Diagnosing Parkinson's Disease with Wearable Sensor-Based activity Recognition using Deep Learning Models

by

Anel Satkan

B.S., Nazarbayev University (2023)

Submitted to the Department of Data Science
in partial fulfillment of the requirements for the degree of

Master of Science in Data Science

at the

NAZARBAYEV UNIVERSITY

April 2025

Author . . . .
Department of Data Science
April 28, 2025

Certified by . . . .
Jurn Gyu Park
Assistant Professor
Thesis Supervisor

Accepted by . . . .
Elizabeth Arkhangelsky
Dean, School of Engineering and Digital Sciences

Diagnosing Parkinson's Disease with Wearable Sensor-Based activity Recognition using Deep Learning Models

by

Anel Satkan

Submitted to the Department of Data Science
on April 28, 2025, in partial fulfillment of the
requirements for the degree of
Master of Science in Data Science

Abstract

Wearable sensor technology presents significant potential for continuous and objective monitoring of patients with Parkinson's Disease (PD), a progressive neurological disorder. However, advanced deep learning techniques-particularly transformer-based architectures-remain underexplored in the context of PD detection using multi-limb accelerometer data. This work studies the use of transformer models for classifying PD from the PD-BioStampRC21 dataset, which contains tri-axial accelerometer readings from five body locations in 34 patients. Methodology of this work follows an adapted CRISP-DM framework, with data preprocessing including sensor calibration, activity segmentation, and windowing into 2.5-second overlapping segments. Multiple deep learning models - including convolutional neural networks (CNN), long short-term memory (LSTM), and a hybrid CNN-Transformer architecture - are developed, tuned, and evaluated. Results demonstrate that the CNN-Transformer model achieved an accuracy of 94.0%, outperforming both CNN (93.4%) and LSTM (89.6%) baselines in posture-specific PD classification.

The main contributions of this work are: (1) development of a hybrid CNN-Transformer model that eliminates extensive feature engineering while achieving competitive classification accuracy; (2) demonstration of the effectiveness of transformer-based architectures for PD detection using wearable sensor data, emphasizing their suitability for sequential time series analysis; (3) analysis of activity-specific data, revealing higher detection accuracy during static (sitting) compared to dynamic (walking) periods; and (4) evaluation of multi-sensor fusion, confirming that combining data from multiple body locations improves detection performance over single-sensor inputs. These findings support the application of transformer-based models for automated, continuous PD monitoring from wearable sensors, enabling scalable at-home solutions that address the limitations of traditional clinical assessments.

Thesis Supervisor: Jurn Gyu Park
Title: Assistant Professor

Acknowledgments

I would like to express my sincere gratitude to Professor Jurn Gyu Park for his valuable guidance, constructive feedback, and excellent management throughout this work. His encouragement and ability to push me at critical moments greatly contributed to the completion of this research.

Special thanks to my classmate Merey Orazaly, with whom I began this project as a course assignment. Her collaboration and early contribution were crucial in shaping the direction of this work.

Finally, I am deeply thankful to my husband for his support and encouragement during this journey.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Parkinson's disease (PD) is a progressive neurological disorder characterized by tremor, impaired coordination, and balance difficulties [16] that severely reduces patients' quality of life, with no current cure or prevention. Early diagnosis and continuous monitoring are the only ways to improve the long-term outcomes for patients. Traditionally, PD diagnosis and monitoring rely on episodic in-clinic assessments, including the Unified Parkinson's Disease Rating Scale (UPDRS) and patient-reported symptoms [4]. However, these subjective evaluations can miss subtle movement abnormalities, increasing misdiagnosis risk [21]. Advancements in machine learning (ML) and deep learning (DL) technologies are helping to address these challenges, improving PD detection and assessment.

The literature review on the use of ML in the diagnosis and assessment of PD shows that audio-video and image datasets are the most commonly used ones in the area [21]. Researchers rely on handwriting, voice recording, neuroimaging, and gait recording datasets. Although these data modalities achieve high diagnostic accuracy, they are limited to in-clinic visits or controlled environments and cannot provide continuous symptom monitoring in patients' daily lives. Clinical recordings or fixed recording setups at home can potentially miss important symptom fluctuations and medication responses. Furthermore, these modalities raise significant privacy concerns. Voice data contain sensitive personal information, since voiceprints are considered personal identifiers [7]. Similarly, video data used for gait and movement pattern monitoring may seem too invasive for patients and change their everyday behavior [20]. In this regard, wearable sensor technology offers continuous, objective monitoring of PD motor symptoms, especially rest tremor present in 70-75% of cases [4, 13]. Unlike audio-video data, sensor signals are less personally identifiable. This enhances privacy while enabling capture of subtle symptom fluctuations in natural environments.

The application of ML to sensor data has shown promising results for PD detection and monitoring [15]. Although encouraging, traditional ML approaches require extensive feature engineering, a time-consuming process that demands domain expertise to extract relevant features from raw sensor data. This manual feature selection process often becomes a bottleneck in developing automated diagnostic systems and may not capture all relevant patterns in complex movement data. DL models, especially transformers, excel at learning complex patterns from raw sensor time series without manual feature engineering [23, 30]. Transformers outperform convolutional neural network (CNN) by 21–30% in human activity recognition (HAR) tasks [26], thanks to their self-attention mechanism that captures long-range temporal dependencies [19]. This makes them promising for detecting subtle PD-related movement abnormalities.

This work aims to develop a PD detection approach that utilizes the strengths of wearable sensor technology and transformer-based DL models. Using sensor data from the *PD-BioStampRC21* dataset [1], this work develops and evaluates DL architectures - including CNN, long short-term memory (LSTM), and a hybrid CNN-Transformer model-to classify PD based on movement data. The proposed transformer-based architecture processes raw accelerometer signals without extensive feature engineering, potentially streamlining PD monitoring system development. The main objective is to demonstrate that a sensor data-driven, transformer-based approach can achieve diagnostic accuracy comparable to methods using other data modalities and traditional ML models, while offering advantages in terms of privacy preservation, continuous monitoring capability, and reduced preprocessing requirements. Additionally, this work investigates the impact of sensor fusion and activity-specific data segmentation on PD detection accuracy.

This work makes the following contributions:

1. Developed a hybrid CNN-Transformer model that eliminates extensive feature engineering while achieving classification accuracy comparable to traditional ML methods.

2. Demonstrated the effectiveness of transformer-based models for PD detection using wearable sensor data, highlighting their suitability for sequential time series analysis.

3. Analyzed activity-specific data, showing higher PD detection accuracy during sitting compared to walking.

4. Evaluated sensor fusion, confirming that combining multiple sensors improves detection performance over single-sensor inputs.

The rest of the paper is organized as follows. Chapter 2 presents the problem background, motivation, and related work. Chapter 3 describes the methodology following the CRISP-DM framework, covering data understanding, preparation, and modeling. Chapter 4 presents the experimental setup, results, and key observations. Chapter 5 discusses the results, limitations, and potential directions for future research. Chapter 6 concludes the work.

Chapter 2

Motivation and Related Work

2.1 Background

PD is a common neurodegenerative disorder whose diagnosis and monitoring are limited by episodic and subjective clinical assessments [16, 28]. As discussed in the Introduction, recent research has explored objective, technology-driven approaches to overcome these limitations. Among these, wearable inertial sensors have emerged as a promising tool for continuous, privacy-preserving monitoring of motor symptoms in real-world settings [31, 23, 25].

Figure 2-1 summarizes the diversity of data modalities and ML techniques applied in recent PD detection research, as reviewed by Mei et al. [21]. The figure shows that traditional modalities such as voice, handwriting, and medical imaging remain dominant in the literature, with support vector machine (SVM) and neural network (NN) being the most frequently used algorithms across modalities. Although gait and multi-sensor wearable data appear in some studies, they represent a smaller fraction compared to voice and imaging modalities, underscoring the underutilization of wearable sensors in PD research. Furthermore, many existing wearable sensor studies use limited sensor placements or single-sensor setups, potentially missing critical inter-limb coordination and temporal dynamics [11, 23]. ML models applied to sensor data often depend on manual feature engineering, which is time-consuming and may not capture the full complexity of PD motor symptoms [15, 11].

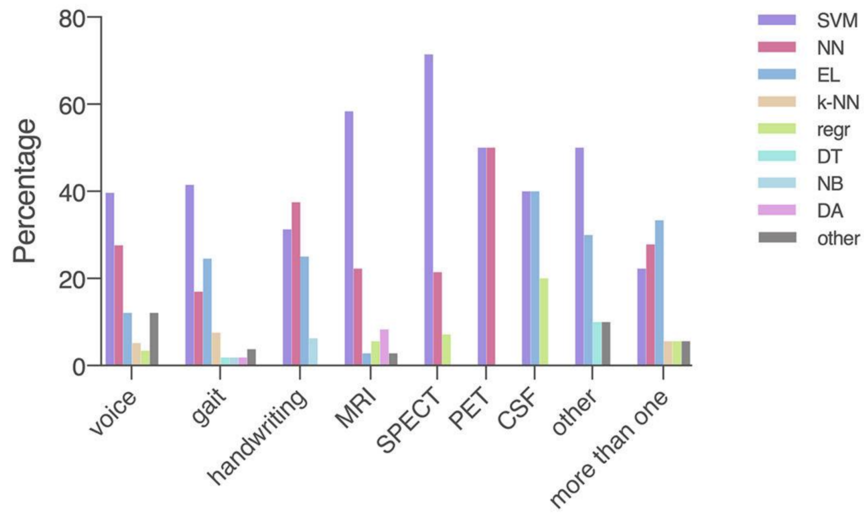

Figure 2-1: Distribution of data modalities and ML techniques used for PD detection. SVM: support vector machine; NN: neural network; EL: ensemble learning; k-NN: k-nearest neighbor; regr: regression; DT: decision tree; NB: Naive Bayes; DA: discriminant analysis. Adapted from [21].

