Preprints.org

Article

Detecting Minor Symptoms of Parkinson's Disease in the Wild Using Bi-LSTM with Attention Mechanism

Vasileios Skaramagkas [*] , Iro Boura , Cleanthi Spanaki , Emilia Michou , Georgios Karamanis , Zinovia Kefalopoulou , Manolis Tsiknakis

Article

Detecting Minor Symptoms of Parkinson's Disease in the Wild Using Bi-LSTM with Attention Mechanism

Vasileios Skaramagkas [1,2,*], **Iro Boura** [3,4], **Cleanthi Spanaki** [3,5], **Emilia Michou** [6], **Georgios Karamanis** [7], **Zinovia Kefalopoulou** [7], **and Manolis Tsiknakis** [1,2]

[1]   Institute of Computer Science, Foundation for Research and Technology Hellas (FORTH), GR-700 13 Heraklion, Greece; vskaramag@ics.forth.gr (V.S.); tsiknaki@ics.forth.gr (M.T.)

[2]   Department of Electrical and Computer Engineering, Hellenic Mediterranean University, GR-710 04 Heraklion, Greece

[3]   School of Medicine, University of Crete, GR-710 03 Heraklion, Greece; boura.iro@gmail.com (I.B.); kliospanaki@gmail.com (C.S.)

[4]   Department of Basic and Clinical Neuroscience, Institute of Psychiatry, Psychology and Neuroscience, King's College London, WC2R 2LS London, UK

[5]   Department of Neurology, University Hospital of Heraklion, GR-715 00 Heraklion, Greece

[6]   School of Health Rehabilitation sciences, Department of Speech and Language Therapy, University of Patras, GR-265 04 Patras, Greece; emiliamichou@upatras.gr (E.M.)

[7]   Department of Neurology, Patras University Hospital, GR-264 04 Patras, Greece; zkefalopoulou@gmail.com (Z.K.); karamanis6eor@gmail.com (G.K.)

*   Correspondence: vskaramag@ics.forth.gr

Abstract: Parkinson's Disease (PD) is a neurodegenerative disorder characterized by motor and non-motor impairment with various implications on patients' quality of life. Since currently available therapies are only symptomatic, identifying individuals with prodromal, preclinical or early-stage PD is crucial, as they would be ideal candidates for future disease-modifying therapies. Our analysis aims to develop a robust model for accurate PD detection using accelerometer data collected from PD and non-PD individuals with mild or no tremor during phone conversations. An open-access dataset comprising accelerometer recordings from 22 PD patients and 11 healthy controls (HC) was utilized. The data were pre-processed to extract relevant time, frequency and energy-related features, and a Bidirectional Long Short-Term Memory (Bi-LSTM) model with attention mechanism was employed for classification. The performance of the model was evaluated using 5-fold cross-validation, and metrics of accuracy, precision, recall, specificity, and f1-score were computed. The proposed model demonstrated high accuracy (98%), precision (99%), recall (98%), specificity (96%), and f1-score (98%) in accurately distinguishing PD patients from HC. Our findings indicate that the proposed model outperforms existing approaches and holds promise for detection of PD with subtle symptoms, like tremor, in the wild. Such symptoms can present in the early or even prodromal stage of the disease, and appropriate mobile-based applications may be a practical tool in real-life settings to alert individuals at risk to seek medical assistance or give patients feedback in monitoring their symptoms.

Keywords: Parkinson's disease; Bi-LSTM with attention; in the wild detection; accelerometer; deep learning

1. Introduction

Parkinson's disease (PD) is the second most common neurodegenerative condition, estimated to currently affect more than 8.5 million people globally according to the World Health Organization (WHO) [1]. An elevated risk for PD driven by population ageing and industrialization-related environmental factors is likely to result in a PD pandemic within the following years [2]. PD has been associated with significantly poorer health-related quality of life, even in the early stage of the disease [3,4]. Interestingly, it was found to be the most rapidly growing disease regarding figures of

prevalence, disability and fatality among neurological disorders, with the latter being the primary source of disability worldwide [5]. Consequently, and since its global burden is expected to surge in the future, interventions for early diagnosis and efficient treatment are imperative.

Care management for PD is personalized and multi-factorial in nature [6]. Early recognition of symptoms allows a timely and informed introduction to efficient multi-level/ multi-disciplinary approaches, both pharmacologic and non-pharmacologic (e.g. exercise, speech/occupational therapy), as clinically indicated, with a direct and significant impact on patients' quality of life [7]. However, to date, available therapeutic options aim solely at symptomatic improvement without being able to slow down, halt or ideally reverse the underlying neurodegenerative process and, consequently, disease progression [8]. On the other hand, heterogeneity of PD in terms of age at symptoms' onset, constellation of symptoms, progression, response to therapy and genetic substrate may relate to distinct disease subtypes (endotypes), which may be linked to different underlying biological processes. Indeed, recent advances in PD genetics have shown that different pathogenetic mechanisms are involved in various degrees and combinations in different patients [9]. Biomarkers of several types (clinical, imaging, biochemical, genetic and digital) may assist in better understanding the subjacent molecular pathways that drive PD development in each patient and guide individualized treatment [10]. Novel disease-modifying therapies, which specifically target some, but not other, pathways are currently being developed in order to address this unmet need in the treatment of PD . Detection of patients at a prodromal or even a pre-symptomatic (preclinical) stage of PD, when neurodegeneration is at more initial stages, seems crucial, as this population is more likely to receive the full therapeutic benefit of the new targeted therapies [11,12].

PD diagnosis is primarily based on clinical assessment with bradykinesia being the core symptom accompanied by rest tremor and/or rigidity [13]. Additional features may support (e.g. clear and positive response to levodopa), exclude (e.g. ataxia) the diagnosis or alert the treating physician for a potential alternative diagnosis of atypical or non-degenerative parkinsonism (e.g. early falls, pyramidal signs), while non-motor symptoms hold an essential role in identifying potential PD subtypes [14,15]. Early diagnosis of PD is challenging and significantly dependent on the clinician's expertise, as to date there is no established diagnostic studies/tests that could efficiently substitute for clinical assessment [16].

Although not a prerequisite for PD diagnosis, resting tremor is a cardinal symptom in PD with the vast majority of patients experiencing other type of tremors as well, such as postural or kinetic tremor [17,18]. Tremor is often the very first symptom noticed by PD patients or their caregivers, leading them to seek medical help. There are also indications that mild tremor may be present in the early or even prodromal phase of PD [19], but also in non-parkinsonian carriers of selected PD-related genetic mutations [20], thus, making it useful as a clinical biomarker of diagnostic or predictive value. Such motor phenomena may be so subtle that only instrumented motor assessments using portable devices can detect them, thus unraveling a population at risk.

In recent years, there has been growing interest in utilizing artificial intelligence (AI) techniques to develop non-invasive and accessible methods for early PD detection [21]. The advent of wearable devices, such as accelerometers embedded in smartphones or smartwatches, has provided a wealth of data which can be leveraged to extract meaningful information about an individual's movement patterns and identify potential biomarkers for PD [22,23]. The use of machine learning (ML) and deep learning (DL) techniques that employ sensor signals (e.g., accelerometers, gyroscopes, pressure sensors) has garnered increased attention [24–26]. These wearable sensors offer a cost-effective solution for long-term monitoring under real-world conditions [27–30] in contrast to other analysis systems (such as video-based motion analysis), which require costly equipment and are restricted to indoor environments end experimental settings [31]. These findings indicate a promising future for monitoring and management of PD patients based on sensor signals and AI-related techniques.

We here present a novel approach for detecting PD in the wild using a Bidirectional Long Short-Term Memory (Bi-LSTM) model with an attention mechanism. The key contribution of our

work lies in the development of a robust and accurate model for PD detection in individuals with slight symptoms of tremor using accelerometer data collected in real-world settings. Unlike previous studies that primarily focused on controlled clinical environments, our approach aims to address the challenges of detecting PD in the wild, where data variability and noise are more prevalent. By leveraging the ubiquity of smartphones and smartwatches, our method allows for convenient and non-intrusive monitoring of individuals in their natural environments. Furthermore, our proposed architecture jointly captures temporal dependencies and focuses on relevant information within the accelerometer sequences. The use of Bi-LSTM layers enables the model to consider both past and future contexts, facilitating the extraction of discriminative features for classification of parkinsonian symptoms severity. Additionally, the attention mechanism enhances the model's ability to emphasize important segments of the input data, providing valuable insights into the underlying patterns contributing to PD detection.

Our paper follows a structured format to present the research findings in a clear and organized manner. It begins with an introduction which provides background information on PD and highlights the significance of detecting PD early when symptoms are minor and clinically questionable. The methodology section outlines the data collection process and describes the implementation of the Bi-LSTM model for classification. The results section presents the experimental outcomes, including accuracy and loss curves, box plots, Receiver Operating Characteristic (ROC) curves, and performance metrics. The discussion section interprets the findings, explores the implications, and provides insights into the model's performance, generalizability, and potential applications in early-stage PD detection. Finally, the conclusions section summarizes the key findings, highlighting the importance of the proposed model and its prospect towards PD detection when only subtle symptoms of tremor are present.

2. Literature review: Deep learning for Parkinson's disease identification based on upper limb motion data

2.1. Artificial Neural Networks

Artificial neural networks (ANN) have revolutionized the field of healthcare and diagnosis, surpassing the efficacy and adaptability of conventional ML techniques [32]. ANN have the potential to improve medical decision-making, patient outcomes, and delivery of care. Inevitably, Multi Layer Perceptron (MLP), but also Deep Neural Network (DNN) models, have been developed additionally for diagnostic purposes related to PD. Authors in [33] proposed a novel mobile phone application system based on measuring the hand acceleration using a mobile phone accelerometer. The study utilized recordings from 21 PD patients and 21 HC, extracting features through wavelet packet analysis to construct a 12-element feature vector. A neural network classifier achieved a 95% accuracy and a 90% Kappa coefficient, indicating the reliable detection of PD. In another study, handwriting patterns were evaluated as potential biomarkers for PD [34]. The researchers collected 935 handwriting assignments from 55 PD patients and 94 HC. Three feature sets were extracted: neuromotor, kinematic, and non-linear dynamic. Among traditional ML classifiers, an MLP network was tested, achieving an accuracy of 97% and 78% for classifying between PD and young or elderly HC respectively.

