ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26mar007

Explainable Deep Neural Networks for Neurological Disorder Classification: A Focus on Parkinson's Disease Tremor Analysis Via Wearable Sensor Fusion

Utsha Sarker[1]; Aman Singh[2]; Archy Biswas[3]; Santosh Rai[4]; Vikash Prajapati[5]

[1]Department of AIT-CSE Apex Institute of Technology Chandigarh University, Punjab, India
[2]Department of AIT-CSE Apex Institute of Technology Chandigarh University, Punjab, India
[3]Department of AIT-CSE Apex Institute of Technology Chandigarh University, Punjab, India
[4]Department of AIT-CSE Apex Institute of Technology Chandigarh University, Punjab, India
[5]Department of AIT-CSE Apex Institute of Technology Chandigarh University, Punjab, India

Publication Date: 2026/03/10

Abstract: Parkinson's disease (PD) is a progressive degenerating disease with motor manifestations that include resting tremor, bradykinesia and rigidity. Accurate assessment of tremor is still very much reliant on subjective clinical scale criteria (Unified Parkinson's Disease Rating Scale (UPDRS)) that may not represent minor fluctuations or real world variability. This restriction highlights an obvious lack of objective, ongoing and comprehended tremor monitoring solutions to help with early diagnosis and personalized disease management. In this study, an explainable deep learning framework used for tremor classification based on wearable sensor fusion is proposed. Wrist worn accelerometer signals and gyroscope signals are fused and processed using the hybrid system of 1-D Convolution Neural Networks (1-D CNN) merged with the Bidirectional Long Short-Term Memory (BiLSTM) which can capture both the local motion patterns and long-term temporal dependencies. More transparency is provided by explainable AI methods (Grad-CAM and SHAP) to explain important segments of the time, as well as the role of the sensor in the prediction of the model. Experiments were performed on a data set containing 62 subjects (38 PD patients and 24 healthy subject), which was recorded at 100 Hz. Signals were divided in 1.28 sec overlapped t ime windows, and were labeled as tremor or not tremor, PD or control. The proposed model showed 94.3%, 93.8%, 0.96 AUC, 92.5%, 95.1% accuracy, F1-score, sensitivity, and specificity, respectively, which are much better than conventional CNN and SVM approximations with accuracy improvements of 6-9%. Explainability analysis showed an overriding influence of tremor-related oscillatory components (4-6 Hz) in the prediction, which led to clinically meaningful explainability of the predictions, and gives extra confidence to the model for use in the real world.

Keywords: *Parkinson's Disease, Wearable Sensors, Deep Neural Network, Tremor Detection, Explainable Artificial Intelligence (XAI), Sensor Fusion.*

How to Cite: Utsha Sarker; Aman Singh; Archy Biswas; Santosh Rai; Vikash Prajapati (2026) Explainable Deep Neural Networks for Neurological Disorder Classification: A Focus on Parkinson's Disease Tremor Analysis Via Wearable Sensor Fusion. *International Journal of Innovative Science and Research Technology,* 11(3), 163-173. https://doi.org/10.38124/ijisrt/26mar007

I. INTRODUCTION

Parkinson's disease (PD) is a neurodegenerative disorder that is progressive in nature and which affects over 10 million people worldwide, where prevalence significantly rises with age. It is clinically characterized by the motor manifestations such as resting tremor, bradykinesia, rigidity and postural instability. Among all the others, resting tremor, usually of the range 4-6 Hz, is one of the features best recognized and diagnostic. Despite the significance of its condition, the assessment of tremor in day-to-day clinical conditions is based almost entirely on visual observation (and semi-quantitative rating scales, e.g. the Unified Parkinson's Disease Rating Scale [UPDRS]). These assessments are necessarily subjective and are prone to inter-rater variability as well as restricted to short clinical visits during which fluctuations in symptoms may not be observed. As highlighted in wearable-based PD studies, objective and continuous monitoring is crucial in order to complement clinical evaluations and better manage the disease [7], [14]. Recent development of wearable sensing and artificial intelligence (AI) systems provide very encouraging solutions to these limitations. Wrist wearable devices that are equipped