Recent reviews [29, 14, 21] highlight the need for scalable, automated analysis methods that leverage multi-sensor data to improve diagnostic accuracy and clinical utility. Furthermore, patient adherence and usability remain important considerations for wearable technology adoption [30, 25]. These challenges in sensor-based PD monitoring motivate the exploration of DL techniques that can extract meaningful patterns from complex multi-sensor data while minimizing manual intervention.

2.2 Motivation

To address these gaps, this work explores the use of transformer-based DL models on multi-sensor accelerometer data for PD detection. The *PD-BioStampRC21* dataset [1] offers a unique opportunity to investigate inter-limb coordination and activity-specific symptom expression using synchronized data from five body locations [2]. To this end, the study is driven by the following key research questions:

1. Can transformer-based models eliminate the need for extensive feature engineering while achieving diagnostic accuracy comparable to or exceeding traditional ML methods on wearable sensor data?

2. Does segmenting data by activity type (e.g., sitting vs. walking) improve PD detection performance, and which activities provide the most discriminative information?

3. How does multi-sensor fusion impact detection accuracy compared to single-sensor approaches?

Figure 2-2: Multi-sensor wearable setup for PD monitoring, with sensors placed on arms, legs, and chest. Adapted from [11, 3].

Figure 2-2 depicts a multi-sensor configuration used in the *PD-BioStampRC21* dataset [1], illustrating the potential to capture complex motor patterns across multiple body segments. Using transformer models on such data can exploit long-range temporal dependencies and inter-sensor relationships, potentially improving PD detection robustness and accuracy [26, 9].

Recent research demonstrates that transformer-based models significantly outperform traditional ML algorithms when applied to sequential accelerometer data for HAR and related time-series tasks. For example, Saidani et al. [26] showed that transformer architectures achieved 21–30% higher accuracy than CNN on HAR tasks using wearable accelerometer data. Accelerometer signals alone achieve high HAR accuracy (F1-score 94.07% subject-dependent, 83.16% subject-independent), with only marginal improvement from adding biosignals like ECG and PPG [5]. These findings underscore that transformers, combined with accelerometer data, provide a practical, efficient, and highly accurate approach to PD detection.

2.3 Related Work

2.3.1 Wearable Sensors

Wearable sensors are highly suitable for PD research due to their ability to provide continuous, objective monitoring of motor symptoms in real-life settings. Early studies, such as Weiss et al. [31], demonstrated the feasibility of at-home, body-worn accelerometers for mobility assessment. Rovini et al. [25] systematically reviewed the application fields of wearable devices in PD, including early diagnosis, tremor, body motion, motor fluctuations, and long-term home monitoring. Recent reviews confirm the growing clinical utility of inertial sensors, noting their ability to detect both motor and non-motor symptoms with high accuracy, especially when combined with IoT and machine learning technologies [27]. Adams et al. [3] demonstrated the feasibility of using accelerometer-based sensors attached to the chest and limbs for both standardized in-clinic assessments and at-home monitoring, achieving high data completeness (99.3%). Their follow-up study [2] expanded this approach by employing five wearable sensors on limbs and trunk, revealing significant differences in activity and tremor patterns between PD patients and controls. Similarly, Dinesh et al. [8] introduced the use of lightweight, tattoo-like sensors to enable unobtrusive, long-term motion measurement, highlighting the potential for continuous symptom tracking. Lonini et al. [18] further investigated sensor placement, finding that a single sensor on the back of the hand sufficed for detecting upper extremity bradykinesia and tremor, emphasizing the importance of sensor location. In contrast to single sensor use, Moreau et al. [23] provides a comprehensive overview, underscoring the promise of multi-sensor fusion to capture complex motor manifestations of PD, which remains underexploited.

While sensor technology provides the hardware foundation for continuous monitoring, their effectiveness also depends on used computational methods.

2.3.2 ML and DL Approaches

ML techniques have been widely applied to wearable sensor data for PD detection. Igene et al. [15] developed an SVM-based model using features extracted from accelerometer data collected from multiple body locations, achieving up to 94.4% accuracy. Hathaliya et al. [12] proposed a CNN model focusing on tremor intensity classification, outperforming traditional ML with 92.4% accuracy. Moon et al. [22] compared various ML methods for differentiating PD from essential tremor using gait and balance features, finding NN yielded the highest F1-score. Li et al. [17] explored frequency-domain features from gait data, suggesting their effectiveness for PD diagnostics using logistic regression and SVM classifiers. These studies highlight the potential of ML but also the dependency on handcrafted features.

DL models, especially those that can learn directly from raw sensor data, have gained traction. Saidani et al. [26] demonstrated that transformer-based models significantly outperform CNNs in HAR tasks using accelerometer data, with 21–30% higher accuracy. Luptáková et al. [19] adapted transformers for smartphone motion sensor data, achieving over 99% accuracy, matching and exceeding CNN-LSTM baselines. Djenouri et al. [9] combined convolutional and transformer architectures, improving HAR classification accuracy from 88% (CNN alone) to 92%. These advances suggest that transformer architectures are well-suited for modeling the complex temporal dynamics of human motion, and potentially PD symptoms.

To maximize the utility of sensor data, signal processing techniques and posture based analysis must be carefully paired with ML/DL architectures.

2.3.3 Signal Processing and Posture Analysis

Signal preprocessing and activity context significantly influence PD detection accuracy. Gil-Martin et al. [11] showed that applying Fast Fourier Transform (FFT) magnitude coefficients to inertial signals improved PD detection accuracy, especially when fusing data from multiple sensors. Integrating information from all five sensors boosted performance from an accuracy of 66.90% to 75.10% compared to single sensor performance. Their subsequent study [10] revealed that posture-specific analysis, particularly focusing on sitting activities, enhanced tremor detection compared to lying or walking, suggesting that symptom expression varies with activity. Similarly, Bächlin et al. [6] developed a wearable system for real-time detection of freezing of gait in PD, using posture-specific features to improve detection accuracy. Another work by Patel et al. [24] demonstrated that posture-specific segmentation improves the sensitivity of motor symptom detection in PD. These findings emphasize the importance of activity segmentation and tailored signal processing in wearable sensor analysis for PD.

2.3.4 Transformer Models

Transformer models have shown promise in time-series analysis for HAR, suggesting potential applications in PD detection. Although developed for natural language processing tasks, the ability of transformers to capture long-range dependencies through self-attention mechanisms have revolutionized sequential sensor data analysis. Luptáková et al. [19] pioneered the application of transformers to wearable sensor data for HAR, achieving state-of-the-art accuracy. Saidani et al. [26] further validated transformers' superiority over CNN in HAR tasks across multiple public datasets. Djenouri et al. [9] demonstrated that hybrid CNN-transformer models outperform pure CNN by effectively combining local feature extraction and global temporal modeling. Despite these advances, transformer models have not been extensively applied to PD-specific multi-sensor datasets. This research addresses this gap by leveraging transformers to exploit inter-limb coordination and temporal dependencies in the *PD-BioStampRC21* dataset [1].

2.3.5 Summary

Existing research demonstrates the potential of wearable sensors for PD monitoring, but robust frameworks for automated feature learning and multi-sensor coordination analysis are still lacking. The *PD-BioStampRC21* dataset's [1] underutilized multi-limb recordings provide an ideal testbed for transformer-based models to address these gaps. Table 2.1 provides a comparative overview of key studies that have applied machine learning and DL methods to the same dataset or closely related sensor data. The table summarizes the analytical methods, windowing strategies, and the extent of posture-specific analysis performed in each work. It highlights the shift from traditional ML approaches and single-sensor analysis to recent advances in DL, multi-sensor fusion, and transformer-based models. This comparison underscores the unique contribution of the present study, which is the first to evaluate transformers for PD detection across multiple postures using multi-sensor data.

Table 2.1: Comparison of Works

Reference	**Methods**	**Dataset**	**Windows**	**Posture analysis**	**Key contribution**
Gil-Martin et al.[11]	FFT+CNN	BioStampRC	6.4s, overlap	None	Multi-sensor fusion and DL for PD detection
Igene et al.[15]	PCA+SVM	BioStampRC	2s, overlap	None	Multi-sensor fusion and ML for PD detection
Hathaliya et al.[12]	CNN	BioStampRC	6.4s, overlap	None	One sensor and DL for PD detection
Luptáková et al.[19]	Transformer	Smartphone sensor data	3s	Analysis of motion signals	First transformer adaptation for motion analysis
Gil-Martin et al.[10]	FFT+CNN	BioStampRC	3.2s, overlap	Posture-specific analysis	Study which activity is better to detect the tremor
This work	CNN, LSTM, Transformer	BioStampRC	2.5s, overlap	Posture-specific analysis	Transformer, comparison across postures

Chapter 3

Methodology

This work follows an adapted CRISP-DM methodology, as illustrated in Figure 3-1. The Problem Understanding stage, outlined in Section 2.1, is addressed in previous chapters. This chapter focuses on the Data Understanding, Data Preparation, and Modeling stages, directly addressing the research questions posed in Chapter 2.