Bazgir et al [35] evaluated tremor using a supervised learning pattern recognition system. The study implemented a classifier block and neural network based on smartphone-recorded acceleration data. The results demonstrated a high accuracy of the neural network, showing a 91% correlation between the tremor item in the Unified Parkinson's Disease Rating Scale (UPDRS) and acceleration values. This approach offered a potential method for evaluating tremor in PD patients. Finally, a study utilized signal processing techniques to examine the consistency between upper limb tremor scores in PD as assessed by the clinician and wearable devices [36]. Acceleration signals were collected during standard movements, and characteristic values were extracted using various domains. The neural network model showed higher accuracy compared to other models, achieving 95.3% accuracy for

resting tremor scores and 91.2% accuracy for postural tremor scores. The combination of acceleration signals and displacement signals effectively simulated the clinician's scores for tremors in the upper limb of PD patients.

2.2. Convolutional Neural Networks

Recently, Convolutional Neural Networks (CNN) have demonstrated remarkable capabilities in medical image analysis, disease detection, diagnosis, and prognosis [21,37]. These networks have the potential to significantly impact healthcare by contributing to early disease detection, treatment planning, and improved clinical outcomes. To this goal, several studies have explored the use of CNN in the context of PD and parkinsonian tremor detection. Xing et al [38] evaluated seven predictive models, including CNN, to differentiate between PD and essential tremor (ET) using demographics and upper limb tremor data. The CNN model, trained on tremor frequency and amplitude-related features, demonstrated modest predictive ability with accuracy above 78% and f1-score above 0.83. In another study, a CNN-based method for detecting PD hand tremor using wrist acceleration information collected from a wearable sensor device was proposed [39]. Their CNN model with nine layers outperformed conventional ML techniques in detecting PD hand tremor, reaching an accuracy of 97%.

The advantages of accelerometers and gyroscopes as wearable sensors on the wrists were exploited towards the development of TremorSense, a PD tremor detection system [40]. Researchers employed an eight-layer CNN model to classify rest, postural, and action tremor in PD, achieving an accuracy of over 94% in various evaluation scenarios. Moreover, in a recent study [41], scientists introduced and evaluated histograms of oriented gradients (HOG) combined with ML and DL methods, including a 2D-CNN, for evaluation of tremor severity in PD patients. The HOG descriptor achieved high discriminating rates in identifying tremor patterns from handwritten drawings and the network achieved a 83% accuracy. Finally, a tremor assessment system utilizing a CNN to distinguish the severity of symptoms based on data collected from a wearable device was developed by Kim et al [42]. Their CNN architecture demonstrated accurate tremor monitoring with 85% correct prediction rate. The aforementioned studies highlight the potential of CNN in PD and tremor detection, showing improved accuracy and performance compared to conventional techniques.

2.3. Long Short Term Memory

Long Short-Term Memory (LSTM) is a type of recurrent neural network (RNN), which has gained significant attention for its ability to model and analyze sequential data [43]. LSTM excels in capturing dependencies and patterns over long sequences, making them particularly well-suited for time series and sequential data analysis. In healthcare, LSTM-based models have been successfully applied on various occasions, such as disease prediction, patient monitoring, and medical signal analysis [21,44]. These models leverage the memory capabilities of LSTM to capture temporal dynamics and contextual information, enabling accurate predictions and insights from healthcare data.

Taking the above into consideration, Thummikarat et al [45] proposed the use of LSTM network, to aid physicians in PD diagnosis. Their study demonstrated an accuracy of 73% in identifying patients with PD at an early stage with the suggested architecture. In another study [46], RNN models were trained using figure drawing data represented as a time series of coordinates, angles, and pressure readings. The paper compared two recurrent network models, namely LSTM and Echo State Networks, to analyze the advantages and limitations of each architecture. Finally, LSTM achieved an identification accuracy of 91% and an f1-score of.94.

Hssayeni et al [47] used wearable devices and introduced two methods based on deep LSTM networks and gradient tree boosting to continuously monitor parkinsonian tremor using gyroscope sensor signals recorded during various body movements. The evaluation of data acquired from 24 PD patients showed that the LSTM method exhibited relatively high correlation with the clinician's assessment of tremor, as rated by UPDRS III subscores (r = 0.77, p < 0.0001) through subject-based leave-one-out cross-validation.

2.4. Hybrid Deep Learning Architectures

Through our literature search, we identified two studies that used sophisticated DL to distinguish between ET and parkinsonian tremor. Initially, Hathaliya et al [48] integrated Gated Recurrent Unit (GRU) and LSTM algorithms. Specifically, accelerometer data were passed through the GRU model, and served as input into the LSTM model to improve its performance. The authors utilized a blockchain network to validate that the trained model's testing accuracy was 74.1%. Additionally, in [49] a data-driven NeurDNet model was proposed for classifying tremor in PD and ET based on the kinematics of the hand. The network was trained on more than 90 hours of hand motion signals, including 250 tremor evaluations from 81 patients, and its differential diagnostic accuracy of 95.55% exceeds its state-of-the-art equivalents.

Concerning the diagnosis of PD, researchers investigated the applicability of CNN and CNN-BLSTM models based on time series classification derived from pen-based signals [50]. In this context, the Multi-Modal Collection (PDMultiMC) containing recordings of online handwriting, vocal signals, and eye movements was created [51]. HandPDMultiMC, a subset of PDMultiMC, includes examples of handwriting from 42 participants (21 PD patients and 21 non-PD controls). When combined with synthetic data augmentation, CNN-BLSTM models trained with jittering and synthetic data augmentation provided the highest performance for early PD identification (97.62% accuracy), according to experimental results on the dataset.

2.5. Limitations in previous studies, addressed in the present study

In comparison to previous studies, our work addresses several limitations and introduces novel contributions to the field of PD detection. Initially, the aforementioned studies have focused on PD population with variable symptoms' severity of the different PD symptoms. Nevertheless, in our study, and as explained in detail in Section 3.1, we selected PD patients experiencing only slight symptoms of tremor, thus focusing on a more homogeneous group of patients. Secondly, while previous studies have primarily used controlled clinical environments, our approach specifically targets PD detection in real-world settings by leveraging data collected in the wild. This is crucial as it allows us to account for the variability and noise commonly encountered in everyday life, making our model more robust and applicable to diverse scenarios. Furthermore, our method leverages the widespread use of smartphones and smartwatches, enabling convenient and non-intrusive monitoring of individuals in their natural environments. This eliminates the need for specialized equipment or clinical visits, making it more accessible and cost-effective. Finally, our proposed architecture comprising of LSTM layers and an attention mechanism (see Section 3.4) enables us to capture temporal dependencies effectively and highlight relevant information within the accelerometer sequences.

3. Methodology

3.1. Dataset Description

The dataset used in this study, contains IMU signals captured in-the-wild via the accelerometer sensor embedded in modern smartphones for the purpose of detecting tremorous episodes related to PD [52]. 31 PD patients and 14 HC contributed accelerometer data from their personal devices over the course of several months. Automatic recording of triaxial acceleration values was performed whenever a phone call was made. The annotation in the dataset is based on UPDRS items 16, 20, and 21, from now on mentioned as updrs16, updrs20, and updrs21 respectively. updrs 16 refers to the value related to tremor as described in item 16 of the UPDRS part II, as reported by the subject, and updrs20 and updrs21 to the values related to rest tremor and action/postural tremor respectively in the right and left hand as described in items 20 and 21 of the UPDRS part III, as reported by the attending neurologist. Finally, a pd_status of '1' indicates that the subject is a PD patient whereas a value of '0' indicates that the subject is a HC.

For purposes of early diagnosis of PD based solely on the motor symptom of upper limb tremor, we used a subset of the initial dataset by keeping those subjects with updrs16 score ≤ 1 (i.e. 0 and 1). The reason we decided to use updrs16 and not updrs20 or updrs21 was to account for the participant's perspective of their tremor. Thus, we only included individuals who may not have been aware of their tremor (even if it was objectively measured by the clinician in items 20 and/or 21 of UPDRS) or, if they had been, this tremor was perceived as minor without any intrusions in activities of everyday life (even if it was not captured in the clinician's assessment using UPDRS). Due to the subtle symptoms, these individuals may not consider visiting a physician at this point, but the findings of portable technologies may alert them to seek medical help. We also thought that UPDRS item 16 may offer a more complete and concise estimation of patients' tremor compared to UPDRS items 20 and 21. The latter items represent instant assessments that may not reflect the true nature of tremor, a symptoms which, similarly to other PD symptoms, tend to fluctuate during the day due to several factors (stress, excitement, fatigue, posturing, etc). Thus, the final utilized dataset comprises of 22 PD patients and 14 HC.

3.2. Data Preprocessing

To lessen the computing burden and network complexity of the system, the accelerometer data were downsampled to 100 Hz before being used in the training and evaluation stages. Notably, PT signals have an instructive spectral region up to 12 Hz in frequency [53]. Given that the Nyquist theorem states that a signal can be reconstructed in its entirety if it is sampled at a rate twice as fast as its highest informative frequency, we conclude that a sampling rate of 100 Hz is adequate for both minimising the system's computational load and preserving the signal's spectral contents of interest. Last but not least, a Butterworth filter with a 12 Hz cutoff was used to efficiently isolate and analyse the unique frequency range associated with PT [54].

3.3. Windowing and Feature Extraction

After completing the data pre-processing phase, we applied a windowing technique with each window size set to 100 data points (1 sec) with 50% overlap. The purpose of introducing overlap between windows is to ensure that important temporal information in the time series is captured. It allows for smoother transitions between adjacent windows and reduces the chances of missing significant patterns or events that might occur near the edges of individual windows [55]. By defining the window size and overlap, we are specifying the granularity at which the time-series data will be analyzed. Smaller window sizes with higher overlaps provide more detailed information but may increase computational complexity, while larger window sizes with lower overlaps offer a more aggregated view of the data but may potentially miss finer patterns [55].

For this work, thirty-three (33) time, frequency and energy related features were calculated for each window, as seen in Table 1.

Table 1. Time, frequency and energy related features calculated per window (x) with size N = 100.