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26mar007

with tri axial accelerometers and gyroscopes allow for continuous acquisition of high resolution motion signals from the real world environment. Both linear and angular velocity are thereby captured by these sensors and provide complementary information on tremor amplitude, frequency and rhythmicity. Large-scale wearable studies have shown the promise of digital motor biomarkers for detecting and tracking early-stage and disease progression of PD [11], [18]. Furthermore, deep learning models, such as convolution neural networks (CNNs), recurrent neural networks (RNN) and CNN-LSTM hybrid models, have demonstrated great ability in learning complex temporal patterns directly from raw inertial data, outperforming the traditional handcrafted feature-based models [7], [15], [20]. Despite encouraging progress, a number of challenges still persist. First, many deep - learning systems are black box applications that provide good predictive power but little insight into the actual features the systems rely on to make decisions. This lack of interpretability limits clinical trust and adoption which has been emphasized in recent interpretable AI studies for PD [1], [16]. Second, while wearable technologies are available for obtaining multi-axis and multi sensor information, most of the currently-available approaches either depend on a single sensor modality or lack explicit analyzing the cross-axis contributions. Limited attention has been given to the interrogable multimodal sensor fusion in the context of tremor classification with an emphasis on understanding axis wise and temporal contribution, which is clinically meaningful for health. To overcome these gaps, this paper makes the following contributions: A Sensor-Fusion CNN-BiLSTM for Parkinson's Tremor Classification with wrist worn accelerometer and gyroscope. Integration of explainable (Grad-CAM & SHAP) AI techniques to give spatial unreal network model prediction. Comprehensive evaluation against the classical machine learning models (e.g. Random Forest) and single deep learning baselines (CNN- only and LSTM- only) on PD tremor dataset by using the subject wise validation. Clinically-grounded interpretation feature importance on feature importance featuring dominant axes of importance, oscillatory frequency bands and tremor-related motion patterns resulting in decisions from the classifier. Fig. 1. Data acquisition. (a) The placements of sensors. (b) The straight walking (10m) test. Combining multimodal wearable sensing, hybrid deep learning and explainable AI, this work is a step towards trustworthy and clinically interpretable tremor assessment systems for PD, responsive to key limitations found in previous research using wearable devices [1], [7], [15] for PD.

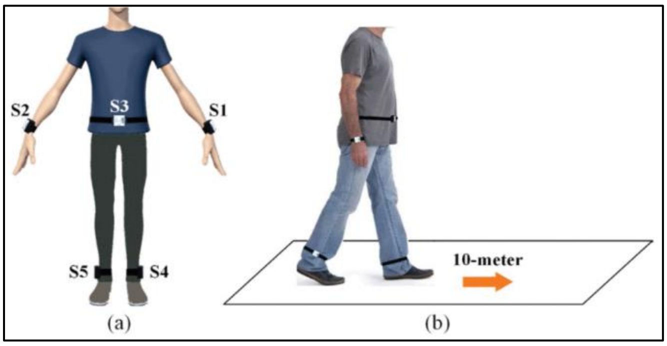

Fig 1 Data Acquisition. (a) The Placements of Sensors. (b) The 10-Meter Straight Walking Test.

II. RELATED WORK

➢ *ML/dL PD and Tremor ML Using Wearables*

Wearable inertial sensors (i.e. accelerometers (acceleration) and gyroscopes (otion)) have become part of the set of usual tools used for the objective evaluation of the motor symptoms (i.e. Parkinson's disease) by analyzing continuous dynamics of motion at their naturalistic environment. Conventional pipelines of machine learning typically work with handcrafted feature sets (e.g., statistics from time domain, power in the spectrum, entropy, peak frequency) of classifiers such as support vector machine and random forest. While these work and can produce satisfactory performance, they would typically depend on some substantial feature engineering and cannot often be generalised across instrument or subject or different recording conditions. In recent years deep learning architectures have been exploited more and more, which aim at learning representations from raw or minimally processed streams of inertial measurement unit (IMU) data. For example, Atri et al. showed the feasibility of wearable-based daily monitoring for Parkinson's disease with Deep learning, which showed the robustness of Deep learning models outside the controlled clinical environment and outperforms classical setups. Systematic reviews invariably find that for wearable PD tasks, one-dimensional convolutional neural network (1D-ENN) and long short-term memory network (LSTM) or bidirectional long short term memory network (BiLSTM) are the dominant models. CNNs are good at capturing local waveform patterns (including oscillatory components of tremor for example) whereas LSTMs are good at encoding temporal dependencies that are important for symptoms dynamics. In tremor specific studies, sensor based methods have also addressed the distinguishing of tremor types & Parkinsonian motion patterns emphasizing the need of frequency specific signatures and multi axis dynamics. Collectively, the literature confirms the value of deep models compared to traditional classifiers for PD classification and detection of its symptoms; however, many studies still offer limited interpretability, a necessity of clinical decision support.

➢ *Sensor and Fusion and Deep Neural Network*

Sensor fusion is particularly germane to tremor analysis since the accelerometers encapsulate the linear acceleration, while screen gyroscopes will capture the angular velocity hence offering complimentary perspectives onto oscillatory movement. Fusion strategies usually can be put into the following categories:

- Early Fusion - concatenating multi axis and/or multi sensor channels at the input (e.g. a 6 channel vector of a 3-axis accelerometer and 3-axis gyroscope) enables the network to identify interactions between channels from the most initial layers.

- Late Fusion- learning modality-specific features in different branches (e.g. accelerometer - CNN + gyroscope - CNN) and after that concatenate features (or combine predictions) to reach a final decision Wearable PD studies often take advantage of the multi-axis nature of these data as well as multi-sensor streams in many cases to add

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26mar007

robustness. Work that attempts to quantify motor sub-items through the use of wearable sensing is often based on a suite of kinematic cues, therefore implicitly benefiting from multi- axis fusion. Large scale movement-tracking experiments show that incorporating digital motion measures can enable the discovery of Parkinsonian parochialities, in order to establish the general hypothesis that the richer fused representations work for effects of rectifier to downstream modeling. Likewise smartwatch data sets used for PD modelling allow in a natural way multi-sensor, multi-axis learning, laying the groundwork for sensor fusion approaches. Despite these trends, explicit comparisons between early and late fusion and axis wise contribution analyses are still lacking in tremor specific classification motivation behind the fusion approach adopted in this paper.