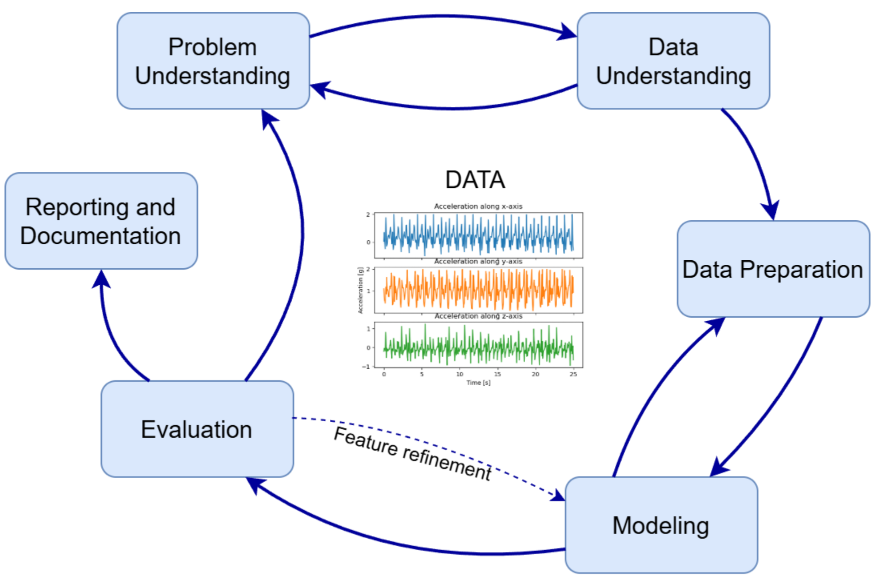

Figure 3-1: Adapted CRISP-DM methodology

3.1 Data Understanding

3.1.1 Collect and Describe Data

This study uses the *PD-BioStampRC21* dataset [1], which contains tri-axial accelerometer readings from five wearable sensors attached to participants' bodies (chest, left leg, right leg, left hand, right hand). Each sensor recorded acceleration along three axes (x, y, z) at 31.25 Hz. Data were collected over two consecutive days per participant, yielding approximately 5.2 million readings per sensor per subject.

The dataset includes 34 participants - 17 PD, 17 healthy controls (HC), with demographic and clinical information, including UPDRS scores for PD patients. The first hour of data was collected in-clinic using standardized UPDRS motor tasks, including walking tests, balance assessments, and hand tremor evaluations; the remainder was during daily life. The dataset (publicly available on Kaggle) consists of raw CSV accelerometer files and clinical annotations.

The dataset comprises synchronized time-series from five sensors, resulting in a 15-dimensional signal (3 axes $\times$ 5 sensors). Each row corresponds to a timestamped reading from all sensors, with accelerometer values reported in units of g ($9.8\ m/s^2$). Figure 3-2 summarizes participant demographics.

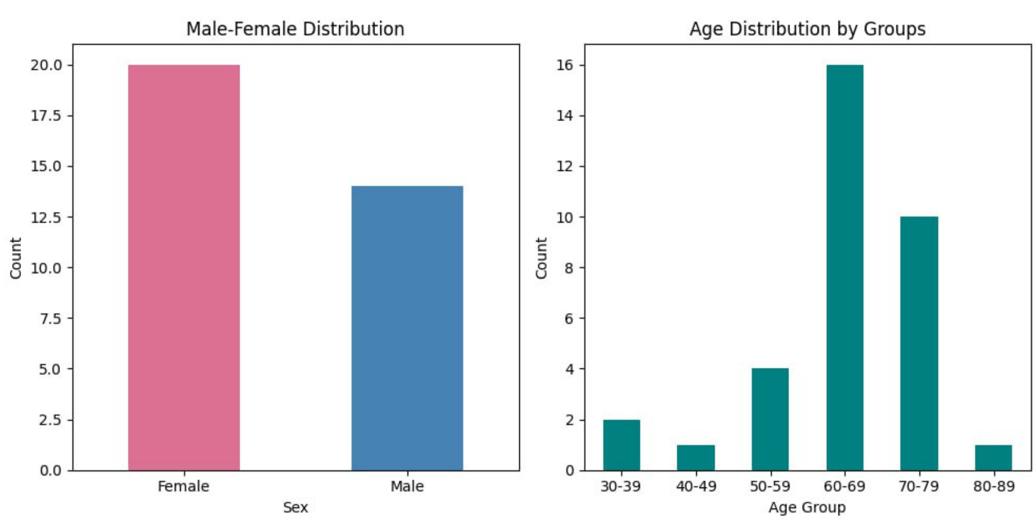

Figure 3-2: Dataset Demographics

With the dataset structure established, the next step is to explore the signal characteristics and behavioral patterns that distinguish PD from HC.

3.1.2 Explore Data

Initial analysis reveals distinct patterns in accelerometer signals between PD and HC.

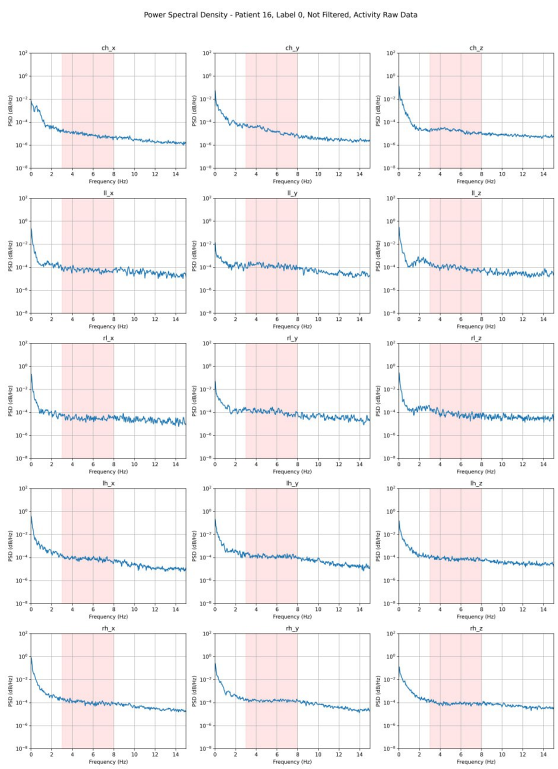

Figure 3-3: Healthy Patient PSD

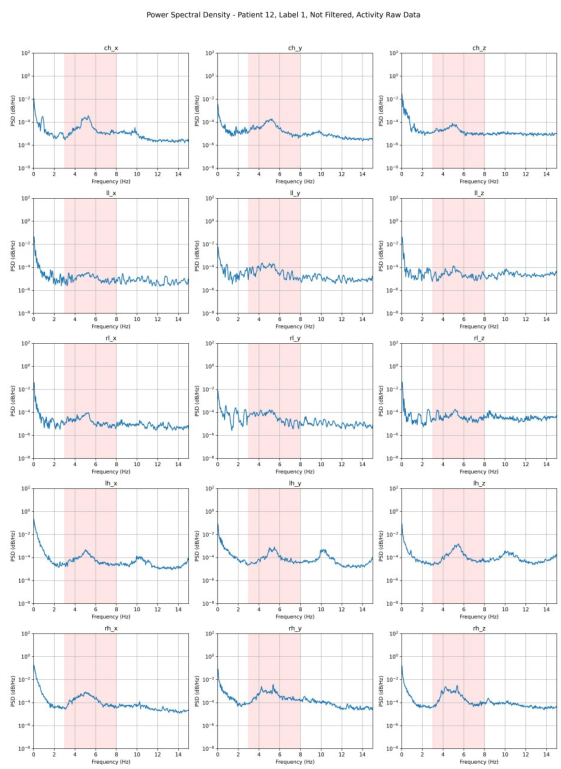

Figure 3-4: PD Patient PSD

Power Spectral Density (PSD) analysis quantifies how the power of a signal is distributed across different frequency components, providing insight into the dominant oscillatory patterns within the data. PSD analysis in Figures 3-3 and 3-4 highlights differences in the tremor frequency band (3–8 Hz, highlighted in red) between groups. In these PSD plots, the **y-axis** represents logarithmic power density, measured in decibels per Hertz (dB/Hz), indicating how much signal power is present at each frequency interval. The **x-axis** denotes frequency in Hertz (Hz), corresponding to cycles per second. This format allows for clear visualization of how signal power is distributed across different frequencies.

PD patients show strong peaks in the 3–8 Hz band, especially in hand sensors. These peaks correspond to classic parkinsonian rest tremor (typically 4–6 Hz). In contrast, HCs display a more uniform power distribution and symmetry between sides. The chest sensor also captures rhythmic tremor, though with lower amplitude.