Feature	Description	Formula
Maximum	The maximum value in the signal (per axis).	$\max(x_1, x_2, \ldots, x_N)$
Minimum	The minimum value in the signal (per axis).	$\min(x_1, x_2, \ldots, x_N)$
Mean	The average value of the signal (per axis).	$\frac{1}{N} \sum_{i=1}^{N} x_i$
Standard deviation (std)	A measure of the dispersion or spread of the signal values around the mean (per axis).	$\sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_i - \text{mean})^2}$

Table 1. *Cont.*

Feature	Description	Formula
Kurtosis	A measure of the peakedness or flatness of the signal's distribution (per axis).	$= \dfrac{\frac{1}{N}\sum_{i=1}^{N}(x_i - \text{mean})^4}{\left(\frac{1}{N}\sum_{i=1}^{N}(x_i - \text{mean})^2\right)^2} - 3$
Zero Crossing Rate	The rate at which the signal changes its sign.	$\frac{1}{N-1}\sum_{i=1}^{N-1}\left(\begin{cases} 1 & \text{if } x_i \cdot x_{i+1} < 0 \\ 0 & \text{otherwise} \end{cases}\right)$
Skewness	A measure of the asymmetry of the signal's distribution (per axis).	$\dfrac{\frac{1}{N}\sum_{i=1}^{N}(x_i - \text{mean})^3}{\left(\frac{1}{N}\sum_{i=1}^{N}(x_i - \text{mean})^2\right)^{\frac{3}{2}}}$
Correlation	The correlation coefficients between different signal components or dimensions (for xy, xz, yz axes).	$\begin{pmatrix} \text{corr}_{xy} & \text{corr}_{xz} & \text{corr}_{yz} \end{pmatrix}$
Maximum PSD	The maximum power spectral density value in the signal.	$\max(\text{psd})$
Average PSD	The average power spectral density value in the signal.	$\frac{1}{N}\sum_{i=1}^{N}\text{psd}_i$
Standard Deviation of PSD	A measure of the variation or spread of the power spectral density values.	$\sqrt{\frac{1}{N}\sum_{i=1}^{N}(\text{psd}_i - \text{avg\_psd})^2}$
Spectral Centroid	The center of mass of the power spectral density distribution, representing the average frequency content of the signal.	$\dfrac{\sum_{i=1}^{N} f_i \cdot \text{psd}_i}{\sum_{i=1}^{N}\text{psd}_i}$
Spectral Rolloff	The frequency below which a specified percentage of the total power of the signal is contained.	$0.85 \cdot \sum_{i=1}^{N}\text{psd}_i$
Spectral Flatness	A measure of the tonality or noisiness of the signal.	$\dfrac{\exp\left(\frac{1}{N}\sum_{i=1}^{N}\log(\text{psd}_i)\right)}{\frac{1}{N}\sum_{i=1}^{N}\text{psd}_i}$
Spectral Skewness	A measure of the asymmetry of the power spectral density distribution around its centroid.	$\frac{1}{N}\sum_{i=1}^{N}\left(\dfrac{f_i - \text{spectral\_centroid}}{\text{std\_psd}}\right)^3$
Spectral Kurtosis	A measure of the peakedness or flatness of the power spectral density distribution around its centroid.	$\frac{1}{N}\sum_{i=1}^{N}\left(\dfrac{f_i - \text{spectral\_centroid}}{\text{std\_psd}}\right)^4 - 3$
Entropy	A measure of the randomness or unpredictability of the signal, calculated using Shannon's entropy formula [56].	$-\sum_{i=1}^{N} p_i \log_2(p_i)$
Total Energy	The total energy or power in the signal, calculated as the sum of the squared values.	$\sum_{i=1}^{N}\left(x_i^2 + y_i^2 + z_i^2\right)$
Signal Magnitude Area	The sum of the absolute values of the signal.	$\sum_{i=1}^{N}\left(|x_i| + |y_i| + |z_i|\right)$

PSD: Power Spectral Density

3.4. Implementation of Bidirectional LSTM with Attention

LSTM is a type of RNN layer that can effectively process sequential data, such as time series or text [57]. The bidirectional variant of LSTM processes the input sequence in both forward and backward directions, allowing the model to capture dependencies in both past and future contexts. The use of a Bi-LSTM with attention model for classifying PD patients and HC using accelerometer signals from smartphones during phone conversations is justified for several reasons [58].

Firstly, the Bi-LSTM architecture is well-suited for capturing temporal dependencies in sequential data, which is essential for analyzing accelerometer signals collected over time. By processing the input sequence in both forward and backward directions, the model can effectively capture patterns and dependencies in both past and future context. This is particularly relevant in the context of PD classification, as the characteristics of movements associated with the disease may manifest in different ways over time. Additionally, the inclusion of an attention mechanism further enhances

the model's ability to focus on relevant parts of the input sequence. PD can exhibit subtle variations in movement patterns and intensity, and the attention mechanism allows the model to dynamically weigh the importance of different time steps in the sequence, emphasizing the most informative parts.

3.4.1. Bidirectional LSTM layer

As mentioned earlier, Bi-LSTM comprises of an LSTM in both the forward and backwards directions. The memory controller in the LSTM is used to determine which information is forgotten and which is retained. It is implemented by three structures: the input gate, the forget gate, and the output gate. The structure of the unit is depicted in Figure 1. The operation process is:

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i) \tag{1}$$
$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_t) \tag{2}$$
$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o) \tag{3}$$
$$g_t = \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \tag{4}$$
$$c_t = f_t c_{t-1} + i_t g_t \tag{5}$$
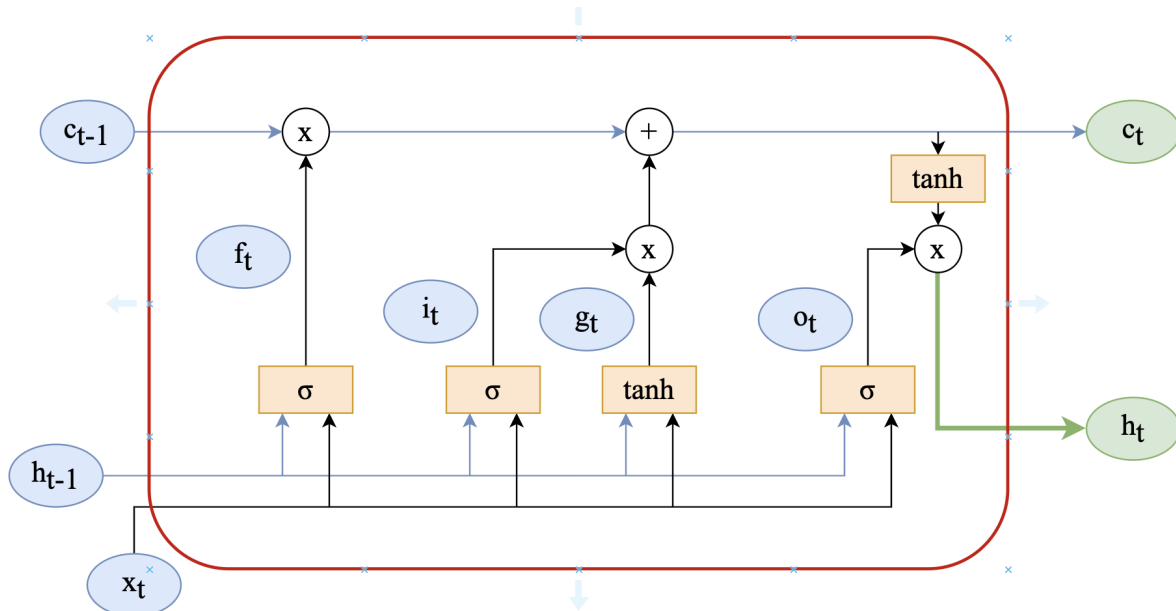$$h_t = o_t \tanh(c_t) \tag{6}$$

Figure 1. LSTM cell internal architecture.

The input gate (1) determines how much new information should be let through to update the cell state c_t at time step t. It takes into account the current input x_t, the previous hidden state h_{t-1}, and bias b_i using the weight matrices W_{xi} and W_{hi}. The sigmoid function σ squashes the combined input, producing a value between 0 and 1 representing the gate's openness (0 means closed, 1 means open). The forget gate (2) determines how much of the previous cell state c_{t-1} should be forgotten at time step t. It takes into account the current input x_t, the previous hidden state h_{t-1}, and bias b_f using the weight matrices W_{xf} and W_{hf}. The sigmoid function σ outputs values between 0 and 1, indicating the amount of information to retain (1 means retain all, 0 means forget all). The output gate (3) determines how much of the cell state c_t should be exposed as the hidden state h_t at time step t. It takes into account the current input x_t, the previous hidden state h_{t-1}, and bias b_o using the weight matrices W_{xo} and W_{ho}. The sigmoid function σ produces values between 0 and 1, controlling the amount of information to be exposed in the hidden state. The new cell state g_t (4) is updated by considering the current input x_t, the previous hidden state h_{t-1}, and bias b_c using the weight matrices

W_{xc} and W_{hc}. The hyperbolic tangent function (tanh) maps the combined input to a value between -1 and 1, capturing new candidate information. The updated cell state c_t (5) is obtained by combining the previous cell state c_{t-1} with the new candidate information g_t. The forget gate f_t determines how much of the previous cell state is forgotten, while the input gate i_t determines how much of the new candidate information is incorporated. Finally, the hidden state h_t is obtained by applying the output gate o_t to the hyperbolic tangent of the updated cell state c_t. The output gate controls the exposure of the cell state information, while the hyperbolic tangent function applies a non-linear transformation to capture relevant features.

The traditional LSTM network, however, can only learn in one direction, thereby disregarding the reverse information. In the Bidirectional LSTM [59], nevertheless, the current input depends not only on the prior accelerometer value but also on the subsequent value. The combination of the two LSTM units thoroughly considers the temporal information preceding and following each acceleration value, as depicted in Figure 2.
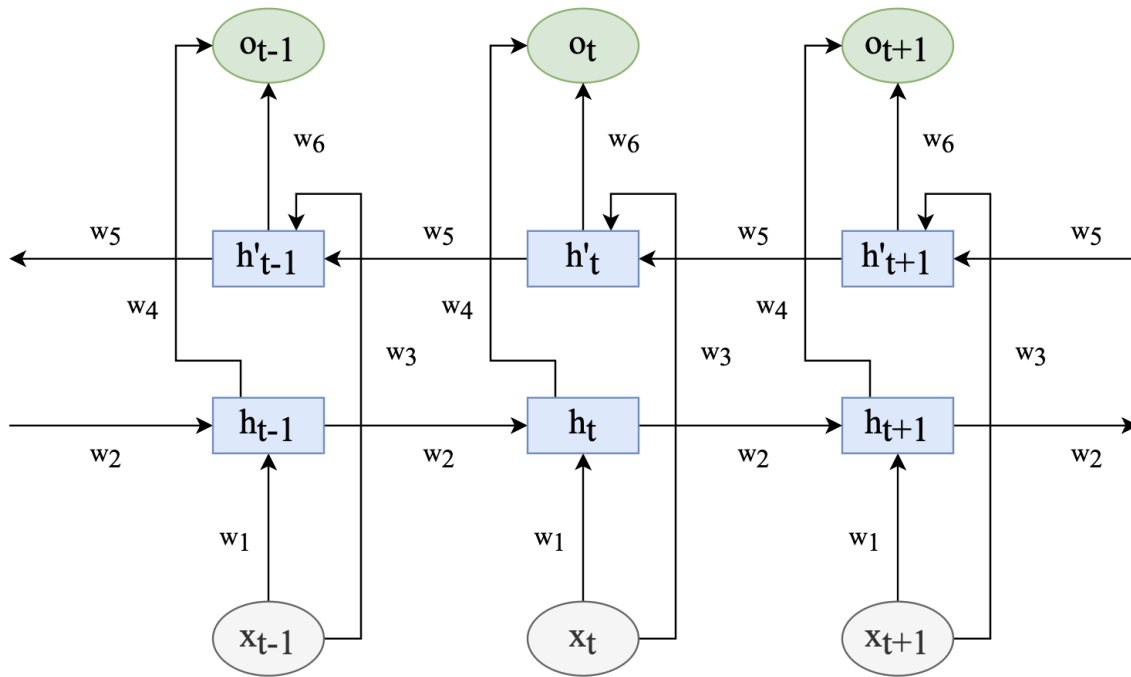

Figure 2. Bidirectional LSTM network model.

