification for post-training. Figure 1 represents the overall methodology framework.

In the self-supervised pre-training phase, the model learns to extract the intrinsic temporal dynamics of a given set of movements without seeking any specific label-based relationships. The primary focus of this phase is to obtain a model capable of describing motion patterns present in gait data, including how movement characteristics change through time.

Once the pre-training phase has been completed, the model is fine-tuned for the classification of gait patterns of individuals with Parkinson's disease. At this point, the model learns representations of gait data that can discriminate between gait patterns associated with individuals with and without Parkinsonism.

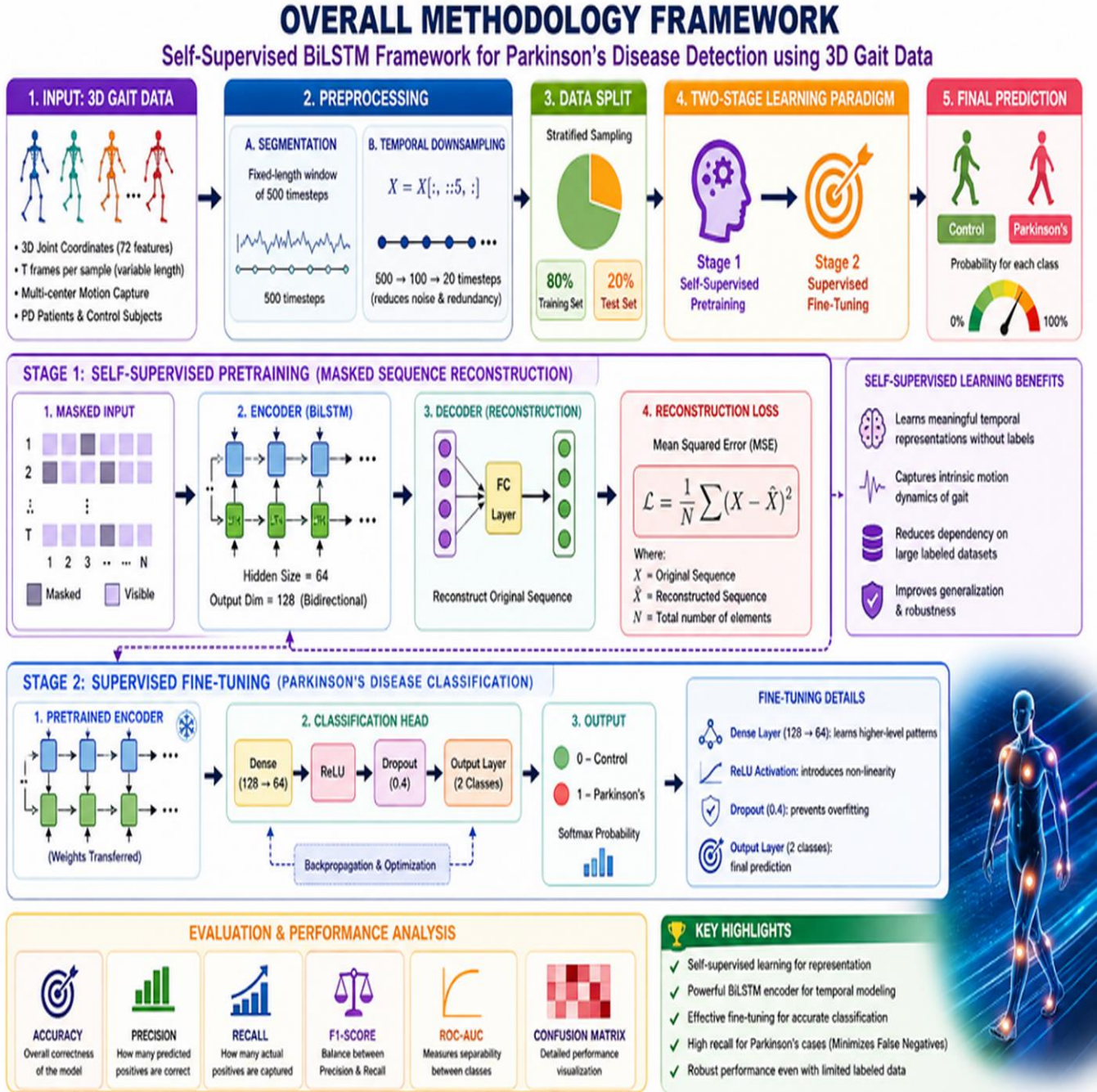


Figure 1. Methodological Framework

B. Self-Supervised Learning

The self-supervised learning phase is executed using masked sequence reconstruction. A given amount of the input sequence is randomly masked and the model is trained to reconstruct the masked areas. This requires the model to learn contextual dependencies between the current segment and the sequence as a whole.