Motor symptom expression and tremor characteristics can vary significantly depending on the patient's posture and activity context. For activity classification, the dataset was trimmed to about 2.7 million rows per participant (one day of monitoring). Activity recognition was performed using a rule-based algorithm [3, 2], classifying postures (lying, upright, sitting, stand/sit, upside-down) and further distinguishing standing from walking via autocorrelation analysis. Sensor orientation mapping is detailed in [3].

Activity classification shown in Figure 3-5 revealed that HCs spent significantly more time walking (about 15,000 s/day) than PD patients (about 12,000 s/day), while PD patients spent more time sitting and in stand/sit transitions. These behavioral differences align with clinical expectations and previous descriptions [3].

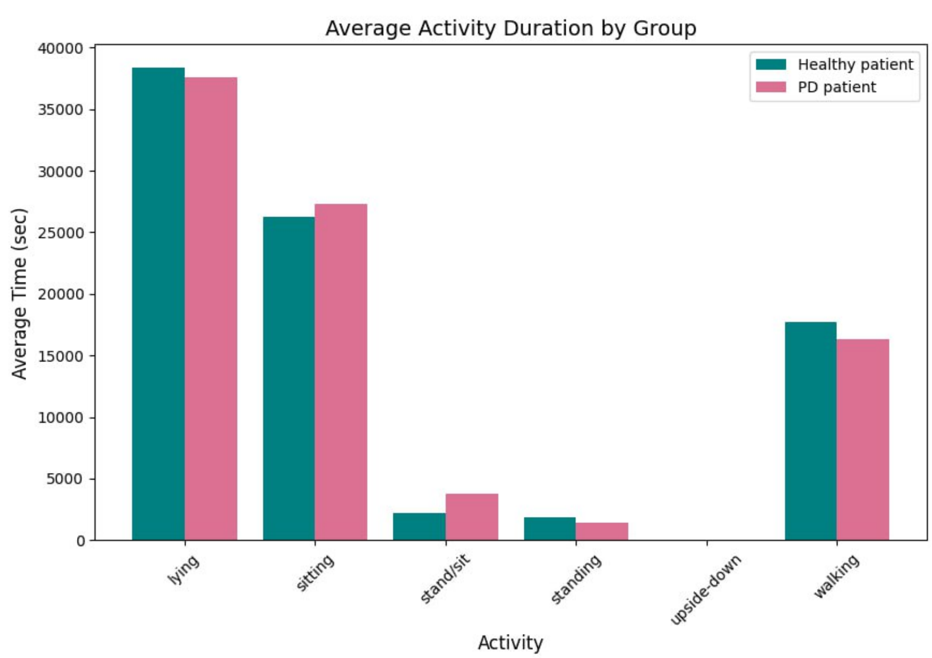

Figure 3-5: Activity Durations

3.1.3 Verify Data Quality

Some participants lacked data from all five sensors. Patients 60, 7, and 14 were excluded, resulting in a final cohort of 14 HC and 17 PD patients for analysis. The time-series data are otherwise complete, with no missing values. Minor parsing issues were corrected manually. Calibration was verified using 'balance test' intervals (patients standing upright), aligning sensor orientation with reference vectors [3]. These intervals are included in the *Annotation* file for each patient. If misalignment exceeded a threshold, the Rodrigues' rotation matrix (Equation 3.2) was applied to align the sensor orientation vectors.

$$\mathbf{u} = \frac{\mathbf{a}_c \times \mathbf{r}_c}{\|\mathbf{a}_c \times \mathbf{r}_c\|}, \qquad \theta = \cos^{-1}(\mathbf{a}_c \cdot \mathbf{r}_c) \tag{3.1}$$

$$R = \begin{bmatrix} \cos\theta + u_x^2(1-\cos\theta) & u_x u_y(1-\cos\theta) - u_z \sin\theta & u_x u_z(1-\cos\theta) + u_y \sin\theta \\ u_y u_x(1-\cos\theta) + u_z \sin\theta & \cos\theta + u_y^2(1-\cos\theta) & u_y u_z(1-\cos\theta) - u_x \sin\theta \\ u_z u_x(1-\cos\theta) - u_y \sin\theta & u_z u_y(1-\cos\theta) + u_x \sin\theta & \cos\theta + u_z^2(1-\cos\theta) \end{bmatrix} \tag{3.2}$$

According to the dataset card, timestamps are consistently reported starting from the instant of the earliest sensor recording [1]. Timestamps are synchronized across sensors, starting from the latest attached sensor. Biological plausibility was checked: as shown in Figure 3-6, the autocorrelation of the chest sensor displays clear periodicity during walking and relative stability during sitting, consistent with expected movement patterns.

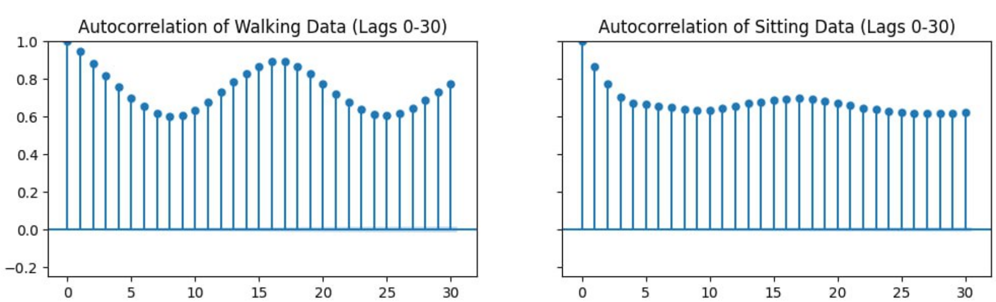

Figure 3-6: Autocorrelation of Walking and Sitting

3.2 Data Preparation

This study evaluates two sensor configurations for PD detection: (1) all five accelerometers, and (2) only the left hand sensor, following prior work on upper limb tremor detection [12, 18]. After calibration and timestamp alignment, data were segmented by activity into static (sitting) and walking periods.

To examine the impact of signal preprocessing, four dataset variants were created: dynamic filtered, dynamic raw, static filtered, and static raw. Here, dynamic refers to data segments recorded during walking activities, which involve continuous movement and more complex motor patterns. In contrast, static refers to data segments recorded during sitting activities, characterized by relatively minimal movement and more stable postures.

For each activity type, 30,000 rows of sensor data per participant were concatenated to form the respective datasets. Each dataset thus contains approximately 900,000 rows and 17 columns, consisting of 15 accelerometer sensor readings (from five sensors, each with three axes), along with patient ID and label columns. In the case of datasets using only the left-hand sensor, the data include three axes of accelerometer readings, patient ID, and label.

For DL, the time series was segmented into fixed-length windows. Window size and overlap ratio were tuned (via random search and refinement), with best results at 80 frames (2.5 s at 31.25 Hz) and 0.75 overlap. Patient-specific splitting was used to prevent data leakage.

Filtered datasets applied a fourth-order Butterworth bandpass (3–8 Hz) to each channel, isolating the tremor band and reducing artifacts [3, 22]. Raw datasets retained only calibrated, segmented signals. This preparation allows evaluation of model performance across sensor setups, activities, and preprocessing pipelines.

3.2.1 Pipeline Comparison

Figures 3-7 and 3-8 illustrate the data preparation pipelines for DL and traditional ML approaches, respectively.

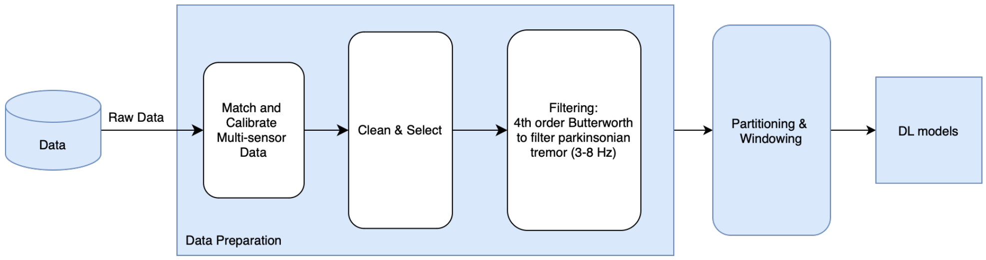

Figure 3-7: DL Data Preparation Pipeline

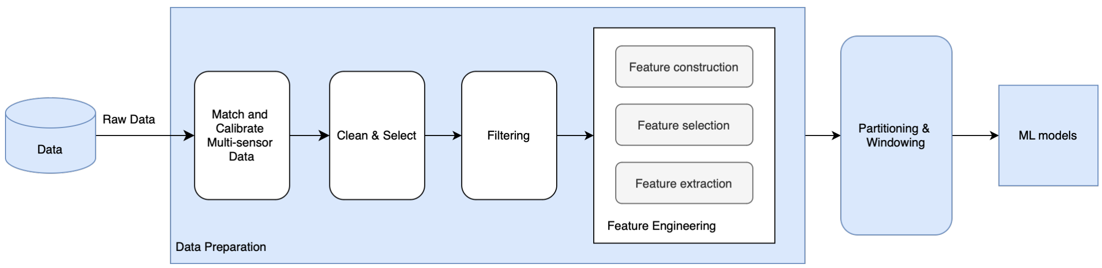

Figure 3-8: ML Data Preparation Pipeline

The ML pipeline (Figure 3-8) reflects the standard methodology widely used in the PD detection literature. In this approach, after acquiring and segmenting raw accelerometer signals, each window undergoes manual feature extraction. A set of time-domain and frequency-domain features-such as mean, variance, skewness, kurtosis, FFT coefficients, and dominant frequency-are computed for each window. These features are then used as input to a traditional ML classifier. While this pipeline has demonstrated success in prior studies, it is inherently dependent on domain expertise for feature selection and may fail to capture subtle or complex signal characteristics present in raw sensor data. As discussed in Chapter 2, this reliance on handcrafted features is a key limitation of conventional ML approaches.