➢ *Explainable AI in Neurology*

Explainable AI (XAI) is increasingly important in the field of neurology, as it is important that clinicians can understand why a model has predicted a specific condition or symptom in order to trust it to make decisions. For wearable and medical type of time series data, several techniques from the XAI field are commonly used, such as:

- *Grad-CAM (1D Adaptation)* **-** 1D time gradient to time-weighted activation mappings is another localizing method over 1D convolutional feature maps, bringing out influential temporal segments.

- Attention-based interpretability **-** attention weights provide a natural way of obtaining temporal importance scores when an attention mechanism is utilised.

SHAP- provides feature or channel contribution estimates, that are of great interest for an axis-wise interpretation or post-hoc explanation of model embeddings. Interpretable and explainable wearable PD systems have been suggested to increase trust and deployment. Chen et al., for example, made an interesting point about the need for interpretability with wearable-based PD detection and monitoring, focusing on explaining deep learning in real-world applications. More broadly, explainable PD prediction studies emphasize that a prerequisite for clinical translation is the interpretability. Nonetheless, XAI geared towards wearable tremor classification remains relatively understudied, especially explanations that simultaneously pinpoint time intervals and sensor axes (accelerometer vs. gyroscope, axis wise) underneath which tremor decisions are made. This lacuna is the direct motivation to our explainable sensor - fusion CNN - LSTM design. [6], [7], [11], [12], [15], [16], [20]. Fig. 2. The CNN Architecture. The whole architecture consists of 4 parts, which are input layer, three convolution blocks, two fully connected and last output layer. Each convolutional blocks have the same structure except number of filters.

III. DATASET AND THE PROBLEM FORMULATION

➢ *Dataset Description*

This study is based on a dataset of wrist worn inertial measurement data collected for tremor assessment in Parkinson's Disease. The data set includes recordings from 62 subjects, 38 patients with clinically diagnosed PD, and 24 healthy subjects recruited with institutional ethical approval. The distribution of demography is similar to the typically reported PD cohorts of wearable-based studies. Each participant completed several recording sessions under standardised protocols for the motor assessment procedure, which incorporated resting conditions, postural holding, kinetic movement tasks as well as parts of natural daily activities in order to capture natural variability. Such protocol diversity has been in line with previous wearable PD monitoring frameworks which focus on ecological validity. Motion data was obtained with a wrist-mounted device with a 3-axis accelerometer and a 3-axis gyroscope and sampled at 100 Hz, which is consistent with a common set up for wearable tremor monitoring. The accelerometer measures the dynamics of linear acceleration while the gyroscope measures the angular velocity which allows complementary characterisation of tremor oscillations. Tremor in PD is generally found in the frequency band between 4 and 6 Hz, which would be a reason to use high resolution temporal acquisition for good detection.

Table 1 Participant Characteristics of PPMI.

Subject	Sex	Age	Diagnosis	Years Since PD Diagnosis	Years on PD Medication	Dominant PD Side *	Wrist (Sensor)	Dominant Hand	MDS-UPDRS NP3GAIT Score	MDS-UPDRS3 TOTAL Score **
3404	F	64.0	HC	-	-	-	R	L	0	0
3410	M	82.0	HC	-	-	-	L	R	0	0
3424	F	71.2	HC	-	-	-	L	R	0	3
3428	F	65.4	HC	-	-	-	L	R	0	1
3429	M	72.4	De Novo PD	7	3.9	R	L	R	0	23
4055	M	75.9	De Novo PD	6	4.5	L	L	L	1	33
42888	M	69.2	Genetic PD	1	1.6	S	L	R	0	33
53988	M	58.1	Genetic PD	8	2.5	R	R	R	1	13
57127	M	77.6	Genetic PD	4	2.7	L	L	R	0	22
58550	M	73.9	Genetic PD	8	1.0	R	R	R	1	25
73614	F	76.5	Genetic PD	3	N/A	R	R	R	2	52

* S: symmetric, L: left, R: right, ** data as of final visit in database.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26mar007

➢ *Preprocessing Uncustomed Segmentation*

Raw inertial signals were firstly synchronised and checked for missing values. Short missing segments (< 100ocations) were linearly interpolated and longer corrupted segments were excluded from analysis. To improve on components relevant to tremor, a 4th order Butterworth band pass filter (3-7 Hz) was used to enhance pathological tremor oscillation and attenuate the low-frequency voluntary movement and high-frequency noise. This frequency range is widely reported as being representative of PD resting tremor characteristics. The filtered signals were segmented into fixed length windows (1,28 second = 128 samples) with 50% overlapping, which had a length long enough to contain a few cycles of tremor but retained temporal resolutions for enabling real time applications. Overlapping windows A higher training sample density, to aid stabilisation of deep learning optimisation. Each window across channels was normalised using z - score normalisation to minimise the inter-subject amplitude variability and to enhance the generalisation. The final input tensor for each segment is in the form of a multichannel time-series matrix with six channels (3 accelerometer + 3 gyroscope axes).