The input sequence is represented as:

$$X \in R^{T \times 72}$$

where T represents the number of time steps and 72 represents the number of pose features.

A mask is used to randomly obscure portions of the input X. The masked input is then sent to the model, which is trained to reconstruct the original input.

The objective is described using the Mean Squared Error (MSE) loss function:

$$\mathcal{L} = \frac{1}{N} \sum (X - \hat{X})^2$$

where N is the number of elements in the sequence, X is the original input, and $\hat{X}$ is the reconstructed output.

The loss function penalizes deviations between the original sequence and the reconstructed sequence, encouraging the model to learn an accurate representation of the underlying dynamics of the motion. One major advantage of this method is the absence of the need for labeled data during the pretraining task.

C. Model Architecture

In this research, a neural network framework that is called Bidirectional Long Short-Term Memory (BiLSTM) is utilized to fulfill the purpose of analyzing sequential data. BiLSTM is different from a traditional LSTM model because it processes input sequences in both directions using past and future information.

The encoder is built up from BiLSTM with 64 hidden units. The output generated at each timestamp contains two feature vector values. Thus, two feature vectors are concatenated to form the feature vector of size 128 (64×2).

Due to the generation of the feature vector of size 128 at each time point, the model would be able to find temporal dependencies included in the behavior of a gait sequence.

The encoder is regarded as the foundation of the architecture because it provides the latent representation necessary to reconstruct the sequence taking into account the presence of certain corruptions.

The decoder is a layer that is connected fully. That means that the decoder receives encoded features and gives back obtained values into the original feature space. The purpose of the decoder is to recreate the sequence of inputs from latent-space representation.

D. Fine-Tuning

The encoder was initially self-trained in gait dynamics before it underwent fine-tuning to create a classification head used to diagnose Parkinson's disease.

The classification head is composed of a chain of fully connected layers starting from layer one which decreases the vector dimensionality from 128 to 64. This layer uses the Rectified Linear Unit (ReLU) activation function allowing for learning complex classification boundaries.

The dropout layer is sometimes used to circumvent overfitting. This layer is set with a dropout ratio of 0.4. Lastly, two output neurons are used to get probability scores for classification of control and PD patients by applying softmax on two neurons.

The complete model with the pretrained encoder was fine-tuned with labeled data.

V. EXPERIMENTAL RESULTS AND EVALUATION

A. Performance Metrics

In order to evaluate the effectiveness of the suggested model, there was a need for specific performance criteria. In general, depending on just one measure is not advisable, as there may be

misleading evaluations. This might be particularly true in medical classification problems that might incur class imbalance and different costs related to misclassification.

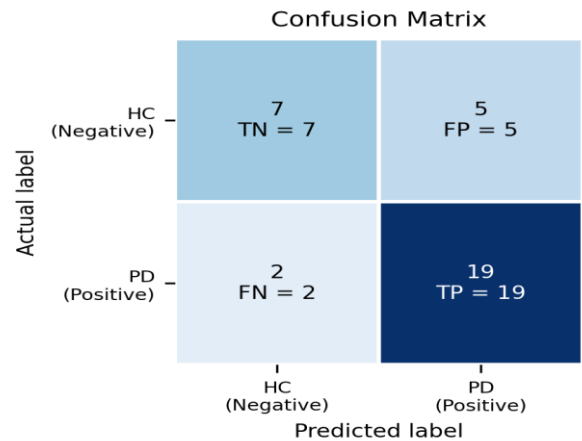
Accuracy denotes generally how well the classifier performs according to how high the percentage of correctly classified cases is from all predictions made. Precision describes the number of correctly predicted positive cases and the total number of cases predicted positive. Recall or sensitivity measures the number of positive cases that have been identified against the total amount of the actual positive cases. F1-score shows combined precision as well as recall metrics. The last aspect that should be mentioned is that ROC and AUC assess how good the classifier is at distinguishing between both classes according to different decision thresholds.