$w_i(i = 1, ..., 6)$ denotes the weight from one unit layer to another, and x_t is the feature vector, h means LSTM units of input feature sequence $(..., x_{t-1}, x_t, x_{t+1}, ...)$, h' indicates LSTM units of input feature sequences $(..., x_{t+1}, x_t, x_{t-1}, ..)$ and o_t is the corresponding output after the feature vector passes through the Bidirectional LSTM network.

The operation process is expressed as:

$$h_t = \sigma(w_1 x_t + w_2 h_{t-1} + b_t^{(1)}) \tag{7}$$

$$h_t' = \sigma(w_3 x_t + w_5 h_{t-1}' + b_t^{(2)}) \tag{8}$$

$$o_t' = \tanh(w_4 h_t + b_t^{(3)}) \tag{9}$$

$$o_t'' = \tanh(w_6 h_t' + b_t^{(4)}) \tag{10}$$

$$o_t = 0.5 * (o_t' + o_t'') \tag{11}$$

where $b_t^{(1)}, b_t^{(2)}, b_t^{(3)}, b_t^{(4)}$ are the biases in the Bidirectional LSTM network at time t. Moreover, o_t', o_t'' are the results of two LSTM units dealing with the vector output at the corresponding time. As

shown in (11), the output feature vector o_t is the mean of the two vectors at the corresponding time. The vector is fed into the attention mechanism to learn the network's weight.

3.4.2. Attention layer

By assigning attention weights to the hidden states, the attention layer provides a mechanism to selectively weigh the contributions of the forward and backward hidden states based on their relevance to the task at hand. The attention mechanism process is as follows:

$$e_t = v_a^T \tanh(W_a[h_t; h_t] + b_a) \tag{12}$$

$$\alpha_t = \text{softmax}(e_t) \tag{13}$$

$$c = \sum_t \alpha_t[h_t; h_t] \tag{14}$$

where in (12), h_t^f and h_t^b are the forward and backward hidden states at time step t, respectively. W_a is a weight matrix, b_a is a bias term, and v_a is a weight vector for calculating attention scores. The attention score e_t is obtained by taking the dot product of v_a with the hyperbolic tangent of the linear transformation of the concatenated hidden states. The attention scores e_t in (13), are passed through the softmax function to obtain attention weights α_t. The softmax function normalizes the attention scores across all time steps, ensuring that the weights sum up to 1. α_t represents the importance or weight assigned to each hidden state h_t in the sequence. Finally in (14), the context vector c is calculated by taking the weighted sum of the concatenated hidden states $[h_t^f; h_t^b]$ across all time steps. The attention weights α_t determine the contribution of each hidden state to the context vector and c represents a summary or representation of the input sequence, capturing the most relevant information.

3.4.3. Network architecture

The proposed model, which is presented in Figure 3, consists of multiple bidirectional LSTM layers to capture sequential patterns in the input accelerometer data. Dropout layers are added for regularization, and an attention layer helps the model focus on relevant information. Finally, a flatten layer and dense layer with a sigmoid activation function produces the binary classification output.
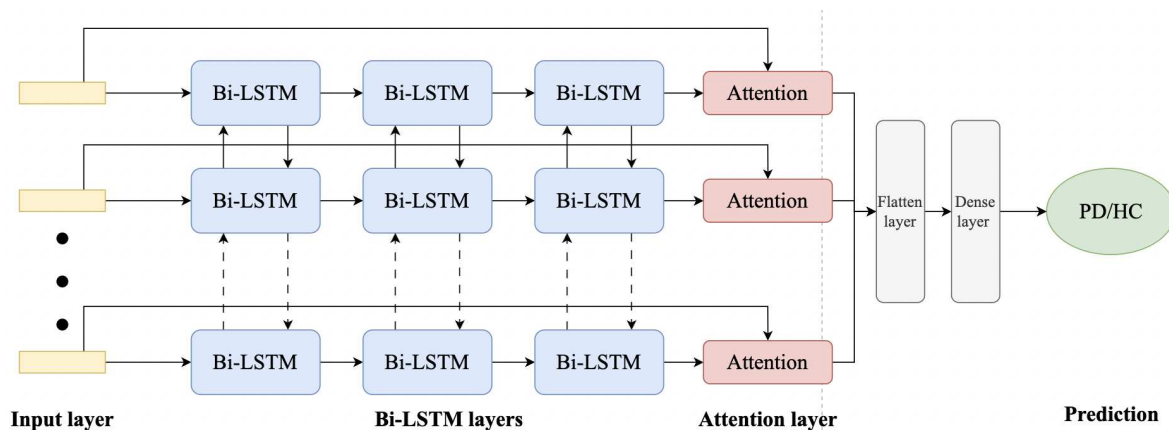

Figure 3. Proposed Bi-LSTM with attention network architecture.

The hyper-parameters parameters of various layers of the presented model are shown in Table 2. Notably, the final model consists of a total of 346,349 parameters.

Table 2. Hyper-parameters of the proposed architecture.

Layer	Neurons	Output shape	Parameters	Activation
Input layer	33	(-, 33, 1)	0	-
Bi-LSTM layer 1	32	(-, 33, 64)	8704	-
Dropout 1	-	(-, 33, 64)	0	-
Bi-LSTM layer 2	64	(-, 33, 128)	66,048	-
Dropout 2	-	(-, 33, 128)	0	-
Bi-LSTM layer 3	128	(-, 33, 256)	263,168	-
Attention layer	-	(-, 33, 256)	0	-
Dropout 3	-	(-, 33, 256)	0	-
Flatten layer	-	(-, 8448)	0	-
Dense layer	1	(-, 1)	8,449	Sigmoid

3.4.4. Experimental setup and implementation

Python version 3.9.7 installed on a personal computer with a GTX GeForce 750 Ti GPU, Intel(R) Core(TM) i7-6700 CPU at 3.40 GHz clock speed, and 32 GB of RAM was used for the implementation and experimentation. TensorFlow-GPU version 2.5.0 with the frontend of Keras-GPU were used for developing, training, and evaluating the DL model. During training, the model's loss was calculated using a binary cross-entropy loss function, and the model's weights were updated using an Adam optimizer with an initial learning rate of $10 - 4$. In addition, 128 minibatch sizes were used to train the suggested model.

4. Results

Regarding our experimental procedure and the classification stage, we split the data into training and testing sets, with the number of test data samples being 30% of the total number of examples [60]. To assess the performance of our model, we employed a 5-fold cross-validation (5-CV) technique, ensuring a robust evaluation. Additionally, hyperparameter optimization was performed using the GridSearch method, allowing us to find the optimal values for the model's hyperparameters. Furthermore, we introduced an Early Stopping callback function with a patience level set to 3, which prevented overfitting and improved the model's generalization capability.

4.1. Training and validation

The training and validation progress of the proposed model was visualized using accuracy and loss curves for each fold of the 5-CV. The accuracy curves depict the performance of the model in correctly classifying PD and HC samples, while the loss curves represent the convergence of the model during training.

Figure 4 illustrates the accuracy curves for the training and validation sets across the five folds. It can be observed that the model achieves high accuracy on both the training and validation sets, indicating its ability to effectively learn the underlying patterns in the accelerometer data. The accuracy steadily increases during the initial epochs and eventually plateaus, demonstrating the convergence of the model. Notably, the validation accuracy closely follows the training accuracy, suggesting that the model generalizes well to unseen data.