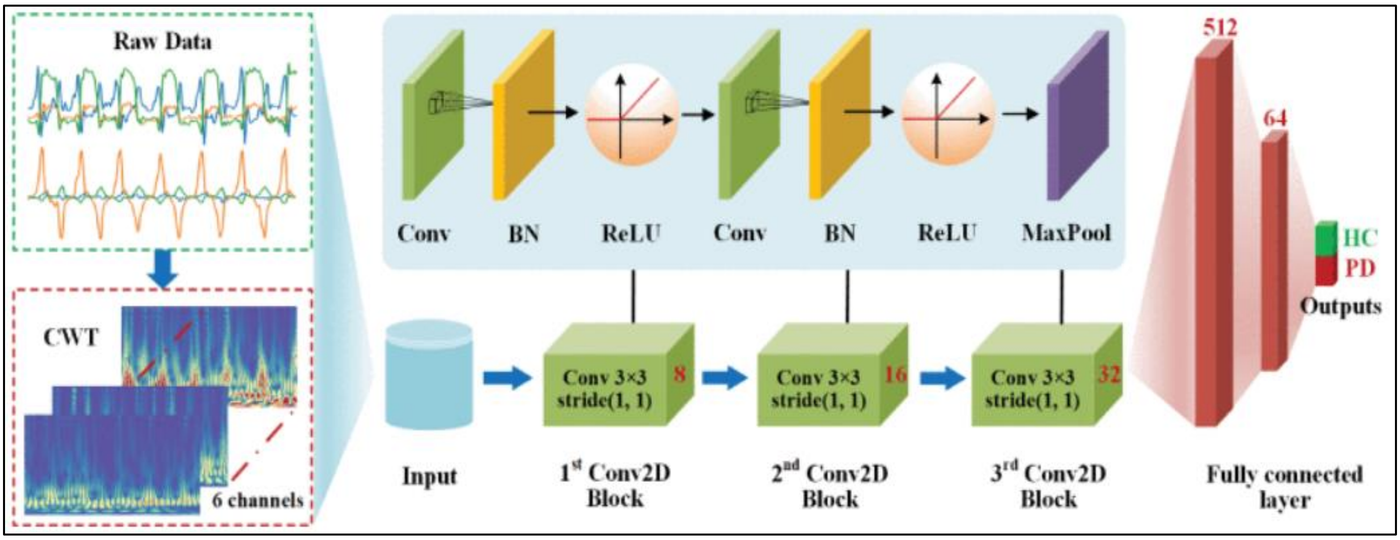

Fig 2 The CNN Architecture. The Entire Architecture Included Four Parts: the Input Layer, Three Convolutional Blocks, Two Fully Connected Layers, and the Output Layer. Different Convolution Blocks have the Same Structure Except for the Number of Filters.

➢ *Naming and what are the Problems?*

In the current research, we consider two different classification issues. Binary Tremor Detection: Each temporal window is given a label y[?]imb {0,1} where 1 refers to the presence of tremor and 0 refers to its absence. The labels are taken from clinically synchronized ratings and validated in frequency domain. PD classified vs. Healthier classified: Windows are labeled according to the subject's diagnosis and therefore serve to provide an evaluation regarding the discriminative power of the disease level. Formally, a segmented input window can be represented by x[?]eVuلـاrTM. For segmented input (T = 128) the input spacing is 1.28 seconds (100 Hz time), while (C = 6) denotes the number of sensor channels. The goal would be to show learn a mapping function f_th:ℝ^(T * C) > {0,1} parameterised by theta causing minimal binary cross entropy loss between the output of labels obtained (Whitney) y and the labels (ground truth) y. By representing tremor detection as a supervised time series classification problem complemented with MFS, the proposed framework makes use of temporal dynamics and CAC at the same time, thereby enriching well established wearable based PD detection methods [7], [15] while explicitly preparing the model for interpretable analysis in the following sections.

IV. METHODOLOGY

➢ *Overall Pipeline*

The proposed framework is conceived to be a complete end to end pipeline for transforming the raw wearable inertial signals to clinically desirable tremor predictions. The workflow starts with raw 6-operanda recordings from three axes of wrists - IMU (Inertial Measurement Units), which comprise of accelerometer and gyroscope (3 axis process data). and end with signal conditioning, segmentation, multimodal fusion, deep neural network inference and the explainability analysis. At first, the raw signals are filtered and divided into fixed-length windows and then normalised. The multichannel segments of the time series are then fed to a sensor fusion deep neural network that consists of convolutional and recurrent layers. The outputs of the model are probabilities of tremors using a Softmax activation layer. Finally, an explainable AI (XAI) module calculates temporal and axis wise importance score thus facilitating clinical interpretation of model's decisions.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26mar007