B. Results

The results achieved through the exercise of the proposed self-supervised BiLSTM model signify that it functioned adequately in regard with the metrics assessed. The model attained a total accuracy of 78.8%. The ROC-AUC of the designed model equals 0.81 showing its discriminating power related to the identification of Parkinson's versus non-Parkinson's walk. The reported recall for the cases of Parkinson's disease is equal to 0.90 meaning that the model was able to recognize the majority of the cases related to the same disease from the tested set. However, the interpretation of the results presented in this paper should not be taken as evidence of the effectiveness of clinical diagnosis based on the pre-evaluated datasets.

C. Confusion Matrix

The confusion matrix in Figure 2 provides a detailed breakdown of the model's predictions.



Accuracy = 78.8% Precision = 79.2% Recall = 90.5% F1 = 84.4%

Figure 2. Confusion Matrix with Color-Coded Cells

The confusion matrix contains:

- TP = 19 — PD cases correctly classified
- FP = 5 — healthy controls misclassified as PD
- FN = 2 — PD cases missed
- TN = 7 — healthy controls correctly classified

The overall accuracy reported is 78.8% while the precision, recall, and F1-score for the positive class is calculated to be 79.2%, 90.5%, and 84.4%, respectively.

The data obtained reveals several important characteristics about the model's classification performance. The observations made from false-positive cases show that normal subjects were diagnosed with the condition of Parkinson's disease. On the other hand, the case of false-negative showed that some samples with Parkinson's disease remained undetected.

The presence of false-positive and false-negative predictions reinforces the argument that accuracy is not the only measure to assess clinical classification performance. The ROC-AUC and class level measures provide more information about model discrimination and detection capability.

D. Comparative Analysis

To assess the effectiveness of the proposed approach, its performance was compared with several baseline models commonly used for sequence-classification tasks. This is shown in Table 2.

TABLE 2. Comparison with Baseline Models

Model	Accuracy
LSTM	72.7%
BiLSTM	75.8%
CNN-BiLSTM	63.6%
Transformer	72.7%
Proposed Model	78.8%

The experimental data indicate that the proposed model achieved the highest accuracy among the comparison models.

Although the traditional LSTM and Transformer models performed adequately and demonstrated an ability to represent temporal patterns, they were trained using labeled datasets. The BiLSTM model performed better than the conventional LSTM because it processes temporal patterns in both directions.

The CNN-BiLSTM model produced the lowest accuracy among the evaluated baselines. This result may be associated with the increased model complexity under the available training conditions.

The comparative results provide evidence supporting the usefulness of the proposed self-supervised pretraining strategy. However, because the comparison is based primarily on accuracy and does not include formal statistical significance testing, the difference between models should be interpreted as an observed performance difference rather than a statistically established superiority.

VI. DISCUSSION

The results of the experiments indicate that using self-supervised learning within the modeling pipeline provides a useful approach for learning temporal representations from gait sequences. Traditional supervised models depend on labeled data, whereas the proposed model first captures intrinsic temporal features from the gait sequences themselves.

This enables the model to learn temporal characteristics such as variations in joint coordination and rhythm of movement before the supervised classification stage. This is particularly relevant when clinical labels are limited.

The proposed framework achieved an accuracy of 78.8% and ROC-AUC of 0.81, while the reported recall for PD cases was 0.90. Among the evaluated baseline models, the proposed model achieved the highest accuracy. These findings suggest that the use of self-supervised representation learning can provide a useful alternative to training a temporal classifier entirely from labeled data. The findings are consistent with the broader direction of recent gait-analysis research. Rangel and Martínez investigated self-supervised generative representation learning for Parkinsonian gait and reported that gait-motion representations could be learned without relying exclusively on conventional supervised rules. Similarly, Li et al. demonstrated in 2025 that masked-reconstruction self-supervised learning improved data efficiency in biomechanical gait modelling, particularly when labeled data were scarce [23], [24].

The present study differs from these approaches in its use of a BiLSTM temporal encoder with masked reconstruction for downstream PD classification using 3D gait sequences. The contribution is therefore centered on temporal representation learning rather than the use of increasingly complex architectures alone.

The results should also be considered in relation to recent systematic reviews. Franco et al. reported that deep learning-based gait analysis for PD increasingly uses both wearable and non-wearable sensing modalities, while dataset heterogeneity and clinical validation remain important challenges [12]. More recent computer-vision literature similarly emphasizes external validation and standardized data acquisition before clinical translation [21].