In Figure 5, the loss curves for the training and validation sets are presented. The curves depict the reduction of the loss function over the training epochs. As expected, the loss decreases rapidly during the initial epochs and gradually stabilizes as the model converges. The loss curves for the training and validation sets exhibit a similar pattern, indicating that the model is not overfitting to the training data and is able to generalize well to new samples.
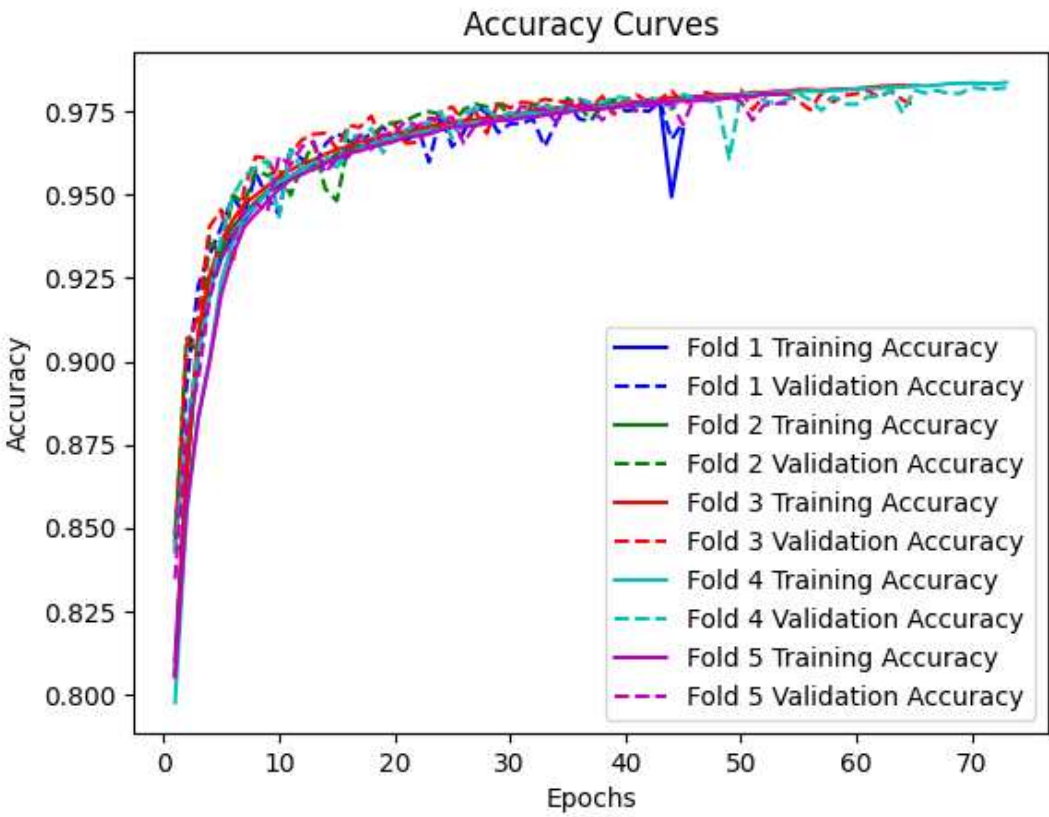
Figure 4. Training and validation accuracy curves for every epoch for BiLSTM with Attention model.
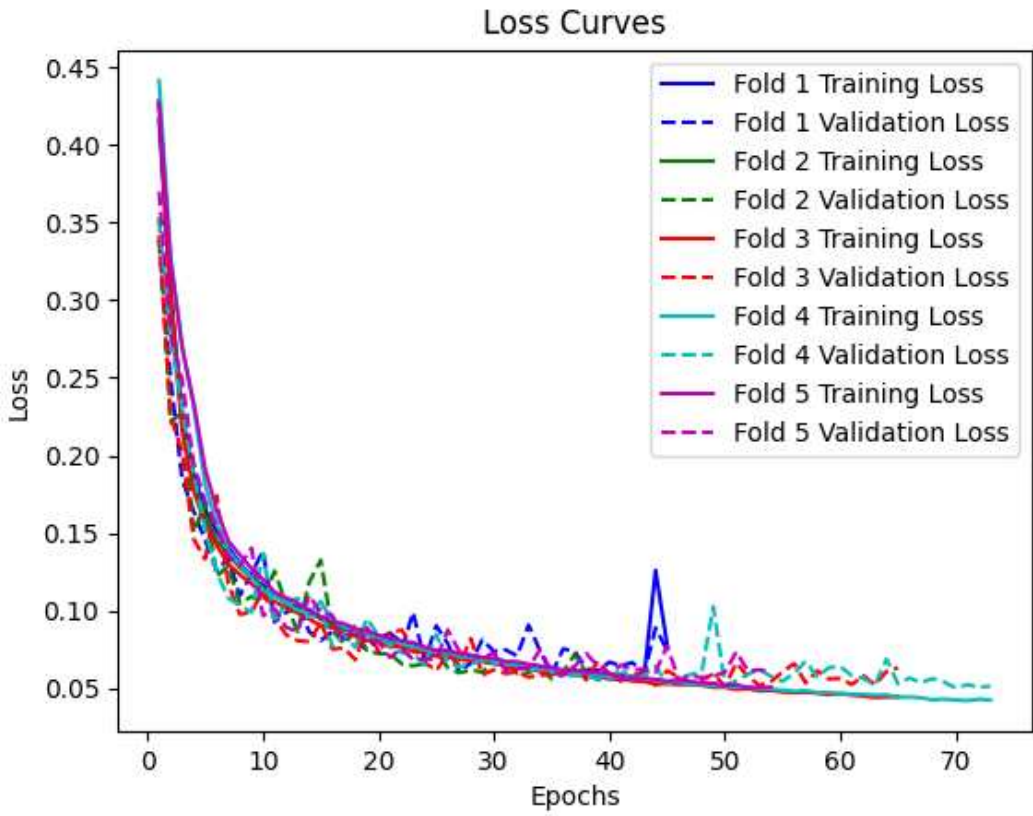

Figure 5. Training and validation loss curves for every epoch for BiLSTM with Attention model.

Furthermore, we visualized the distribution of performance metrics across the different folds and epochs using box plots, as observed in Figure 6. These plots provide valuable insights into the variability and consistency of our model's performance during training and validation. The first set of box plots represents the training and validation loss across different folds at each epoch. The vertical axis denotes the loss values, while the horizontal axis represents the folds. The boxes in the plots indicate the interquartile range (IQR) of the loss distribution, with the median value represented by the horizontal line inside the box. The whiskers extending from the boxes indicate the minimum and maximum values, excluding outliers.
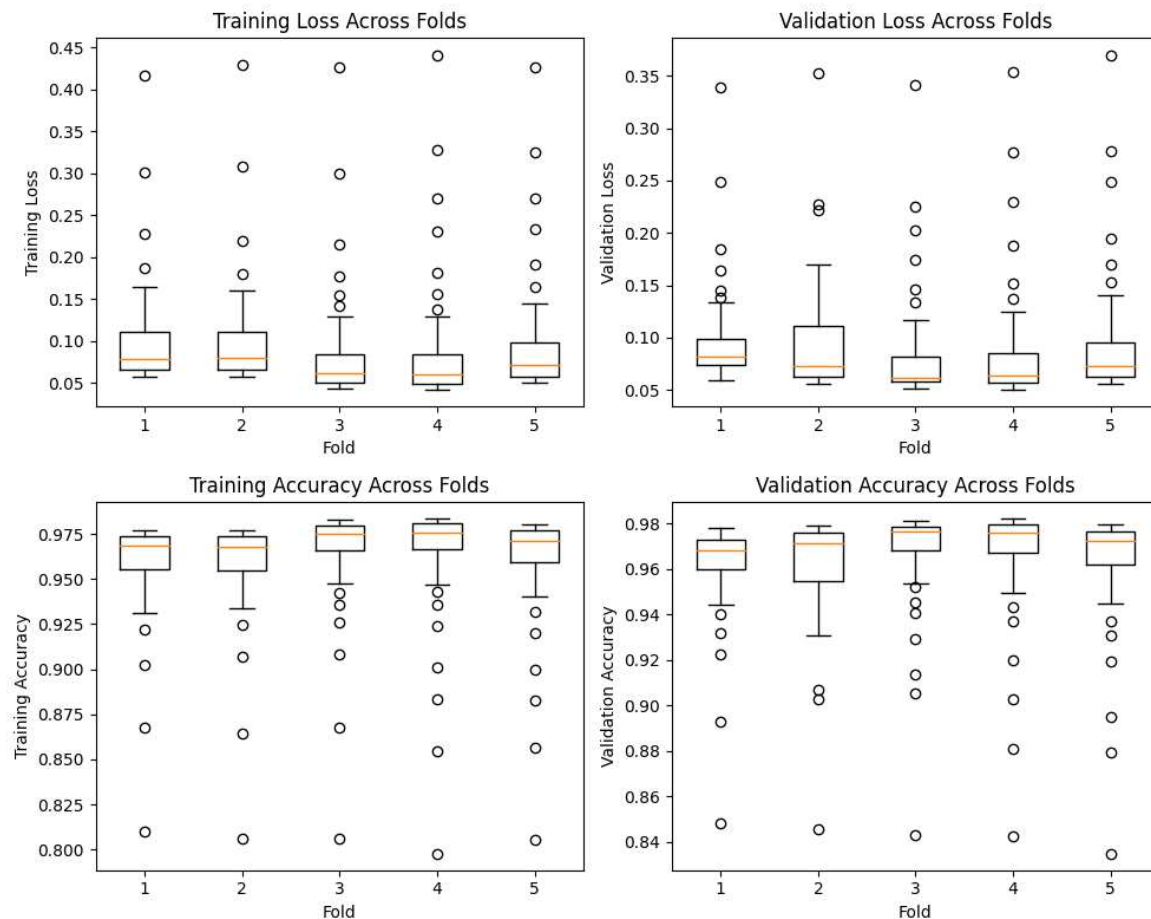

Figure 6. Box plots showing the distribution of training loss, validation loss, training accuracy, and validation accuracy across different folds at each epoch.

Similarly, we created box plots to visualize the training and validation accuracy across different folds at each epoch. These plots showcase the distribution of accuracy values achieved by our model during training and validation. Once again, the boxes represent the IQR, the horizontal line inside the box denotes the median accuracy, and the whiskers display the minimum and maximum values, excluding outliers.

Analyzing these box plots provides a comprehensive understanding of our model's performance characteristics. We can observe the consistency of our model's performance across different folds, with minimal variability in both loss and accuracy. This consistency indicates the robustness of our proposed model in effectively capturing and generalizing patterns in the data.

To further assess the discriminative power and performance of our proposed model, we plotted the ROC curves and calculated the corresponding Area Under the Curve (AUC) for each fold, as presented

in Figure 7. The curves show the model's performance at different classification thresholds, indicating the balance between correctly identifying PD patients and minimizing false positive predictions.
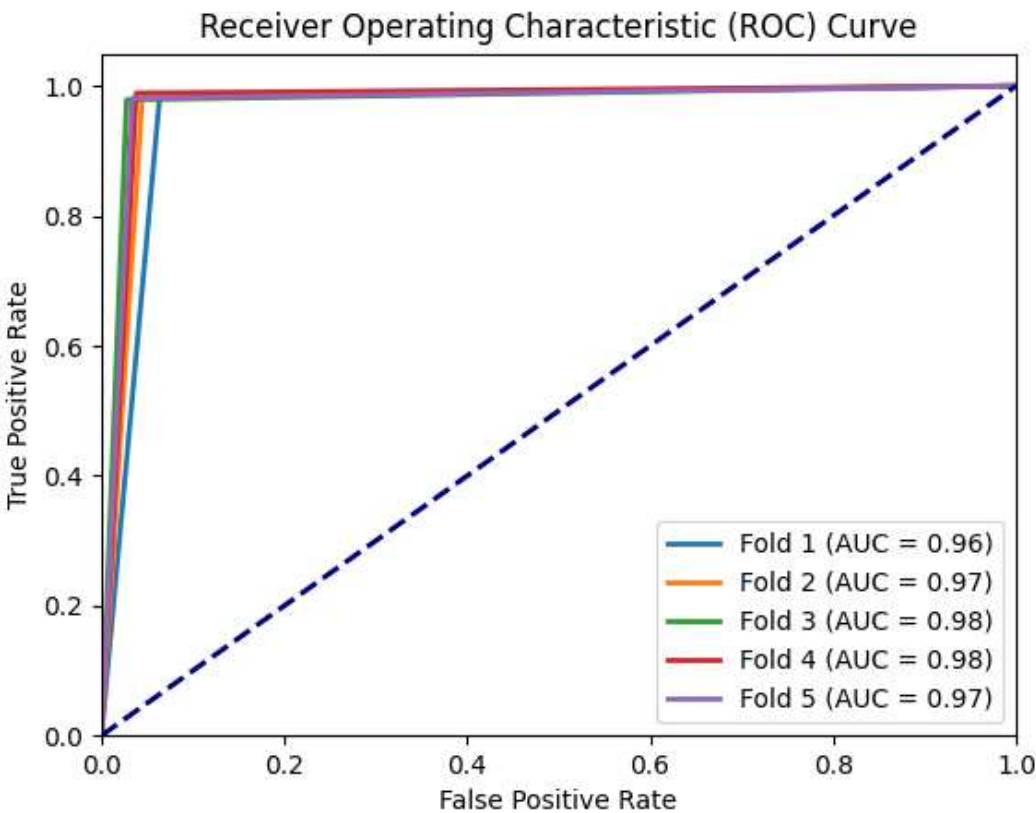

Figure 7. ROC-curves for each fold for BiLSTM with Attention model.