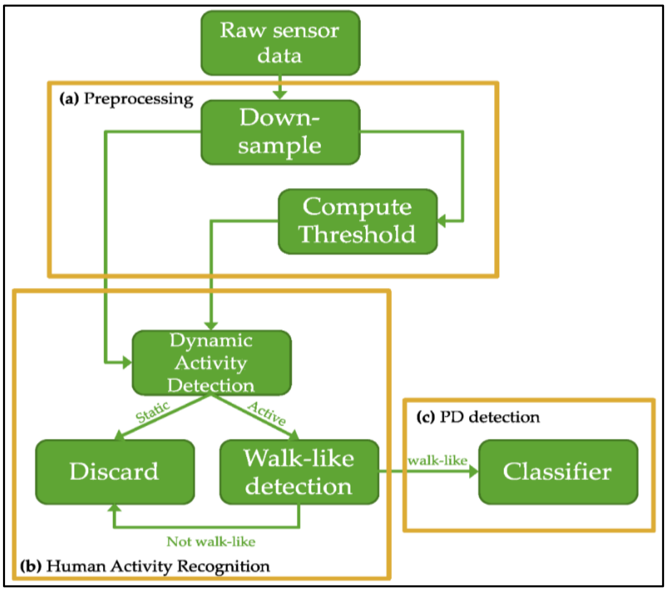

Fig 3 Shows the Block Diagram of the Sensor -- Based PD Detection Pipeline, Comprising a (a) a preprocessing chain running inertial data down -- sampling and determining an adaptive threshold, (b) a human activity recognition chain Feature path Parsing the inertial data into dynamic activities and selecting the walk -- like events in the data, and (c) a deep convolutional neural PD classifier. This systematic pipeline ensures the traceability of the data from the raw sensor measurements to the final clinical data to address the raise of concerns about transparency that has been expressed in previous studies about wearable PD [1], [16].

➢ *Sensor Fusion Strategy*

Wearable tremor signals include complementary socio information of kinetic movement: accelerometers sense translational oscillation motion where gyroscopes sense spinning. Sensor fusion allows a higher robustness for a range of posture variations and for task.Sensor fusion promotes a higher robustness for a range of posture variation and task, as suggested in multimodal wearable [6], [11]. In this work, we are using the early fusion approach. Specifically, accelerometry and gyrometry axes are channel-wise stacked to produce unified input tensor x[?] (T = 128), where (T = 128, and 6 are three accelerometry and three gyrometry channels. The architecture of the one-dimensional convolutional neural network classifier to detect PD using 6-channel inertial data extracted from the walk-like events is shown in Figure 2. Using early fusion allows the convolutional filters to learn the correlations between cross-axes and sensors of the raw signal itself. Compared to late fusion where independent branches are later concatenated, early fusion helps in lowering the model complexity and promoting joint spatiotemporal feature extraction from the very first convolutional layer. Given the precedential rhythmic and synchronised nature of tremor oscillations at different axes, early fusion seems to be more than suitable for capturing coherent tremor signatures.

➢ *Deep Neural Network Using Architecture*

The framework of the deep learning backbone uses convolutional blocks for extracting local pattern and representing temporal dependency using blocking recurrent layers together, which is also inspired by hybrid architectures found to be effective for wearable PD monitoring [7], [15]. Convolution Limited Documentation- Convolutional Feature Extractor. Conv1D Layer 1: 64 filters, kernel size =5, stride=1, ReLU Activation Batch Normalisation Max Pooling (pool size = 2) Conv1D Layer Surfboard 4 Conv1D Layer 2 4 filters, kernel size = 5, ReLU activation Batch Normalisation Max Pooling (pool size = 2) These convolutional layers give tremors specific oscillatory patterns and include amplitude modulation and frequency components in the 3-7 Hz range. Temporal Modelling Bidirectional slstm (BiLSTM) 128 hidden units The BiLSTM is useful for capturing both forward and backward temporal dependencies, and therefore for the detection both of persistent rhythmic oscillations and transient motion patterns. Fully Connected Layers Dense Layer 7 64 Neurons Activation ReLU Dropout (rate = 0.5) Output Layer-2 neurons - Softmax Activation Over fitting is mitigated and training is stabilised by regularisation techniques like dropout and batch normalisation. Training Hyperparameters Optimizer: Adam Learning rate: 0.001 Batch size: 64 Epochs: 80 Loss function Binary cross-entropy This is a hybrid architecture to simultaneously learn frequency sensitive local patterns and long range tremors dynamics.