An important implication of the present study is therefore not that the proposed model is ready for clinical deployment, but that self-supervised temporal representation learning provides a feasible research direction for reducing dependence on disease-specific labels. The 78.8% accuracy also indicates that additional validation and model development remain necessary.

The confusion matrix further demonstrates that the model does not eliminate classification errors. False-positive predictions were observed among control participants, while false-negative predictions were also present. Therefore, the model should currently be considered a research prototype rather than a standalone diagnostic system.

The results also highlight the importance of validation beyond a single train-test split. Recent systematic evidence indicates substantial heterogeneity among AI-based PD gait studies and emphasizes the need for methodological rigor and external validation [42].

Future studies should therefore evaluate the framework using larger cohorts, independent datasets, repeated experiments, confidence intervals, and statistical comparisons among models. Such validation would provide stronger evidence regarding the reproducibility and generalizability of the observed performance.

VII. STUDY LIMITATIONS

Several limitations should be considered when interpreting the findings of this study.

First, the evaluation was conducted using the available CARE-PD data and the experimental train-test split. The final implementation used a stratified sample-level split rather than participant-wise group splitting; therefore, possible overlap of trials from the same participant across training and testing is a limitation. The size and diversity of the final evaluation set also limit the extent to which the results can be generalized to broader clinical populations.

Second, the current framework uses 3D gait information without incorporating additional clinical, demographic, imaging, or wearable-sensor information. Such complementary information may provide additional evidence for differentiating PD from control participants.

Third, the proposed model achieved the highest accuracy among the evaluated baseline models, but the absolute difference between some models was relatively modest. The current study does not include formal statistical significance testing across repeated independent experiments. Therefore, the observed performance difference should not be interpreted as definitive evidence that the proposed approach will consistently outperform all alternative architectures.

Fourth, independent external validation has not yet been performed. The model should therefore be evaluated on independent cohorts from different acquisition environments and populations before any clinical deployment is considered.

Fifth, the current work focuses on binary PD detection rather than disease staging, severity estimation, medication-state prediction, or longitudinal monitoring.

Finally, the model's computational efficiency, calibration, interpretability, robustness to acquisition variations, and real-time performance have not been comprehensively evaluated. These factors will be important for translation from a research prototype to a clinically useful decision-support system.

VIII. NOVELTY AND CONTRIBUTION

The novelty of this study lies primarily in the learning strategy rather than in introducing a new recurrent neural-network architecture. The proposed framework integrates three components into a single PD gait-detection pipeline: (1) 3D gait-sequence representation, (2) masked sequence reconstruction for self-supervised temporal pretraining, and (3) supervised fine-tuning of a pretrained BiLSTM encoder for PD classification.

Existing studies have demonstrated several related directions, including supervised PD gait classification, graph-based gait analysis, Transformer-based 3D pose estimation, and self-supervised motion representation learning. However, the proposed framework specifically investigates masked reconstruction of 3D gait sequences using a BiLSTM temporal encoder and subsequently transfers the learned representation to PD classification under limited labeled-data conditions.

The contribution is therefore the integration and evaluation of self-supervised temporal representation learning with clinical 3D gait classification. The study further provides a direct comparison against LSTM, BiLSTM, CNN-BiLSTM, and Transformer baselines using the same experimental setting.

This distinction is important because the objective is not simply to increase model complexity, but to reduce the dependence on disease-specific annotations by allowing the encoder to learn temporal gait characteristics before supervised learning.

IX. FUTURE RESEARCH

The current framework shows promise for detecting Parkinson's disease through the processing of 3D gait data, and numerous opportunities may be explored to enhance its functionality and extend its capabilities.

One avenue of opportunity is the integration of Graph Neural Networks (GNNs) for modeling at the joint level. Because human gait offers an inherent structure in that humans have multiple joints connected in a biomechanical graph, traditional sequence-modeling methods treat pose features as flat, linear vectors without explicitly recognizing their spatial relationships with one another. GNNs can provide an explicit representation of the biomechanical structure of the human body and may therefore help model subtle spatial relationships between body parts.

Multimodal data fusion is another area of exploration. In multimodal data fusion, gait data can be combined with other types of data, for example clinical records, information from wearables or sensors, and imaging, to create more complete characterizations of patients with Parkinson's disease. Various fusion techniques may be investigated, including attention-based fusion and cross-modal learning.