In contrast, this work adopts a DL pipeline (Figure 3-7) that eliminates the need for extensive feature engineering. Here, synchronized and calibrated accelerometer data are segmented by activity type and divided into fixed-length, overlapping windows. For filtered dataset variants, a Butterworth bandpass filter (3–8 Hz) is applied to each channel to isolate the tremor frequency band and reduce noise. Each window is normalized and reshaped into a three-dimensional tensor, which is then provided directly as input to the DL models-CNN, LSTM, or CNN-Transformer. These models are capable of learning complex, discriminative features automatically from raw or minimally processed data, enabling a fully data-driven approach to PD detection.

By prioritizing the DL pipeline, this study directly addresses the first research question posed in Chapter 2 regarding the elimination of manual feature engineering. The effectiveness and advantages of this data-driven strategy are further explored in the following modeling section, where the selected DL architectures are described and evaluated in detail.

3.3 Modeling

DL models are prioritized in this work due to their capacity to extract complex features directly from raw sensor data, thus eliminating the need for extensive manual feature engineering. This methodological choice directly addresses the first research question outlined in Chapter 2 and is motivated by the limitations of traditional ML approaches discussed in the literature review. Based on best practices and identified research gaps, three DL architectures were selected for evaluation: CNN, LSTM, and a hybrid CNN-Transformer model.

For all architectures, hyperparameters were tuned using random search followed by targeted refinement, ensuring optimal model performance for the *PD-BioStampRC21* dataset [1]. The following subsections detail each model and its optimal configuration.

3.3.1 CNN

The CNN model was chosen for its proven effectiveness in PD detection tasks using wearable sensor data, as demonstrated in recent studies employing the *PD-BioStampRC21* dataset [12, 11, 10]. CNNs are well-suited for this dataset because they can automatically extract discriminative spatial and temporal features from raw accelerometer signals, reducing reliance on handcrafted features. Prior work has shown that CNNs can achieve high accuracy, especially when leveraging multi-sensor fusion and carefully selected windowing strategies [11]. In this study, the optimal CNN model was configured to receive fixed-length windows of 80 frames (2.5 seconds) with a 75% overlap, capturing sufficient temporal context to detect tremor and movement abnormalities. The architecture consists of two convolutional layers - the first with 96 filters and a kernel size of 7, second with 80 filters and a kernel size of 3 - followed by a fully connected layer with 96 units. Batch normalization was not applied, and dropout was set at 0.4 to prevent overfitting. The model was trained with a learning rate of approximately 0.001, as determined by hyperparameter optimization. The full hyperparameter search space and best values for the CNN are summarized in Table 3.1.

Table 3.1: CNN Hyperparameter Search Space and Best Values

Hyperparameter	**Search Space**	**Best Value**
Window Size	[32, 64, 80, 96]	80
Overlap Ratio	[0.25, 0.50, 0.75, 0.80]	0.75
Filters (1st Layer)	[32, 64, 96, 128]	96
Kernel Size (1st Layer)	[3, 5, 7]	7
BatchNorm (1st Layer)	{True, False}	False
Add Second Conv Layer	{True, False}	True
Filters (2nd Layer)	[32, 64, 80, 128]	80
Kernel Size (2nd Layer)	[3, 5, 7]	3
BatchNorm (2nd Layer)	{True, False}	False
Dense Units	[64, 128, 256]	96
Use Dropout	{True, False}	True
Dropout Rate	Uniform(0.1, 0.5)	0.4
Learning Rate	LogUniform(1e-4, 1e-2)	0.00099

3.3.2 LSTM

The LSTM model was selected to capture long-term temporal dependencies in the sequential accelerometer data, which are critical for identifying movement abnormalities that may only become apparent over several seconds. LSTM networks are particularly effective for modeling time series in which tremor characteristics and movement patterns evolve over time. In this configuration, the input consists of fixed-length windows of 80 frames (2.5 seconds) with a 75% overlap. The architecture includes two LSTM layers: the first with 112 units and batch normalization to stabilize training, and the second with 128 units without batch normalization. The first LSTM layer is set to return only the final output, focusing on the overall temporal summary of each window. A fully connected layer with 96 units integrates the learned temporal features, and dropout is applied at a rate of 0.3 to reduce overfitting. The model is trained with a learning rate of 0.00116. The hyperparameter search space and optimal values for the LSTM model are detailed in Table 3.2. This architecture balances model complexity and regularization, enabling effective learning from sequential sensor data.

Table 3.2: LSTM Hyperparameter Search Space and Best Values

Hyperparameter	**Search Space**	**Best Value**
Window Size	[32, 64, 80, 96]	80
Overlap Ratio	[0.25, 0.5, 0.75]	0.75
LSTM Units (1st Layer)	[32, 64, 96, 112, 128]	112
Return Sequences (1st)	{True, False}	False
BatchNorm (1st Layer)	{True, False}	True
Add Second LSTM Layer	{True, False}	True
LSTM Units (2nd Layer)	[32, 64, 96, 128]	128
BatchNorm (2nd Layer)	{True, False}	False
Dense Units	[64, 96, 128, 256]	96
Use Dropout	{True, False}	True
Dropout Rate	Uniform(0.1, 0.5)	0.3
Learning Rate	LogUniform(1e-4, 1e-2)	0.00116

3.3.3 CNN-Transformer

The hybrid CNN-Transformer model was developed to combine the strengths of both convolutional and transformer architectures. While transformers have shown superior performance in HAR and other time-series tasks, their application to PD detection with wearable accelerometer data remains limited. Initial experiments with a pure transformer model on the dataset [1] yielded suboptimal results, suggesting that transformers alone may not effectively extract low-level movement features from moderate-sized datasets. To address this, the hybrid model combines local feature extraction via convolutional layers with global context modeling through transformer encoder blocks, as recommended in recent literature [9]. The model receives input windows of 112 frames ($\sim$3.6 s) with 75% overlap. It begins with two Conv1D layers:

the first with 112 filters and a kernel size of 7; the second with 80 filters and a kernel size of 3, without batch normalization. This is followed by two transformer encoder blocks, each with a head size of 64, two attention heads, a feed-forward dimension of 64, and a dropout rate of 0.2. A dense layer with 128 units integrates the extracted features. The model was trained with a learning rate of 0.00013. Table 3.3 summarizes the hyperparameter search space and optimal values for the hybrid model. This architecture is designed to capture both local and global temporal dependencies, enabling comprehensive analysis of multi-sensor time-series data for PD detection.

Table 3.3: CNN-Transformer Hyperparameter Search Space and Best Values

Hyperparameter	**Search Space**	**Best Value**
Window Size	[32, 48, 64, 80, 96, 112, 128]	112
Overlap Ratio	[0.25, 0.5, 0.75, 0.80]	0.75
Conv1D Filters (1st Layer)	[16, 32, 48, 64, 80, 96, 112, 128]	112
Kernel Size (1st Layer)	[3, 5, 7]	7
BatchNorm (1st Layer)	{True, False}	False
Add Second Conv1D Layer	{True, False}	True
Conv1D Filters (2nd Layer)	[16, 32, 48, 64, 80, 96, 112, 128]	80
Kernel Size (2nd Layer)	[3, 5]	3
BatchNorm (2nd Layer)	{True, False}	False
Number of Transformer Blocks	[1, 2]	2
Transformer Head Size	[16, 32, 64]	64
Transformer Num Heads	[2, 4, 8]	2
Transformer FF Dim	[32, 64, 128]	64
Transformer Dropout	[0.1, 0.2, 0.3, 0.4, 0.5]	0.2
Dense Units	[32, 48, 64, 80, 96, 112, 128]	112
Use Dropout	{True, False}	False
Dropout Rate	[0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8]	0.6
Learning Rate	LogUniform(1e-4, 1e-2)	0.00013

In summary, the modeling approach in this work is designed to systematically compare the effectiveness of three state-of-the-art DL architectures for PD detection using multi-sensor wearable data. By eliminating manual feature engineering and utilizing advanced model architectures, this study aims to demonstrate the advantages of a fully data-driven approach, as motivated by the research questions and literature review in the preceding chapters.

Chapter 4

Results and Analysis

4.1 Experimental Setup

The experiments in this study utilized the publicly available PD-BioStampRC21 dataset [1], which consists of tri-axial accelerometer recordings from 30 participants. The dataset was segmented by activity type, focusing on two primary conditions: **static** (sitting) and **dynamic** (walking) periods.

For DL model training and evaluation, the continuous sensor data were partitioned into fixed-length overlapping windows. Specifically, the CNN and LSTM models were trained on a total of 44,997 windows, which were split into 35,997 for training, 4,500 for validation, and 4,500 for testing. In contrast, the transformer-based models utilized 32,139 windows in total, divided into 25,711 for training, 3,214 for validation, and 3,214 for testing. This difference in window counts reflects model-specific preprocessing and segmentation strategies optimized during hyperparameter tuning.