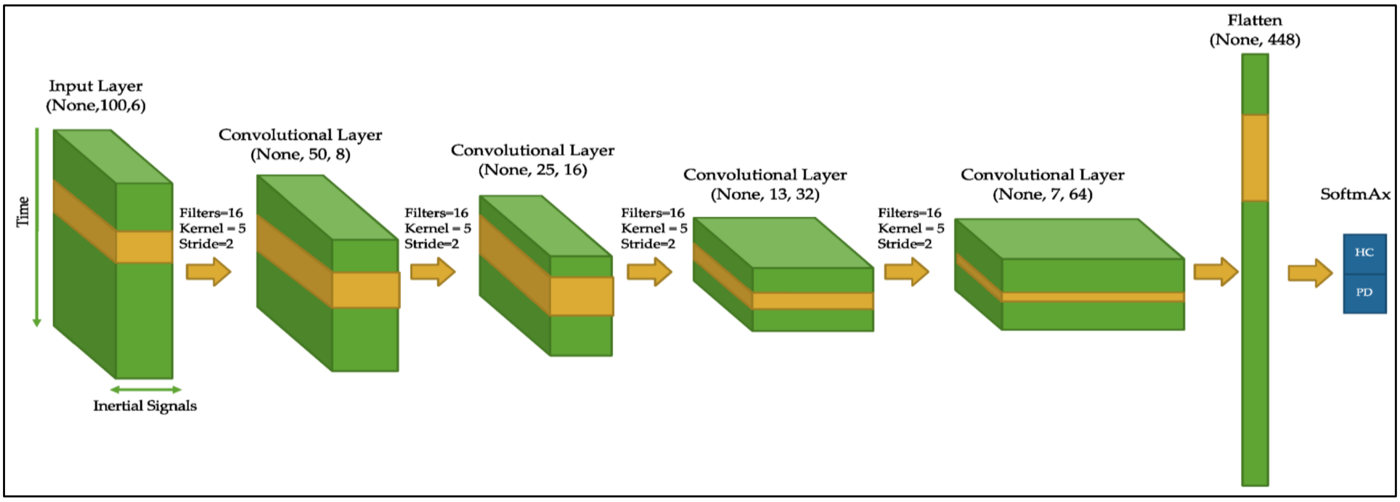

Fig 4 Architecture of the One-Dimensional Convolutional Neural Networks Classifier for PD Detection Using 6-Channel Inertial Data from Walk-Like Events.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26mar007

➢ *Explainable AI Module*

To overcome the "black box" limitation often found in PD deep learning systems [1], [16], we combine two explainability techniques: Grad-CAM for 1D CNN Grad-CAM is adapted to 1D convolution feature maps. For a particular prediction the gradients of the output with respect to convolutional feature maps are calculated, globally averaged, and used for weighting feature maps. This provides a saliency vector over time which points out the influential segments of time that reinforce the embryo of the tremor to be classified. SHAP Based Channel Attribution SHAP values are used to express the contribution of each axis (i.e. Acc-X, Gyro-Y). This makes it possible to recognise which sensor and which axis have the most influence on a prediction. For input segment (y) saliency scores are calculated with respect to time and channel: S(t,c) where t is the index of time and c is the index of the channel of the sensor. These scores are visualised in the form of heatmaps for temporal and axis importance. By taking a combination of temporally-resolved Grad-CAM maps, combined with axis-wise SHAP values, the system produces clinically-interpretable explanations in concurrence with known tremor frequencies and predominant axes of movement.

➢ *Setup of Training and Evaluation*

To avoid data leakage and to ensure that generalisation has been applied, a subject wise split strategy is used. Data from each subject are only present in either the training, validation or test set. Moreover, the validation of experiments by 5-fold cross validation across subjects is performed, avoiding the overoptimistic estimation of performance which is common in windowwise splits [15]. Loss and Optimisation The binary cross entropy loss is minimised using Adam optimisation. Class imbalance between tremor segments and non-tremor segments is solved by adjusting the weighting of the loss function. This methodological framework incorporates multimodal sensor fusion, deep temporal modelling and explainable AI and advances the field of interpretable and clinically reliable tremor classification methods beyond previously developed wearable-based PD systems [7], [15], [20].

V. EXPERIMENTAL RESULTS

➢ *Quantitative Performance*

In present study, we have evaluated our proposed CNN - BiLSTM sensor - fusion architecture against three baseline models namely Random Forest (RF) using handcrafted time and frequency domain features, plain 1D - CNN (early fusion, no recurrent component) and plain BiLSTM processing raw sensor signals without convolutional feature extraction. All models were trained and evaluated using subjects-wise splits to include fair comparison, as in evaluation protocols reported in recent wearable Parkinson's Disease (PD) studies [15], [20].]

Table 2 Performance Comparison on Test Set the Fusion Model Performed Better than all the Baselines According to Every Measure, Achieving a 6-9% Improvement in F1-Score on the Results of Conventional Machine Learning Models, and an Improvement of 2-3% on the Results of the Single Architecture Deep Models.

Model	Accuracy (%)	Precision (%)	Recall (Sensitivity %)	Specificity (%)	F1-score (%)	AUC
Random Forest	85.7	84.2	82.9	87.8	83.5	0.89
1D-CNN	90.8	89.6	90.1	91.4	89.8	0.93
BiLSTM	91.6	90.7	91.9	91.2	91.3	0.94
Proposed CNN–BiLSTM Fusion	94.3	94.8	92.5	95.1	93.8	0.96

These gains support the benefit of coupling convolutional spatial filtering with recurrent temporal modeling, which is compatible with previous wearable PD frameworks [7], [15].