The application of the developed model to actual real-time wearable systems is another potential direction. Smartwatch and motion-detection technologies provide opportunities for passive gait monitoring in community settings. Adapting the gait model for efficient inference on resource-constrained devices could support future remote-monitoring applications.

Future research should also investigate larger and more diverse cohorts, independent external validation, repeated cross-validation, confidence intervals, calibration, model interpretability, and statistical significance testing. These steps would strengthen the evidence required for evaluating the generalizability and clinical applicability of the proposed framework.

X. CONCLUSION

The study presents a self-learning BiLSTM model that can detect Parkinson's diseases based on 3D gait information, thus overcoming the problem of limited labeled data in medical machine learning. The learning method in question consists of two stages; first, information about the time is obtained and then used in the goal of supervised learning.

The self-learning methodology has achieved an accuracy of 78.8% and ROC-AUC of 0.81, while recall in the case of PD was recorded as 0.90. The method in question showed the best results among those assessed by the authors, including LSTM, BiLSTM, CNN-BiLSTM, and Transformer methods.

In general, it is possible to state that the importance of self-learning technique lays in its possibility of learning about gait representation before performing the specific disease classification tasks. The results prove the statement that the techniques used in the research can be beneficial for obtaining data in cases when labeled clinical data is absent.

Nevertheless, it is important to note that the results have preliminary nature and cannot be regarded as enough to launch the clinical research.

Data Availability

The dataset used in this study is publicly available through the CARE-PD repository:

<https://borealisdata.ca/dataset.xhtml?persistentId=doi:10.5683/S-P3/TWIKMK>

Declaration of Competing Interest

The authors declare that they have no competing interests

Author Contributions

All the authors contributed to the conceptualization, methodology, data analysis, implementation, interpretation of results, and preparation of the manuscript. All authors have read and approved the final manuscript.

REFERENCES

- [1] B. R. Bloem, M. S. Okun, and C. Klein, "Parkinson's disease," *The Lancet*, vol. 397, no. 10291, pp. 2284–2303, 2021.
- [2] M. J. Armstrong and M. S. Okun, "Diagnosis and treatment of Parkinson disease: A review," *JAMA*, vol. 323, no. 6, pp. 548–560, 2020.
- [3] Y. Luo, L. Qiao, M. Li, X. Wen, W. Zhang, and X. Li, "Global, regional, national epidemiology and trends of Parkinson's disease from 1990 to 2021: Findings from the Global Burden of Disease Study 2021," *Frontiers in Aging Neuroscience*, vol. 16, Art. no. 1498756, 2025, doi: 10.3389/fnagi.2024.1498756.
- [4] H. Gu et al., "A non-contact real-time muscle activity measurement system using dual IMUs for knee kinematic analysis during walking," *IEEE Sensors Journal*, 2022.
- [5] R. Mallat et al., "A portable visual-inertial motion capture system integrating IMU data with visual information for accurate pose estimation," *IEEE Transactions on Instrumentation and Measurement*, 2022.
- [6] T. Virmani, F. Moskowitz, R. Bhatt, and J. J. Bhatt, "A tablet-based application to assist gait training in Parkinson's disease," *Frontiers in Neurology*, vol. 6, p. 9, 2015.
- [7] L. Di Biase, A. Di Santo, M. E. Caminiti, A. De Liso, S. A. Shah, L. Ricci, and V. Di Lazzaro, "Gait analysis in Parkinson's disease: An overview of the most accurate markers for diagnosis and symptoms monitoring," *Sensors*, vol. 20, no. 12, p. 3529, 2020.
- [8] P. Meriaux et al., "A study of Vicon system positioning to measure human movement," *Sensors*, vol. 17, no. 7, p. 1591, 2017.
- [9] M. Parati et al., "XSENS system for clinical gait analysis," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2022.
- [10] C. Price, D. Parker, and C. J. Nester, "Validity and repeatability of three in-shoe pressure measurement systems," *Gait & Posture*, vol. 46, pp. 69–74, 2016, doi: 10.1016/j.gaitpost.2016.01.026.
- [11] C. Altham et al., "Markerless human pose estimation for clinical motion analysis: A systematic review," *Gait & Posture*, 2024.
- [12] Franco, M. Russo, M. Amboni, A. M. Ponsiglione, F. Di Filippo, M. Romano, F. Amato, and C. Ricciardi, "The role of deep learning and gait analysis in Parkinson's disease: A systematic review," *Sensors*, vol. 24, no. 18, p. 5957, 2024, doi: 10.3390/s24185957.
- [13] Y. Li et al., "Vision-based assessment of Parkinson's disease severity and levodopa-induced dyskinesia," in *Proc. AAAI Conference on Artificial Intelligence*, 2018.
- [14] R. Zeng et al., "Video analysis for MDS-UPDRS gait score prediction using skeleton-contour fusion convolutional network," *Frontiers in Neurology*, 2023.
- [15] Y. Kondo et al., "Video-based automatic detection of freezing of gait combining target tracking and 3D pose estimation," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2024.
- [16] Z. Guo et al., "Automated five-category freezing of gait assessment using graph convolutional networks," in *Proc. IEEE International Conference on Robotics and Automation (ICRA)*, 2024.
- [17] Z. Tang, Z. Qiu, Y. Hao, R. Hong, and T. Yao, "3D human pose estimation with spatio-temporal criss-cross attention," in *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023.
- [18] X. Yu, Z. Zhang, Y. Liu, S. Zhong, Y. Liu, and C. W. Chen, "GLA-GCN: Global-local adaptive graph convolutional network for 3D human pose estimation from monocular video," in *Proc. IEEE/CVF International Conference on Computer Vision (ICCV)*, 2023, pp. 8818–8829.
- [19] S. Mehraban, V. Adeli, and B. Taati, "MotionAGFormer: Enhancing 3D human pose estimation with a Transformer-GCNFormer network," in *Proc. IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 2024, pp. 6920–6930.
- [20] M. Ye, L. Yang, H. Zhu, Z. Zheng, X. Wang, and Y. Lou, "Dual-stream Transformer-GCN model with contextualized representations learning for monocular 3D human pose estimation," *arXiv preprint arXiv:2504.01764*, 2025.
- [21] L. di Biase, P. M. Pecoraro, and F. Bugamelli, "AI video analysis in Parkinson's disease: A systematic review of the