Each model was trained for up to **150 epochs** with early stopping based on validation loss to prevent overfitting. Training progress was monitored through loss and accuracy curves recorded for both training and validation sets.

Model performance was evaluated on the held-out test set using classification accuracy and confusion matrices, providing insights into overall effectiveness and class-specific prediction quality.

The computational environment for all experiments is summarized in Table 4.1.

Table 4.1: Hardware and Software Setup for Experiments

Hardware	
CPU	Intel Core i5-9600K, 4.29 GHz
GPU	Nvidia GeForce RTX 2070, 1410 MHz, 8 GB
RAM	16 GB, 2667 MHz
Operating System	Windows 10
Software	
Programming Language	Python 3.11
DL Framework	TensorFlow
IDE / Environment	Jupyter Notebook
Additional Libraries	NumPy, Pandas, Scikit-learn

With the experimental design and computational environment established, the following sections present a comprehensive analysis of model performance across multiple dataset variants. The results are organized to systematically address the key research questions posed in Chapter 2, including the impact of data preprocessing, the effectiveness of different DL architectures, and the role of sensor fusion and activity-specific segmentation in PD detection.

4.2 Data Preparation related Results

This section examines the influence of signal preprocessing and activity segmentation on model performance. As shown in Table 4.2, filtering the sensor data generally improved classification accuracy for the LSTM and CNN-Transformer models across both static and dynamic datasets. This indicates that isolating the tremor frequency band via filtering enhances the models' ability to detect PD motor symptoms.

Interestingly, the CNN model exhibited a different trend on the dynamic dataset, achieving higher accuracy on raw data compared to filtered data. This suggests that CNNs may extract useful features from the broader frequency spectrum during walking activities, where movement patterns are more complex.

One of the primary research questions addressed in this work was whether posture-specific analysis improves PD detection. The rationale is that walking data are inherently rhythmic and noisy, potentially obscuring subtle tremor signals, whereas sitting data provide a more stable window for tremor detection. Findings of this work support this hypothesis: static datasets consistently yielded higher accuracies than dynamic datasets, with improvements of up to 15%. These results align with prior work by Gil-Martin et al. [10], reinforcing the importance of activity segmentation in wearable sensor analysis for PD.

In addition to accuracy gains, static datasets reduced training time across all models, indicating improved computational efficiency. These observations highlight the dual benefit of thoughtful data preprocessing: enhancing both model performance and training resource utilization, which is critical for practical deployment in real-world monitoring systems.

Table 4.2: Comparison of CNN, LSTM, and Transformer Results

Model	**Dataset**	**Accuracy**	**Training Time (s)**
CNN	dynamic_raw	89.3%	176.41
	dynamic_filtered	85.2%	94.24
	static_raw	92.5%	83.90
	static_filtered	93.4%	87.40
LSTM	dynamic_raw	73.3%	1823.33
	dynamic_filtered	74.6%	1663.44
	static_raw	87.0%	1208.05
	static_filtered	89.6%	601.35
CNN + Transformer	dynamic_raw	85.8%	464.45
	dynamic_filtered	86.5%	444.20
	static_raw	92.1%	204.80
	static_filtered	94.0%	344.44

4.3 Model Assessment

To evaluate model learning dynamics and convergence behavior, training histories were analyzed for each architecture on the static filtered dataset, which yielded the best overall performance.

4.3.1 Static Filtered Dataset Results

This section evaluates the learning dynamics, convergence behavior, and overall performance characteristics of the CNN, LSTM, and CNN-Transformer models on the **static filtered dataset**, which demonstrated the best overall results. The analysis of training and validation curves provides insight into each model's ability to learn discriminative features relevant to PD detection from wearable sensor data.

CNN Model

As shown in Figure 4-1, the CNN model exhibits strong and efficient learning on the static filtered dataset. The training accuracy rapidly increases, reaching approximately 80% validation accuracy within the first 5 epochs. The model continues to improve steadily, achieving a final training accuracy of 97% and a validation accuracy stabilizing around 94%. Early stopping was triggered at epoch 29, effectively preventing overfitting.

The loss curves mirror this pattern, with a rapid initial decrease followed by steady improvement. The final gap between training loss ($\sim$0.07) and validation loss ($\sim$0.18) suggests some degree of overfitting, but the validation loss stabilizes after epoch 20, indicating a good balance between fitting the training data and generalizing to unseen examples.

This behavior aligns with the first research question (RQ1) concerning the ability of DL models to learn from raw or minimally processed sensor data without extensive feature engineering, demonstrating that CNN can effectively extract relevant features for PD detection.

LSTM Model

Figure 4-2 presents the LSTM model's training history, which reveals a slower and less stable convergence compared to the CNN. Validation accuracy starts near 60% and gradually improves to approximately 87% by epoch 14. The wider gap between training (92%) and validation accuracy (87%) indicates a higher risk of overfitting.

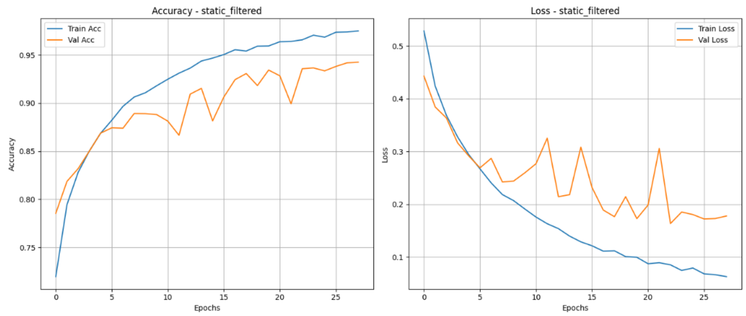

Figure 4-1: CNN Training on Static Filtered Dataset

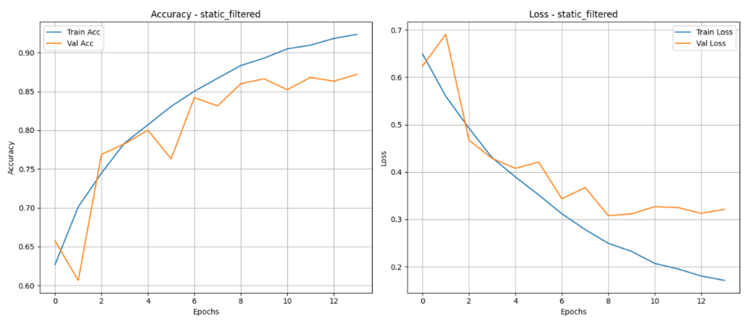

Figure 4-2: LSTM Training on Static Filtered Dataset

The loss curves reinforce this observation, with training loss (0.18) lower than validation loss (0.32), suggesting that the LSTM model struggles more to generalize. Despite this, the model achieves acceptable performance, though it does not match the accuracy or stability of the CNN.

These findings highlight the challenges of modeling long-term temporal dependencies in PD sensor data and suggest that while LSTMs are capable, alternative architectures may offer advantages.

CNN-Transformer Model

The hybrid CNN-Transformer model (Figure 4-3) combines convolutional feature extraction with transformer-based temporal modeling. It demonstrates rapid initial improvement similar to the CNN but achieves even higher final accuracy. Validation accuracy reaches approximately 80% within the first 5 epochs and steadily improves to stabilize around 94–95% by epoch 20.

The training and validation loss curves closely resemble those of the CNN, with a

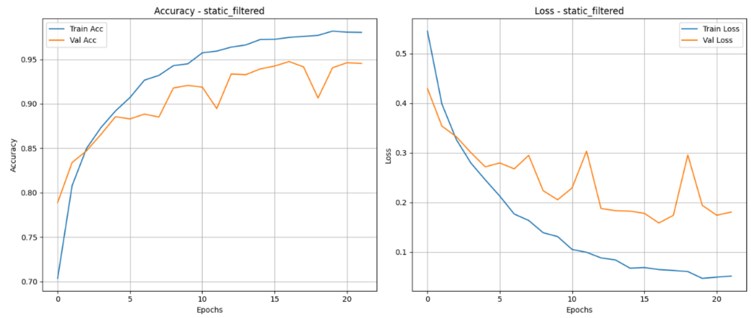

Figure 4-3: CNN-Transformer Training on Static Filtered Dataset

small final gap, indicating strong generalization. Notably, the CNN-Transformer exhibits smoother validation accuracy with fewer fluctuations after epoch 10, suggesting that the transformer blocks help capture more robust temporal representations.

This superior performance addresses RQ1 and supports the thesis contribution of demonstrating the effectiveness of transformer-based architectures for PD detection using wearable sensor data.

4.3.2 Dynamic Filtered Dataset Results

Training curves for the dynamic filtered dataset (Figures 4-4, 4-5, 4-6) show similar overall patterns but with reduced stability and lower validation accuracies compared to static data. This confirms the increased difficulty of tremor detection during walking, where rhythmic and noisy signals complicate feature extraction.

Among the models, the CNN-Transformer consistently achieves the best generalization and stability, outperforming CNN and LSTM by a larger margin on dynamic data. This underlines the importance of advanced architectures capable of modeling complex temporal and spatial dependencies in real-world movement.