Analyzing the ROC curves and AUC values across the 5 folds, we observed consistent and high performance across all iterations. The AUC values for each fold were consistently high, indicating that the Bi-LSTM with Attention model achieved effective discrimination between PD patients and HC. These results highlight the robustness and effectiveness of our model in accurately classifying PD patients with minor tremor symptoms.

4.2. Performance evaluation on the test set

The performance of our network model was evaluated using several metrics, including accuracy, precision, recall, specificity, and f1-score. These metrics provide insights into different aspects of the model's performance. The average performance metrics obtained from the 5-CV are presented in Table 3.

Table 3. Performance metrics for classification of PD and HC, and comparison with the relevant literature

Model	Experiment	Accuracy	Precision	Recall	Specificity	f1-score
LSTM [45]	Typing	.73				
2 Stacked LSTM [46]	Writing	.91	-	1.00	.65	.94
CNN-BiLSTM [50]	Writing	.98	-	.95	1.00	
DNN [33]	Hand poses	.95	-	-	-	-
BiLSTM + Attention (proposed model)	Talking on the phone	.98	.99	.98	.96	.98

The results demonstrate the effectiveness of our proposed model in accurately classifying the accelerometer data into PD and HC. The high mean accuracy of 0.98 indicates that our model achieved a high overall classification accuracy across all folds. The mean precision of 0.99 indicates a low rate

of false positives, emphasizing the model's ability to correctly identify positive instances. The mean recall of 0.98 suggests that the model effectively captures the majority of positive instances. The mean specificity of 0.96 indicates a high rate of correctly identifying negative instances. Lastly, the mean f1-score of 0.98 represents a balanced measure of precision and recall, indicating the model's strong overall performance.

Additionally, to provide a comprehensive comparison with relevant literature, Table 3 presents the performance metrics of our proposed model alongside the results reported in other research works. Notably, our Bi-LSTM with Attention model achieved superior performance compared to previous studies. The comparison underscores the effectiveness of our proposed model in accurately differentiating PD patients from healthy individuals, particularly when minor symptoms are present, highlighting its significance and potential as a valuable tool for early PD detection and monitoring when presenting symptoms may be subtle. Importantly, our model's success can be attributed to its ability to analyze accelerometer data collected in real-world settings, capturing the true complexities and variabilities of individuals' everyday movements. This feature enhances the generalizability and practicality of our approach, making it suitable for real-life applications and paving the way for widespread adoption in the field of PD research and healthcare.

4.3. An additional experiment with a subset of the dataset

We conducted an additional experiment for purposes of early diagnosis of PD based solely on the motor symptom of upper limb tremor. This time, we utilized an even smaller sample of the initial dataset by keeping those subjects with only zero (0) updrs16 score. We, thus, wanted to address this group of patients who either do not experience any symptoms of tremor (e.g. akinetic type of PD) or if they did (updrs20/21 score of at least 1) they were not aware of it. In this case, we ended up with a much smaller dataset which comprised of 8 PD patients and 13 HC. For this particular experiment, we followed the same data processing, feature engineering, training, validation and testing procedure as described in the previous sections. Our Bi-LSTM with Attention network yielded once more encouraging results with the metrics of accuracy, precision, recall, specificity and f1-score being .97, .98, .96, .96, .96 respectively.

5. Discussion

The results of this study demonstrate the effectiveness of our proposed Bi-LSTM with Attention model in accurately discriminating PD patients from HC based on accelerometer data when individuals experience slight or no tremor (updrs16 score $\leq$ 1), irrespective of the clinician's objective assessment. The model achieved high accuracy, precision, recall, specificity, and f1-score, indicating its robustness and potential for detection and monitoring of PD when only minor motor symptoms of tremor are present, a concept which may apply in patients with early or even prodromal PD.

One key advantage of our model is its ability to efficiently analyze accelerometer data collected in real-world settings. By capturing the true complexities and variabilities of individuals' everyday movements our model enhances the generalizability and practicality of methods of PD detection. This capability is particularly crucial for early-stage PD detection, as it enables the model to handle diverse and dynamic movement patterns commonly observed in real-life scenarios.

The high accuracy achieved by our model surpasses the performance of previous studies in PD classification using different types of data, such as typing, writing, and hand poses. This superiority can be attributed to the effectiveness of the Bi-LSTM architecture combined with the attention mechanism. The Bi-LSTM allows the model to capture temporal dependencies in the accelerometer data, while the attention mechanism focuses on relevant segments of the input, enhancing the model's ability to extract discriminative features.

The results of the 5-fold cross-validation demonstrate the model's stability and consistency in performance across different folds. The close alignment between training and validation accuracy curves suggests that our model generalizes well to unseen data, indicating its robustness and ability to

capture the underlying patterns in the accelerometer data. Moreover, the analysis of ROC curves and AUC values further supports the discriminative power of our model. The well-separated ROC curves and consistently high AUC values across the 5 folds indicate the model's ability to distinguish between PD patients and HC with high sensitivity and specificity. This implies that our model can effectively balance between correctly identifying PD patients and minimizing false positive predictions.

Comparing our results with previous works, our proposed model outperformed existing studies in PD classification. Specifically, in the work by [45], an LSTM model was applied to classify PD and HC based on typing data, achieving an accuracy of 0.73. Researchers in [46], utilized a 2-stacked LSTM model for PD detection during writing tasks, resulting in an accuracy of 0.91. Moreover, a CNN-BiLSTM model was employed in [50] for classifying PD based on writing data, achieving an accuracy of 0.98. Scientists in [33] used a DNN model for PD classification using hand pose data, obtaining an accuracy of 0.95. Comparatively, our proposed model exhibited superiority in terms of accuracy, precision, recall, specificity, and f1-score.

While our proposed model outperforms previous works, several limitations need to be acknowledged. Firstly, the participants' clinical and accelerometer data was derived from an open access public dataset comprised of PD patients with different characteristics in terms of disease severity/duration and levodopa equivalent dose [61]. Although all included patients had comparative tremor scores in the relevant UPDRS items ($\leq$ 1 in updrs16 and $\leq$ 2 in updrs20/updrs 21), patients may have not been similar regarding their disease severity, which is defined by other symptoms as well, like bradykinesia or gait impairment. It is, thus, not clear whether the participants' differentiation between PD patients and controls was based solely on the detection of slight symptoms of tremor, potentially macroscopically unperceived by human senses (both of the patient and of the examiner), or due to co-existent motor features, such as bradykinesia. Indeed, accelerometers capture motor performance collectively, based on the recorded acceleration of movements. In this particular experiment, data was collected for brief time periods while individuals were talking on their smartphone and their upper limb was relatively stable. Although a true state of immobility cannot be realistically achieved, it is considered less likely to capture symptoms of bradykinesia when no particular movements are actively performed (e.g. after a command). Moreover, bradykinesia, in contrast to rest or postural tremor, is usually detected from accelerometers when upper or lower limbs are actively moving [62,63]. We could, thus, argue that the selected group of patients was rather homogeneous from the scope of tremor characteristics.

The high heterogeneity of PD though cannot be captured by a single symptom and, in our case, tremor-free PD patients of the akinetic type are bound to be missed by such a detection system. In order to address this concern, we did a further analysis of a subset of PD patients who reported no symptoms of tremor (0 in updrs16). Interestingly, our model achieved high levels of PD detection in this subgroup of individuals as well, highlighting that the built-in accelerometer may be able to capture particular patterns of motor impairment that patients are not aware of. Although the results of this latter subanalysis should be received with skepticism due to the small sample size which can possibly introduce bias, high variance, limited complexity and generalization, they pave the way for future studies. Expanding the dataset size and diversity could enhance the generalizability of the model. Moreover, tremor tends to improve with dopaminergic therapy and occasionally decline as the disease progresses and symptoms of rigidity and bradykinesia predominate. Thus, low tremor scores may easily be encountered both in the early and advanced stage of PD; this cannot be inferred by UPDRS, which is mostly a descriptive tool. A dataset of drug naïve, early-stage PD patients, or even at risk individuals in the preclinical or prodromal stage of PD would give more focused and robust results considering early PD detection.

Additionally, tremor is a common symptom, which may be associated with numerous other neurological (e.g. ET) and systemic disorders (e.g. thyroid dysfunction), may also arise as a drug-related adverse event (e.g. valproic acid), or even be present in healthy individuals (physiological tremor) [64]. Including participants with such characteristics in future datasets would also be essential.

Incorporating other types of sensor data, such as voice recordings or gait analysis, could provide a more comprehensive and multimodal approach to PD detection. Finally, exploring the interpretability of the model's predictions and investigating the specific features learned by the attention mechanism could also provide valuable insights into the underlying mechanisms of PD.

6. Conclusions

In this study, we proposed a novel approach for detection of PD using accelerometer data collected during phone conversations from subjects demonstrating minor tremor of the upper limb, which might have gone undetected. The results of our experiments showcased the robustness and effectiveness of our model in differentiating PD patients from HC. Our findings may have significant implications for the field of PD research and healthcare. The ability to accurately detect subtle PD symptoms, such as tremor, could potentially facilitate the identification of PD patients in the early stage of the disease and prompt individuals at risk to seek medical help and proper consultation and eventually lead to its effective treatment. Such a possibility could also lead to effective patients' selection for clinical trials of new targeted, disease-modifying precision medicine therapies and for biomarkers development. Our proposed model provides a valuable contribution to this end, as mobile-based applications capturing data in the wild can revolutionize screening procedures for outpatients. Future research directions include expanding the dataset size and diversity, incorporating other sensor modalities, such as voice recordings or gait analysis, and investigating the interpretability of the model's predictions. These efforts may further enhance our understanding of the underlying mechanisms of PD.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

PD	Parkinson's disease
HC	Healthy control
ML	Machine Learning
DL	Deep Learning
AI	Artificial Intelligence
Bi-LSTM	Bidirectional Long Short-Term Memory
CNN	Convolutional Neural Network
DNN	Deep Neural Network
RNN	Recurrent Neural network
ANN	Artificial Neural Network
IQR	Interquartile Range
AUC	Area Under the Curve
ROC	Receiving Operating Characteristic