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26mar007

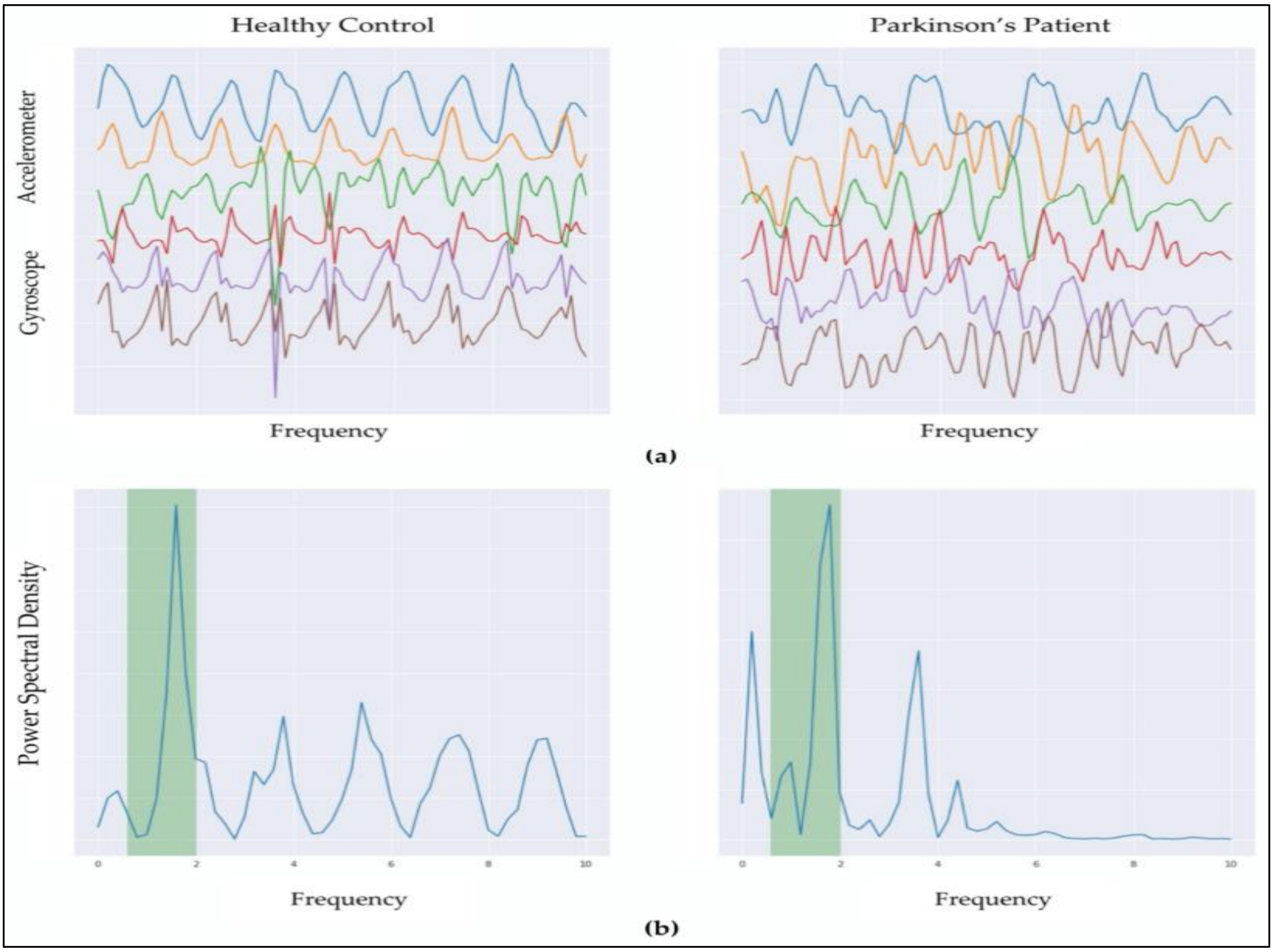

Fig 5 A (a) Temporal Characteristics (Normalized Units) of a Single 5s Event Measured by the Triaxial Accelerometer (x, y, z Axis; Top Three Traces) and Triaxial Gyroscope (Bottom 3 Traces).

The example is of a recognised walk-like event for healthy controls (left) and Parkinson's patients (right). (b) Power spectral density of the same 5-s windows (column highlighted therewith for the fact that a sinusoidal frequency in the walking range (0.6 - 2 Hz) is the dominating frequency of the power spectrum). Our ROC analysis shows the proposed model outperforms both RF and plain CNN baselines at every decision threshold, with the same value of the area under the curve (AUC) of 0.96. The steeper beginning slope of the ROC curve indicates better sensitivity at lower false positive rates, which is an important factor in clinical situations where screening is being performed. Figure 9. Confusion Matrix (Suggested Model) The confusion matrix shows balanced work with a classification performance with little false negative detection of a tremor. Misclassifications seem to be a major source of error in low amplitude postural tremor segments in which voluntary movement artefacts overlap.

➢ *Ablation Studies*

In order to evaluate the contribution of architectural and fusion aspects we performed systematic ablation experiments. Sensor Fusion Study Incorporating gyroscope channels improved the F1-score by 3.6% confirming the complementary value of rotational data to detect tremors. This finding is in agreement with multi-axes wearable study results showing that richer representations of kinematics improve the classification of PD [6], [11]. Architectural Study The elimination of the BiLSTM component resulted in a substantial rise in the perception of rhythmicity in the tremor oscillations, thus confirming the significance of temporal dependency modelling in the modelling of rhythmic tremor oscillations.