- most accurate computer vision tools for diagnosis, symptom monitoring, and therapy management,” *Sensors*, vol. 25, no. 20, p. 6373, 2025, doi: 10.3390/s25206373.
- [22] H. Rabie and M. A. Akhloufi, “A review of machine learning and deep learning for Parkinson’s disease detection,” *Discover Artificial Intelligence*, vol. 5, p. 24, 2025, doi: 10.1007/s44163-025-00241-9.
- [23] Y. Li, J. He, B. Liew, D. S. Hollinger, Q. Mei, B. Gholami, M. Fasli, K. McDonald-Maier, and X. Zhai, “Self-supervised learning enhances accuracy and data efficiency in lower-limb joint moment estimation from gait kinematics,” *Frontiers in Bioengineering and Biotechnology*, vol. 13, p. 1633513, 2025, doi: 10.3389/fbioe.2025.1633513.
- [24] E. Rangel and F. Martínez, “Parkinsonian gait modelling from an anomaly deep representation,” *Multimedia Tools and Applications*, vol. 84, pp. 21605–21623, 2025, doi: 10.1007/s11042-024-19961-8.
- [25] M. H. Al-Adhaileh, A. Wadood, T. H. Aldhyani, A. Khan, M. I. Uddin, and A. H. Al-Nefaie, “Deep learning techniques for detecting freezing of gait episodes in Parkinson’s disease using wearable sensors,” *Frontiers in Physiology*, vol. 16, 2025, doi: 10.3389/fphys.2025.1581699.
- [26] R. Kim et al., “Assessment of temporospatial and kinematic gait parameters using human pose estimation in patients with Parkinson’s disease: A comparison between near-frontal and lateral views,” *PLOS ONE*, 2025, doi: 10.1371/journal.pone.0317933.
- [27] X. Chen et al., “Anatomy-aware 3D human pose estimation with bone-based pose decomposition,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 32, no. 1, pp. 198–209, 2022.
- [28] H. Li et al., “3D human pose estimation using skeletal structure information and uncertainty-aware self-attention,” *IEEE Transactions on Image Processing*, 2023.
- [29] R. Hu, Y. Diao, Y. Wang, G. Li, R. He, Y. Ning, N. Lou, G. Li, and G. Zhao, “A novel multi-level 3D pose estimation framework for gait detection of Parkinson’s disease using monocular video,” *Frontiers in Bioengineering and Biotechnology*, vol. 11, p. 1335251, 2024, doi: 10.3389/fbioe.2023.1335251.
- [30] M. Endo et al., “GaitForeMer: Self-Supervised Pre-Training of Transformers via Human Motion Forecasting for Few-Shot Gait Impairment Severity Estimation,” 2022.
- [31] M. Gholami et al., “Automatic weakly-supervised gait labeling for Parkinson’s disease,” in *Proc. IEEE Engineering in Medicine and Biology Conference (EMBC)*, 2023.
- [32] E. Özel, “Self-Supervised Masked Gait Modeling: An Explainable Low-Resource Learning Approach for Parkinson’s Disease Detection,” *IEEE Access*, vol. 14, pp. 82890–82899, 2026, doi: 10.1109/ACCESS.2026.3697731.