References

1. Parkinson disease.
2. Rossi, A.; Berger, K.; Chen, H.; Leslie, D.; Mailman, R.B.; Huang, X. Projection of the prevalence of Parkinson's disease in the coming decades: Revisited. *Movement Disorders: Official Journal of the Movement Disorder Society* **2018**, *33*, 156–159. https://doi.org/10.1002/mds.27063.
3. Lubomski, M.; Davis, R.L.; Sue, C.M. Health-related quality of life for parkinson's disease patients and their caregivers. *Journal of Movement Disorders* **2021**, *14*, 42–52. https://doi.org/10.14802/jmd.20079.
4. Duncan, G.W.; Khoo, T.K.; Yarnall, A.J.; O'Brien, J.T.; Coleman, S.Y.; Brooks, D.J.; Barker, R.A.; Burn, D.J. Health-related quality of life in early Parkinson's disease: the impact of nonmotor symptoms. *Movement Disorders: Official Journal of the Movement Disorder Society* **2014**, *29*, 195–202. https://doi.org/10.1002/mds.25664.
5. Dorsey, E.R.; Elbaz, A.; Nichols, E.; Abbasi, N.; Abd-Allah, F.; Abdelalim, A.; Adsuar, J.C.; Ansha, M.G.; Brayne, C.; Choi, J.Y.J.; et al Global, regional, and national burden of Parkinson's disease, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet Neurology* **2018**, *17*, 939–953. https://doi.org/10.1016/S1474-4422(18)30295-3.
6. Titova, N.; Chaudhuri, K.R. Personalized medicine and nonmotor symptoms in parkinson's disease. *International Review of Neurobiology* **2017**, *134*, 1257–1281. https://doi.org/10.1016/bs.irn.2017.05.015.
7. Armstrong, M.J.; Okun, M.S. Diagnosis and treatment of parkinson disease: a review. *JAMA* **2020**, *323*, 548–560. https://doi.org/10.1001/jama.2019.22360.
8. Ntetsika, T.; Papathoma, P.E.; Markaki, I. Novel targeted therapies for Parkinson's disease. *Molecular Medicine (Cambridge, Mass.)* **2021**, *27*, 17. https://doi.org/10.1186/s10020-021-00279-2.
9. Ye, H.; Robak, L.A.; Yu, M.; Cykowski, M.; Shulman, J.M. Genetics and pathogenesis of parkinson's syndrome. *Annual Review of Pathology* **2023**, *18*, 95–121. https://doi.org/10.1146/annurev-pathmechdis-031521-034145.
10. Emamzadeh, F.N.; Surguchov, A. Parkinson's disease: biomarkers, treatment, and risk factors. *Frontiers in Neuroscience* **2018**, *12*, 612. https://doi.org/10.3389/fnins.2018.00612.
11. Straccia, G.; Colucci, F.; Eleopra, R.; Cilia, R. Precision medicine in parkinson's disease: from genetic risk signals to personalized therapy. *Brain Sciences* **2022**, *12*, 1308. https://doi.org/10.3390/brainsci12101308.
12. Noyce, A.J.; Lees, A.J.; Schrag, A.E. The prediagnostic phase of Parkinson's disease. *Journal of Neurology, Neurosurgery & Psychiatry* **2016**, *87*, 871–878. https://doi.org/10.1136/jnnp-2015-311890.
13. Bloem, B.R.; Okun, M.S.; Klein, C. Parkinson's disease. *Lancet (London, England)* **2021**, *397*, 2284–2303. https://doi.org/10.1016/S0140-6736(21)00218-X.
14. Postuma, R.B.; Berg, D.; Stern, M.; Poewe, W.; Olanow, C.W.; Oertel, W.; Obeso, J.; Marek, K.; Litvan, I.; Lang, A.E.; et al. MDS clinical diagnostic criteria for Parkinson's disease: MDS-PD Clinical Diagnostic Criteria. *Movement Disorders* **2015**, *30*, 1591–1601. https://doi.org/10.1002/mds.26424.
15. Sauerbier, A.; Jenner, P.; Todorova, A.; Chaudhuri, K.R. Non motor subtypes and Parkinson's disease. *Parkinsonism & Related Disorders* **2016**, *22 Suppl 1*, S41–46. https://doi.org/10.1016/j.parkreldis.2015.09.027.
16. Tolosa, E.; Garrido, A.; Scholz, S.W.; Poewe, W. Challenges in the diagnosis of {Parkinson}'s disease. *The Lancet Neurology* **2021**, *20*, 385–397. https://doi.org/10.1016/S1474-4422(21)00030-2.
17. Pasquini, J.; Ceravolo, R.; Qamhawi, Z.; Lee, J.Y.; Deuschl, G.; Brooks, D.J.; Bonuccelli, U.; Pavese, N. Progression of tremor in early stages of Parkinson's disease: a clinical and neuroimaging study. *Brain: A Journal of Neurology* **2018**, *141*, 811–821. https://doi.org/10.1093/brain/awx376.
18. Dirkx, M.F.; Zach, H.; Bloem, B.R.; Hallett, M.; Helmich, R.C. The nature of postural tremor in Parkinson disease. *Neurology* **2018**, *90*, e1095–e1103. https://doi.org/10.1212/WNL.0000000000005215.
19. de Lau, L.M.L.; Koudstaal, P.J.; Hofman, A.; Breteler, M.M.B. Subjective complaints precede Parkinson disease: the rotterdam study. *Archives of Neurology* **2006**, *63*, 362–365. https://doi.org/10.1001/archneur.63.3.noc50312.
20. Marras, C.; Schüle, B.; Munhoz, R.P.; Rogaeva, E.; Langston, J.W.; Kasten, M.; Meaney, C.; Klein, C.; Wadia, P.M.; Lim, S.Y.; et al. Phenotype in parkinsonian and nonparkinsonian LRRK2 G2019S mutation carriers. *Neurology* **2011**, *77*, 325–333. https://doi.org/10.1212/WNL.0b013e318227042d.
21. Skaramagkas, V.; Pentari, A.; Kefalopoulou, Z.; Tsiknakis, M. Multi-modal deep learning diagnosis of parkinson's disease—a systematic review. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **2023**, *31*, 2399–2423. https://doi.org/10.1109/TNSRE.2023.3277749.

22. Yang, Y.; Yuan, Y.; Zhang, G.; Wang, H.; Chen, Y.C.; Liu, Y.; Tarolli, C.G.; Crepeau, D.; Bukartyk, J.; Junna, M.R.; et al. Artificial intelligence-enabled detection and assessment of Parkinson's disease using nocturnal breathing signals. *Nature Medicine* **2022**, *28*, 2207–2215. https://doi.org/10.1038/s41591-022-01932-x.

23. Tsakanikas, V.; Ntanis, A.; Rigas, G.; Androutsos, C.; Boucharas, D.; Tachos, N.; Skaramagkas, V.; Chatzaki, C.; Kefalopoulou, Z.; Tsiknakis, M.; et al. Evaluating gait impairment in parkinson's disease from instrumented insole and imu sensor data. *Sensors* **2023**, *23*, 3902. https://doi.org/10.3390/s23083902.

24. Chatzaki, C.; Skaramagkas, V.; Tachos, N.; Christodoulakis, G.; Maniadi, E.; Kefalopoulou, Z.; Fotiadis, D.I.; Tsiknakis, M. The Smart-Insole Dataset : Gait Analysis Using Wearable Sensors with a Focus on Elderly and Parkinson ' s Patients. *Sensors (Switzerland)* **2021**, *21*, 1–22.

25. Skaramagkas, V.; Andrikopoulos, G.; Kefalopoulou, Z.; Polychronopoulos, P. Towards differential diagnosis of essential and parkinson's tremor via machine learning. *2020 28th Mediterranean Conference on Control and Automation, MED 2020* **2020**, pp. 782–787. https://doi.org/10.1109/MED48518.2020.9182922.

26. Goschenhofer, J.; Pfister, F.M.; Yuksel, K.A.; Bischl, B.; Fietzek, U.; Thomas, J. Wearable-based Parkinson's Disease Severity Monitoring using Deep Learning. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* **2019**, *11908 LNAI*, 400–415. https://doi.org/10.1007/978-3-030-46133-1_24.

27. Kefalopoulou, Z.; Chatzaki, C.; Skaramagkas, V.; Chroni, E.; Tachos, N.; Fotiadis, D.; Tsiknakis, M. Pressure Sensor Insole Gait Assessment for Parkinson's Disease patients: A pilot study. In Proceedings of the MOVEMENT DISORDERS. WILEY 111 RIVER ST, HOBOKEN 07030-5774, NJ USA, 2022, Vol. 37, pp. S193–S193.

28. Chatzaki, C.; Skaramagkas, V.; Kefalopoulou, Z.; Tachos, N.; Kostikis, N.; Kanellos, F.; Triantafyllou, E.; Chroni, E.; Fotiadis, D.I.; Tsiknakis, M. Can Gait Features Help in Differentiating Parkinson's Disease Medication States and Severity Levels? A Machine Learning Approach. *Sensors* **2022**, *22*, 9937. https://doi.org/10.3390/s22249937.

29. Skaramagkas, V.; Andrikopoulos, G.; Kefalopoulou, Z.; Polychronopoulos, P. A Study on the Essential and Parkinson's Arm Tremor Classification. *Signals 2021, Vol. 2, Pages 201-224* **2021**, *2*, 201–224. https://doi.org/10.3390/SIGNALS2020016.

30. Lu, M.; Zhao, Q.; Poston, K.L.; Sullivan, E.V.; Pfefferbaum, A.; Shahid, M.; Katz, M.; Kouhsari, L.M.; Schulman, K.; Milstein, A.; et al. Quantifying Parkinson's disease motor severity under uncertainty using MDS-UPDRS videos. *Medical Image Analysis* **2021**, *73*, 102179. https://doi.org/10.1016/j.media.2021.102179.

31. Tsakanikas, V.D.; Dimopoulos, D.G.; Tachos, N.S.; Chatzaki, C.; Skaramagkas, V.; Christodoulakis, G.; Tsiknakis, M.; Fotiadis, D.I. Gait and balance patterns related to Free-Walking and TUG tests in Parkinson's Disease based on plantar pressure data. In Proceedings of the 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2021, pp. 236–239. https://doi.org/10.1109/EMBC46164.2021.9629637.

32. Rajkomar, A.; Dean, J.; Kohane, I. Machine learning in medicine. *The New England Journal of Medicine* **2019**, *380*, 1347–1358. https://doi.org/10.1056/NEJMra1814259.

33. Fraiwan, L.; Khnouf, R.; Mashagbeh, A.R. Parkinson's disease hand tremor detection system for mobile application. *Journal of Medical Engineering & Technology* **2016**, *40*, 127–134. https://doi.org/10.3109/03091902.2016.1148792.