These results directly support RQ2 regarding the impact of activity-specific segmentation and reinforce the motivation to focus on posture-specific data to improve PD detection accuracy.

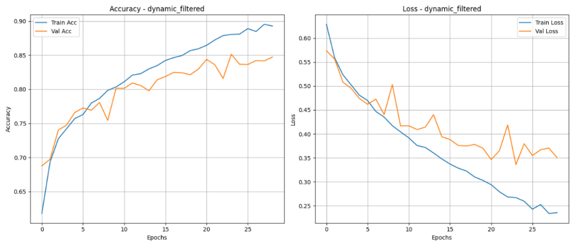

Figure 4-4: CNN on Dynamic Filtered

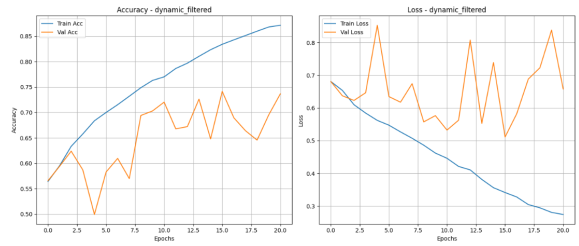

Figure 4-5: LSTM on Dynamic Filtered

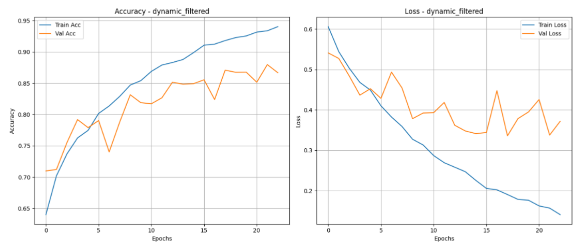

Figure 4-6: CNN-Transformer on Dynamic Filtered

4.4 Classification Performance and Sensor Fusion

To complement accuracy metrics, confusion matrices were constructed for each model on the best-performing static filtered dataset (Table 4.3). The CNN-Transformer achieves the best balance between true positives and true negatives, with the lowest false positive and false negative rates. This balance is critical in clinical settings, where misdiagnosis can have serious consequences: false negatives delay treatment for PD patients, while false positives may lead to unnecessary interventions for healthy individuals.

The CNN and LSTM models exhibit higher misclassification rates, particularly the LSTM with a notable number of false negatives, indicating a higher risk of missed PD cases.

Table 4.3: Confusion Matrix Summary for Static_Filtered Dataset

Model	**TN**	**FP**	**FN**	**TP**
CNN	3818	349	257	4576
LSTM	3472	695	408	4425
CNN-Transformer	2833	177	178	3240

Figure 4-7: Performance of Models across Datasets

Regarding the sufficiency of a single sensor (specifically the left-hand sensor) for tremor detection, results were less promising (Table 4.4). Despite extensive hyperparameter tuning, models trained on single-sensor data yielded lower accuracies (CNN: 85.0%, CNN-Transformer: 82.0%) compared to multi-sensor fusion. This contrasts with Hathaliya et al. [12], who reported 92.4% accuracy using only the left-hand sensor, but aligns with Gil-Martin et al. [11], who observed an 8.2% accuracy boost when combining data from all sensors.

Table 4.4: Comparison of CNN, LSTM, and Transformer results for one sensor

Model	Dataset	Accuracy	Training Time (s)
CNN (one hand)	dynamic_raw	75.5%	165.04
	dynamic_filtered	74.6%	110.59
	static_raw	82.2%	46.54
	static_filtered	85.0%	110.09
CNN + Transformer (one hand)	dynamic_raw	68.9%	192.59
	dynamic_filtered	67.7%	148.96
	static_raw	79.7%	669.04
	static_filtered	82.0%	977.70

These findings address RQ3 and highlight the importance of multi-sensor fusion for capturing inter-limb coordination and improving PD detection robustness.

To provide a comprehensive overview of model performance, Figure 4-7 summarizes the classification accuracies across all models and dataset variants using multi-sensor data.

The results presented in this chapter demonstrate the effectiveness of transformer-based architectures and the importance of posture-specific data segmentation and multi-sensor fusion for PD detection using wearable sensors. While these findings match and slightly advance the state-of-the-art, certain limitations exist, such as reliance on a single dataset and challenges in dynamic activity analysis. The next chapter explores these limitations in depth, contextualizes the results within the broader research landscape, and outlines future research directions to enhance clinical applicability and robustness.

Chapter 5

Discussion and Future Work

This chapter discusses the key limitations of the current work on PD detection using wearable sensor data and DL models, and proposes directions for future research to address these challenges.

This work was motivated by the need for continuous, privacy-preserving, and accurate PD detection using wearable sensor data and advanced DL models. The research questions focused on (1) eliminating manual feature engineering through DL architectures, (2) improving detection accuracy via posture-specific (activity-based) data segmentation, and (3) evaluating the impact of multi-sensor fusion versus single-sensor setups.

5.0.1 Discussion

A key limitation related to RQ1 is the reliance on a single publicly available dataset [1], collected in a semi-controlled environment with a limited number of participants. While the DL models demonstrated promising performance on this dataset, their generalizability to broader populations, diverse sensor types, and fully uncontrolled real-world settings remains untested.

Regarding RQ2, this study assumed that activity type (static vs. dynamic) is known a priori, enabling posture-specific segmentation. This assumption simplifies the modeling pipeline but is not always realistic in practical scenarios where activity recognition itself is challenging. Moreover, uniform filtering was applied across activities without tailoring preprocessing to the specific characteristics of static or dynamic movements. These factors may constrain the models' ability to fully exploit activity-specific information, potentially limiting detection accuracy in complex, real-life conditions.

For RQ3, the findings highlight the superiority of multi-sensor fusion over single-sensor data for PD detection. However, the single-sensor experiments focused only on the left hand, which may not capture the most affected limb for all patients due to symptom asymmetry or progression [16]. This limitation underscores the need for flexible sensor configurations and adaptive modeling to accommodate patient variability.

Finally, while the DL architectures effectively learned from raw or minimally processed data, their "black box" nature poses challenges for clinical interpretability. Understanding which features or temporal patterns drive classification decisions is crucial for clinician trust and adoption but was not addressed in this work.

Building on the limitations identified in relation to the thesis motivation and research questions, several directions emerge to advance PD detection using wearable sensors and DL.

5.0.2 Future Work

To address the limitation of dataset scale and diversity highlighted in RQ1, future work should prioritize external validation of the proposed models on independent datasets. These datasets should ideally encompass a wider range of demographics, sensor types, and real-world conditions to rigorously test model generalizability.

In response to the assumption of known activity type in RQ2, future research should develop integrated activity recognition frameworks or end-to-end models capable of jointly detecting patient activity and PD symptoms. This would remove the need for manual or prior activity labeling, enabling truly continuous and automated monitoring. Additionally, exploring adaptive or activity-aware preprocessing pipelines that tailor filtering and feature extraction to the specific characteristics of static and dynamic movements could further improve detection accuracy.

Regarding RQ3 and the limitations of single-sensor data, future studies should investigate flexible sensor configurations that can adapt to patient-specific symptom presentation, including asymmetry and progression. This may involve dynamic sensor selection or weighting schemes within multi-sensor fusion frameworks to maximize diagnostic accuracy while minimizing sensor burden.

Finally, to overcome the interpretability challenges of DL models, future work should incorporate explainability techniques such as attention visualization, saliency maps, or model-agnostic interpretability methods. Enhancing model transparency will be critical for clinical acceptance and could provide novel insights into the physiological and behavioral markers of PD.

Together, these future directions aim to bridge the gap between promising experimental results and practical, scalable, and interpretable PD monitoring solutions, fulfilling the motivation of this work.

Chapter 6

Conclusion

This work has investigated the application of advanced deep learning architectures, particularly transformer-based models, for the detection of Parkinson's disease using wearable sensor data. Motivated by the need for continuous, privacy-preserving, and accurate PD monitoring, this work focused on three key research questions: (1) the ability of DL models to learn directly from raw or minimally processed sensor data without manual feature engineering, (2) the impact of posture-specific (activity-based) data segmentation on detection accuracy, and (3) the effectiveness of multi-sensor fusion compared to single-sensor setups.

The experimental results demonstrate that the hybrid CNN-Transformer model consistently outperforms traditional CNN and LSTM architectures, achieving a peak accuracy of 94.0% on the best-performing static filtered dataset. The convolutional layers excel at extracting local temporal features-such as short bursts and micro-patterns characteristic of tremor signals-while the transformer's self-attention mechanism effectively models long-range temporal dependencies and cross-sensor correlations. This combination enables the model to detect subtle, distributed patterns of motor impairment, including asymmetric tremor propagation and coordination deficits that unfold over time. In contrast, LSTMs, although theoretically capable of capturing long-term dependencies, exhibited slower convergence and less stable training dynamics.

The work also highlights the critical importance of posture-specific analysis. Static (sitting) data consistently yielded higher accuracy and reduced training time compared to dynamic (walking) data, reflecting the more stable and less noisy nature of tremor signals during stationary periods. This finding aligns with the thesis motivation to use activity segmentation to improve detection robustness.

Furthermore, multi-sensor fusion proved essential for robust PD detection. Experiments using data from a single sensor-specifically the left hand-showed substantially lower accuracy, underscoring the value of capturing inter-limb coordination and symptom heterogeneity through multiple wearable sensors.