- [33] Y. E. Brand, A. S. Buchman, F. Kluge, L. Palmerini, C. Becker, A. Cereatti, W. Maetzler, B. Vereijken, A. J. Yarnall, L. Rochester, S. Del Din, A. Mueller, J. M. Hausdorff, and O. Perlman, “Continuous assessment of daily-living gait using self-supervised learning of wrist-worn accelerometer data,” *npj Digital Medicine*, vol. 9, Art. no. 338, 2026, doi: 10.1038/s41746-026-02528-2.
- [34] Y.-R. Lai, C.-C. Huang, C.-Y. Lien, Y.-F. Chiang, C.-P. Ting, C.-F. Kung, and C.-H. Lu, “A comprehensive deep learning model for motor phenotypes of Parkinson’s disease using three-dimensional Kinect V2 detectors,” *Gait & Posture*, vol. 123, Art. no. 110000, Jan. 2026, doi: 10.1016/j.gaitpost.2025.110000.
- [35] M. I. Mahali, C. Avian, N. A. S. Putro, S. W. Prakosa, and J.-S. Leu, “Hybrid ConvTransLSTM for spatio-temporal classification: identifying early Parkinson’s disease from gait patterns,” *Neural Computing and Applications*, vol. 38, no. 7, Art. no. 224, 2026, doi: 10.1007/s00521-025-11829-8.
- [36] X. Wang, X. Xu, Z. Zhao, F. Li, F. Qi, and S. Liang, “VGRF Signal-Based Gait Analysis for Parkinson’s Disease Detection: A Multi-Scale Directed Graph Neural Network Approach,” *IEEE Journal of Biomedical and Health Informatics*, vol. 30, no. 1, pp. 100–112, Jan. 2026, doi: 10.1109/JBHI.2025.3589772.
- [37] D. Wang, Y. Shi, J. Pang, X. Zhu, L. Meng, and D. Ming, “Data-driven subtyping of early Parkinson’s disease via mutual cross-attention fusion of EEG and dual-task gait features,” *npj Parkinson’s Disease*, vol. 12, Art. no. 45, 2026, doi: 10.1038/s41531-026-01258-2.
- [38] M. Ji, R. Ren, Q. Xu, C. Liang, and C. Zhang, “Gait analysis and severity rating of Parkinson’s disease based on pressure and inertial sensors,” *Biomedical Signal Processing and Control*, vol. 120, Part B, Art. no. 110103, Jul. 2026, doi: 10.1016/j.bspc.2026.110103.
- [39] M. Carter and H. Wang, “Leveraging asymmetric gait features and machine learning preferences for Parkinson’s disease diagnosis,” *Biomedical Signal Processing and Control*, vol. 122, Art. no. 110421, Aug. 2026, doi: 10.1016/j.bspc.2026.110421.
- [40] M. Michaud, N. de l’Escalopier, C. Voisard, L. Oudre, and D. Ricard, “Real-world gait study of Parkinson’s disease using wearable sensors: a systematic review,” *Journal of NeuroEngineering and Rehabilitation*, vol. 23, Art. no. 251, 2026, doi: 10.1186/s12984-026-01957-0.
- [41] Adeli et al., “CARE-PD: A Multi-Site Anonymized Clinical Dataset for Parkinson’s Disease Gait Assessment,” in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 38, 2025, doi: 10.52202/085713-4391.
- [42] P. Gorce and J. Jacquier-Bret, “Current Trends in AI Gait Analysis for the Detection and Assessment of Parkinson’s Disease Severity: Systematic Review and Meta-Analysis of Performance Using Logit Transformation,” *Healthcare*, vol. 14, no. 13, p. 1820, 2026, doi: 10.3390/healthcare14131820.