34. Castrillón, R.; Acien, A.; Orozco-Arroyave, J.R.; Morales, A.; Vargas-Bonilla, J.F.; Vera-Rodríguez, R.; Fierrez, J.; Ortega-Garcia, J.; Villegas, A. Characterization of the Handwriting Skills as a Biomarker for Parkinson's Disease. *2019 14th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2019)* **2019**, pp. 1–5.

35. Bazgir, O.; Frounchi, J.; Habibi, S.A.H.; Palma, L.; Pierleoni, P. A neural network system for diagnosis and assessment of tremor in Parkinson disease patients. *2015 22nd Iranian Conference on Biomedical Engineering, ICBME 2015* **2016**, pp. 1–5. https://doi.org/10.1109/ICBME.2015.7404105.

36. Wu, H.; Zhang, Y.; Wu, X.; Yang, F. Assessment of Upper Limb Tremors in Patients with Parkinson's Disease Based on Displacement and Acceleration Information. *Proceedings - 5th International Conference on Automation, Control and Robotics Engineering, CACRE 2020* **2020**, pp. 177–182. https://doi.org/10.1109/CACRE50138.2020.9230024.

37. Lekadir, K.; Osuala, R.; Gallin, C.; Lazrak, N.; Kushibar, K.; Tsakou, G.; Aussó, S.; Alberich, L.C.; Marias, K.; Tsiknakis, M.; et al. Future-ai: guiding principles and consensus recommendations for trustworthy artificial intelligence in medical imaging, 2021. arXiv:2109.09658 [cs], https://doi.org/10.48550/arXiv.2109.09658.

38. Xing, X.; Luo, N.; Li, S.; Zhou, L.; Song, C.; Liu, J. Identification and Classification of Parkinsonian and Essential Tremors for Diagnosis Using Machine Learning Algorithms. *Frontiers in Neuroscience* **2022**, *16*. https://doi.org/10.3389/FNINS.2022.701632.

39. Tong, L.; He, J.; Peng, L. CNN-Based PD Hand Tremor Detection Using Inertial Sensors. *IEEE Sensors Letters* **2021**, *5*. https://doi.org/10.1109/LSENS.2021.3074958.

40. Sun, M.; Watson, A.; Blackwell, G.; Jung, W.; Wang, S.; Koltermann, K.; Helm, N.; Zhou, G.; Cloud, L.; Pretzer-Aboff, I. TremorSense: Tremor Detection for Parkinson's Disease Using Convolutional Neural Network. *Proceedings - 2021 IEEE/ACM Conference on Connected Health: Applications, Systems and Engineering Technologies, CHASE 2021* **2021**, pp. 1–10. https://doi.org/10.1109/CHASE52844.2021.00009.

41. Folador, J.P.; Santos, M.C.S.; Luiz, L.M.D.; de Souza, L.A.P.S.; Vieira, M.F.; Pereira, A.A.; de Oliveira Andrade, A. On the use of histograms of oriented gradients for tremor detection from sinusoidal and spiral handwritten drawings of people with Parkinson's disease. *Medical & biological engineering & computing* **2021**, *59*, 195–214. https://doi.org/10.1007/S11517-020-02303-9.

42. Kim, H.B.; Lee, W.W.; Kim, A.; Lee, H.J.; Park, H.Y.; Jeon, H.S.; Kim, S.K.; Jeon, B.; Park, K.S. Wrist sensor-based tremor severity quantification in Parkinson's disease using convolutional neural network. *Computers in biology and medicine* **2018**, *95*, 140–146. https://doi.org/10.1016/J.COMPBIOMED.2018.02.007.

43. Kaushik, S.; Choudhury, A.; Sheron, P.K.; Dasgupta, N.; Natarajan, S.; Pickett, L.A.; Dutt, V. Ai in healthcare: time-series forecasting using statistical, neural, and ensemble architectures. *Frontiers in Big Data* **2020**, *3*.

44. Absar, N.; Uddin, N.; Khandaker, M.U.; Ullah, H. The efficacy of deep learning based LSTM model in forecasting the outbreak of contagious diseases. *Infectious Disease Modelling* **2021**, *7*, 170–183. https://doi.org/10.1016/j.idm.2021.12.005.

45. Thummikarat, H.; Chongstitvatana, P. An implementation of machine learning for parkinson's disease diagnosis. *ECTI-CON 2021 - 2021 18th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology: Smart Electrical System and Technology, Proceedings* **2021**, pp. 258–261. https://doi.org/10.1109/ECTI-CON51831.2021.9454784.

46. A V, A.S.; Lones, M.A.; Smith, S.L.; Vallejo, M. Evaluation of Recurrent Neural Network Models for Parkinson's Disease Classification Using Drawing Data. In Proceedings of the 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2021, pp. 1702–1706. https://doi.org/10.1109/EMBC46164.2021.9630106.

47. Hssayeni, M.D.; Jimenez-Shahed, J.; Burack, M.A.; Ghoraani, B. Wearable Sensors for Estimation of Parkinsonian Tremor Severity during Free Body Movements. *Sensors (Basel, Switzerland)* **2019**, *19*. https://doi.org/10.3390/S19194215.

48. Hathaliya, J.J.; Modi, H.; Gupta, R.; Tanwar, S. Deep learning and Blockchain-based Essential and Parkinson Tremor Classification Scheme. *IEEE INFOCOM 2022 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)* **2022**, pp. 1–6. https://doi.org/10.1109/INFOCOMWKSHPS54753.2022.9798053.

49. Shahtalebi, S.; Atashzar, S.F.; Patel, R.V.; Jog, M.S.; Mohammadi, A. A deep explainable artificial intelligent framework for neurological disorders discrimination. *Scientific reports* **2021**, *11*. https://doi.org/10.1038/S41598-021-88919-9.

50. Taleb, C.; Likforman-Sulem, L.; Mokbel, C.; Khachab, M. Detection of Parkinson's disease from handwriting using deep learning: a comparative study. *Evolutionary Intelligence* **2020**. https://doi.org/10.1007/S12065-020-00470-0.

51. Taleb, C.; Khachab, M.; Mokbel, C.; Likforman-Sulem, L. Feature selection for an improved Parkinson's disease identification based on handwriting. *1st IEEE International Workshop on Arabic Script Analysis and Recognition, ASAR 2017* **2017**, pp. 52–56. https://doi.org/10.1109/ASAR.2017.8067759.

52. Papadopoulos, A.; Kyritsis, K.; Klingelhoefer, L.; Bostanjopoulou, S.; Chaudhuri, K.R.; Delopoulos, A. Detecting parkinsonian tremor from imu data collected in-the-wild using deep multiple-instance learning. *IEEE journal of biomedical and health informatics* **2020**, *24*, 2559–2569. https://doi.org/10.1109/JBHI.2019.2961748.

53. Chan, P.Y.; Ripin, Z.M.; Halim, S.A.; Arifin, W.N.; Yahya, A.S.; Eow, G.B.; Tan, K.; Hor, J.Y.; Wong, C.K. Motion characteristics of subclinical tremors in Parkinson's disease and normal subjects. *Scientific Reports* **2022**, *12*, 4021. https://doi.org/10.1038/s41598-022-07957-z.

54. Oppenheim, A.; Schafer, R. *Discrete-time signal processing: pearson new international edition*, 3rd edition ed.; Pearson Education Limited: Harlow, 2013.

55. Yamansavaşçılar, B.; Güvensan, M.A. Activity recognition on smartphones: efficient sampling rates and window sizes. In Proceedings of the 2016 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops), 2016, pp. 1–6. https://doi.org/10.1109/PERCOMW.2016.7457154.

56. Djordjevic, I. Chapter 3 - quantum circuits and quantum information processing fundamentals. In *Quantum Information Processing and Quantum Error Correction*; Djordjevic, I., Ed.; Academic Press: Oxford, 2012; pp. 91–117. https://doi.org/10.1016/B978-0-12-385491-9.00003-4.

57. Akandeh, A.; Salem, F.M. Slim lstm networks: lstm_6 and lstm_c6. In Proceedings of the 2019 IEEE 62nd International Midwest Symposium on Circuits and Systems (MWSCAS), 2019, pp. 630–633. ISSN: 1558-3899, https://doi.org/10.1109/MWSCAS.2019.8884912.

58. Siami-Namini, S.; Tavakoli, N.; Namin, A.S. The performance of lstm and bilstm in forecasting time series. In Proceedings of the 2019 IEEE International Conference on Big Data (Big Data), 2019, pp. 3285–3292. https://doi.org/10.1109/BigData47090.2019.9005997.

59. Graves, A.; Fernández, S.; Schmidhuber, J. Bidirectional lstm networks for improved phoneme classification and recognition. In Proceedings of the Artificial Neural Networks: Formal Models and Their Applications – ICANN 2005; Duch, W.; Kacprzyk, J.; Oja, E.; Zadrożny, S., Eds.; Springer: Berlin, Heidelberg, 2005; Lecture Notes in Computer Science, pp. 799–804. https://doi.org/10.1007/11550907_126.

60. Afendras, G.; Markatou, M. Optimality of training/test size and resampling effectiveness in cross-validation. *Journal of Statistical Planning and Inference* **2019**, *199*, 286–301. https://doi.org/https://doi.org/10.1016/j.jspi.2018.07.005.

61. Papadopoulos, A.; Iakovakis, D.; Klingelhoefer, L.; Bostantjopoulou, S.; Chaudhuri, K.R.; Kyritsis, K.; Hadjidimitriou, S.; Charisis, V.; Hadjileontiadis, L.J.; Delopoulos, A. Unobtrusive detection of Parkinson's disease from multi-modal and in-the-wild sensor data using deep learning techniques. *Scientific Reports* **2020**, *10*, 21370. https://doi.org/10.1038/s41598-020-78418-8.

62. Habets, J.G.V.; Herff, C.; Kubben, P.L.; Kuijf, M.L.; Temel, Y.; Evers, L.J.W.; Bloem, B.R.; Starr, P.A.; Gilron, R.; Little, S. Rapid dynamic naturalistic monitoring of bradykinesia in parkinson's disease using a wrist-worn accelerometer. *Sensors (Basel, Switzerland)* **2021**, *21*, 7876. https://doi.org/10.3390/s21237876.

63. Mahadevan, N.; Demanuele, C.; Zhang, H.; Volfson, D.; Ho, B.; Erb, M.K.; Patel, S. Development of digital biomarkers for resting tremor and bradykinesia using a wrist-worn wearable device. *NPJ Digital Medicine* **2020**, *3*, 5. https://doi.org/10.1038/s41746-019-0217-7.

64. Bhidayasiri, R. Differential diagnosis of common tremor syndromes. *Postgraduate Medical Journal* **2005**, *81*, 756–762. https://doi.org/10.1136/pgmj.2005.032979.