While these findings establish a proof of concept, the work acknowledges limitations including reliance on a single dataset collected in semi-controlled conditions and the assumption of known activity types during preprocessing. Addressing these limitations through external validation, integrated activity recognition, adaptive preprocessing, and model interpretability are important areas for future research.

In summary, this work contributes to the advancement of wearable sensor-based PD detection by demonstrating the feasibility and advantages of transformer-based deep learning models combined with thoughtful data segmentation and sensor fusion strategies. These results pave the way for scalable, and objective monitoring systems that have the potential to transform PD diagnosis and management in real-world settings.

Bibliography

[1] Jamie L. Adams, Karthik Dinesh, Christopher W. Snyder, Mulin Xiong, Christopher G. Tarolli, Saloni Sharma, E. Ray Dorsey, and Gaurav Sharma. Pdbiostamprc21: Parkinson's disease accelerometry dataset from five wearable sensor study, 2020.

[2] Jamie L. Adams, Karthik Dinesh, Christopher W. Snyder, Mulin Xiong, Christopher G. Tarolli, Saloni Sharma, E. Ray Dorsey, and Gaurav Sharma. Multiple wearable sensors in parkinson and huntington disease individuals: A pilot study in clinic and at home. *Digital Biomarkers*, 5(1):33–44, 2021.

[3] Jamie L. Adams, Karthik Dinesh, Mulin Xiong, Christopher G. Tarolli, Saloni Sharma, Nita Sheth, Alexander J. Aranyosi, Wenlong Zhu, Sara Goldenthal, Kevin M. Biglan, E. Ray Dorsey, and Gaurav Sharma. A real-world study of wearable sensors in parkinson's disease. *Digital Biomarkers*, 1(1):52–63, 2017.

[4] Warwick R. Adams. High-accuracy detection of early parkinson's disease using multiple characteristics of finger movement while typing. *Scientific Reports*, 7:13665, 2017.

[5] Mahsa Sadat Afzali Arani. *Human Activity Recognition: A Comparative Study to Assess the Contribution Level of Accelerometer and Vital Signals*. PhD thesis, Concordia University, 2022.

[6] Marc Bächlin, Meir Plotnik, Daniel Roggen, Inbal Maidan, Jeffrey M. Hausdorff, Nir Giladi, and Gerhard Tröster. Wearable assistant for parkinson's disease patients with the freezing of gait symptom. *IEEE Transactions on Information Technology in Biomedicine*, 14(2):436–446, 2010.

[7] JC Bélisle-Pipon, M Powell, R English, MF Malo, V Ravitsky, Bridge2AI-Voice Consortium, and Y Bensoussan. Stakeholder perspectives on ethical and trustworthy voice ai in health care. *Digital Health*, 10:20552076241260407, 2024.

[8] Karthik Dinesh, Mulin Xiong, Jamie Adams, Ray Dorsey, and Gaurav Sharma. Signal analysis for detecting motor symptoms in parkinson's and huntington's disease using multiple body-affixed sensors: A pilot study. In *2016 IEEE Western New York Image and Signal Processing Workshop (WNYISPW)*, pages 1–5, 2016.

[9] Youcef Djenouri, Ahmed Nabil Belbachir, Gautam Srivastava, and Jerry Chun-Wei Lin. A hybrid visual transformer for efficient deep human activity recognition. In *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*, pages 18816–18826, 2023.

[10] Manuel Gil Martín, Sergio Esteban Romero, Fernando Fernández Martínez, and Rubén San Segundo Hernández. Parkinson's disease detection through inertial signals and posture insights. In *Proceedings of the 16th International Conference on Agents and Artificial Intelligence*, pages 1144–1151. SciTePress, 2024.

[11] M. Gil-Martín, S. Esteban-Romero, F. Fernández-Martínez, and R. San-Segundo. A comprehensive analysis of parkinson's disease detection through inertial signal processing. *Proceedings of the 16th International Conference on Agents and Artificial Intelligence - Volume 3: ICAART*, pages 462–469, 2024.

[12] J. J. Hathaliya, H. Modi, R. Gupta, S. Tanwar, P. Sharma, and R. Sharma. Parkinson and essential tremor classification to identify the patient's risk based on tremor severity. *Computers and Electrical Engineering*, 101:107946, 2022.

[13] L. E. Heusinkveld, M. L. Hacker, M. Turchan, T. L. Davis, and D. Charles. Impact of tremor on patients with early stage parkinson's disease. *Frontiers in Neurology*, 9:628, August 2018.

[14] Aiesha Ibrahim and Mazin Mohammed. A comprehensive review on advancements in artificial intelligence approaches and future perspectives for early diagnosis of parkinson's disease. *International Journal of Mathematics, Statistics, and Computer Science*, 2:173–182, 01 2024.

[15] Lambert Igene, Anika Alim, Masudul H Imtiaz, and Stephanie Schuckers. A machine learning model for early prediction of parkinson's disease from wearable sensors. In *2023 IEEE 13th Annual Computing and Communication Workshop and Conference (CCWC)*, pages 0734–0737, 2023.

[16] R. N. L. Lamptey, B. Chaulagain, R. Trivedi, A. Gothwal, B. Layek, and J. Singh. A review of the common neurodegenerative disorders: Current therapeutic approaches and the potential role of nanotherapeutics. *International Journal of Molecular Sciences*, 23(3):1851, 2022.

[17] Alex Li and Chenyu Li. Detecting parkinson's disease through gait measures using machine learning. *Diagnostics*, 12:2404, 10 2022.

[18] Luca Lonini, Andrew Dai, Nicholas Shawen, Tanya Simuni, Cynthia Poon, Leo Shimanovich, Margaret Daeschler, Roozbeh Ghaffari, John Rogers, and Arun Jayaraman. Wearable sensors for parkinson's disease: which data are worth collecting for training symptom detection models. *npj Digital Medicine*, 1, 12 2018.

[19] Iveta Luptáková, Martin Kubovčík, and Jiří Pospíchal. Wearable sensor-based human activity recognition with transformer model. *Sensors*, 22:1911, 03 2022.

[20] C Maidhof, J Offermann, and M Ziefle. Eyes on privacy: acceptance of video-based aal impacted by activities being filmed. *Frontiers in Public Health*, 11:1186944, 2023.

[21] J. Mei, C. Desrosiers, and J. Frasnelli. Machine learning for the diagnosis of parkinson's disease: A review of literature. *Frontiers in Aging Neuroscience*, 13:633752, May 6 2021.

[22] S. Moon, H. J. Song, V. D. Sharma, K. E. Lyons, R. Pahwa, A. E. Akinwuntan, and H. Devos. Classification of parkinson's disease and essential tremor based on balance and gait characteristics from wearable motion sensors via machine learning techniques: a data-driven approach. *Journal of NeuroEngineering and Rehabilitation*, 17(1):125, Sep 11 2020.

[23] C. Moreau, T. Rouaud, D. Grabli, et al. Overview on wearable sensors for the management of parkinson's disease. *npj Parkinson's Disease*, 9:153, 2023.

[24] Shyamal Patel, Konrad Lorincz, Richard Hughes, Nathaniel Huggins, John Growdon, David Standaert, and Michael Welsh. Monitoring motor fluctuations in patients with parkinson's disease using wearable sensors. *IEEE Transactions on Information Technology in Biomedicine*, 13(6):864–873, 2009.

[25] Elisa Rovini, Cecilia Maremmani, and Filippo Cavallo. How wearable sensors can support parkinson's disease diagnosis and treatment: A systematic review. *Frontiers in Neuroscience*, 11:555, 2017.

[26] Oumaima Saidani, Majed Alsafyani, Roobaea Alroobaea, Nazik Alturki, Rashid Jahangir, and Leila Jamel. An efficient human activity recognition using hybrid features and transformer model. *IEEE Access*, 11, 09 2023.

[27] Silvia Sapienza, Olga Tsurkalenko, Mindaugas Giraitis, et al. Assessing the clinical utility of inertial sensors for home monitoring in parkinson's disease: a comprehensive review. *npj Parkinson's Disease*, 10:161, 2024. Open Access under CC BY 4.0.

[28] Fabrizio Stocchi, Daniele Bravi, Aron Emmi, and Angelo Antonini. Parkinson disease therapy: current strategies and future research priorities. *Nature Reviews Neurology*, 20(12):695–707, 2024.

[29] William Tam, Mohannad Alajlani, and Alaa Abd-alrazaq. An exploration of wearable device features used in uk hospital parkinson disease care: Scoping review. *J Med Internet Res*, 25:e42950, Aug 2023.

[30] C. Virbel-Fleischman, Y. Rétory, S. Hardy, C. Huiban, J. C. Corvol, and D. Grabli. Body-worn sensors for parkinson's disease: A qualitative approach with patients and healthcare professionals. *PLoS One*, 17(5):e0265438, 2022.

[31] Aner Weiss, Sarvi Sharifi, Meir Plotnik, Jeroen P. P. van Vugt, Nir Giladi, and Jeffrey M. Hausdorff. Toward automated, at-home assessment of mobility among patients with parkinson disease, using a body-worn accelerometer. *Neurorehabilitation and Neural Repair*, 25(9):810–818, 2011. PMID: 21989633.