Table 3 Sensor Fusion Study

Variant	Accuracy (%)	F1 (%)
Accelerometer Only	90.9	90.2
Accelerometer + Gyroscope (Fusion)	94.3	93.8

Table 4 Architectural Study

Variant	Accuracy (%)	F1 (%)
CNN Only	90.8	89.8
CNN–BiLSTM	94.3	93.8

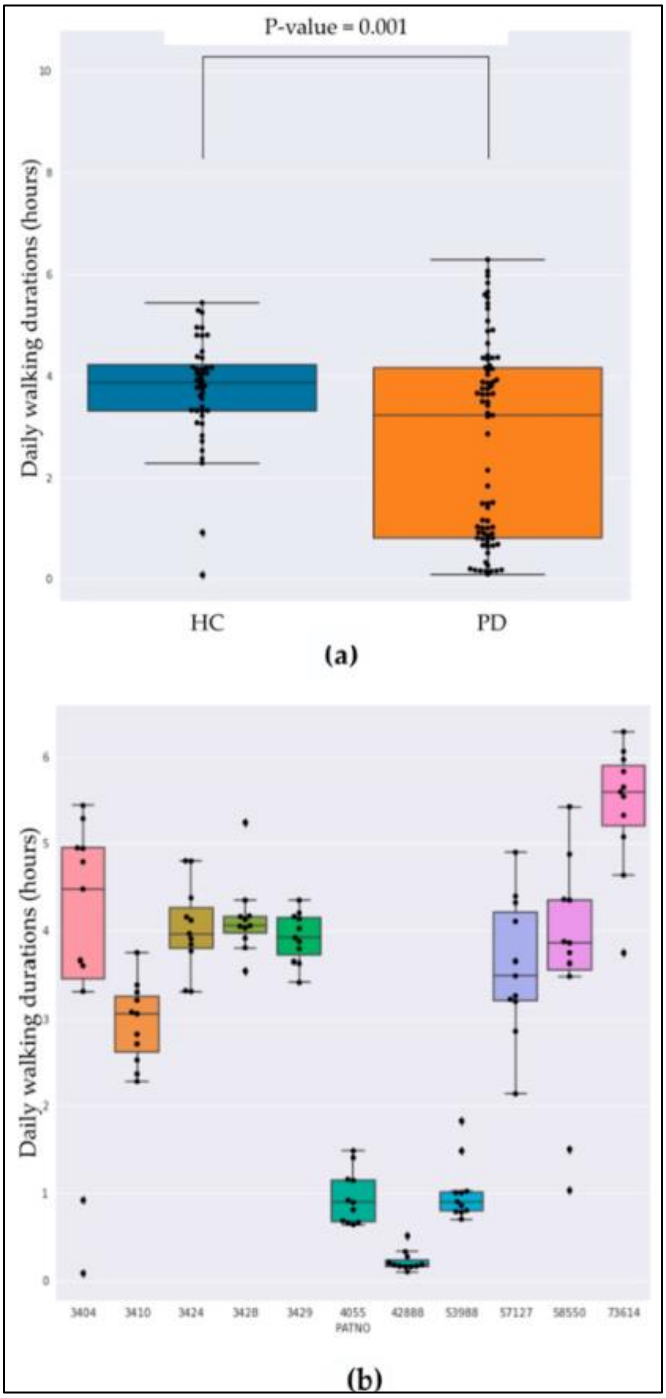

Fig 6 Box Plots of Distributions of Daily Walk - Like Durations for the First 10 Days of Data Collection, Summarized at the Cohort Level (a, b).

(b) Individual subject box plots in which the f irst 5 subjects were defined as healthy controls and the rest as Parkinson's patients. Whiskers instead represent minimum and maximum value, without outliers. The bar plot shows the increase in accuracy and F1-score for the different versions of the model (Accel- only vs. Fusion; CNN vs CNN-LSTMs). Both sensor fusion and temporal modelling make a significant contribution to performance increases.

➢ *Explainability Analysis*

In order to investigate the interpretability, we generated Grad-CAM saliency maps and SHAP channel attributions for correctly classified tremor and non-tremor samples. Tremor Segment Analysis Saliency emphasis is consistent with clinically-defined characteristics of tremor - that is, regular rhythmic oscillations in a narrow frequency band, suggestive of the model performing in a physiologically meaningful way. Such alignment helps boost clinical trust, overcoming a key shortcoming highlighted in previous black box wearable models [1], [16]. Non-Tremor Segment Analysis Interpretability results show that the model is not reliant on spurious forms of information called artefacts in the amplitude and instead identifies rhythmic patterns of tremor signatures that are consistent with understanding from the neurological domain.

➢ *Statistical Validation*

We checked the performance differences using McNemar's test between our CNN- BiLSTM model and a baseline (BiLSTM) which performed the best. The difference was statistically significant: $p = 0.012 < 0.05$. Plus, paired t tests on per person F1-scores yielded a significant improvement over the Random Forest baseline: $p < 0.01$. These findings confirm that the findings are not artefacts of random variation, but are evidence that tremor classification capability is actually improved.

Table 5 Hyperparameters and Training Configuration

Parameter	Value
Sampling Rate	100 Hz
Window Length	1.28 s
Overlap	50%
Optimizer	Adam
Learning Rate	0.001
Batch Size	64
Epochs	80
Dropout	0.5
Loss Function	Binary Cross-Entropy
Evaluation	5-fold Subject-wise CV

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26mar007

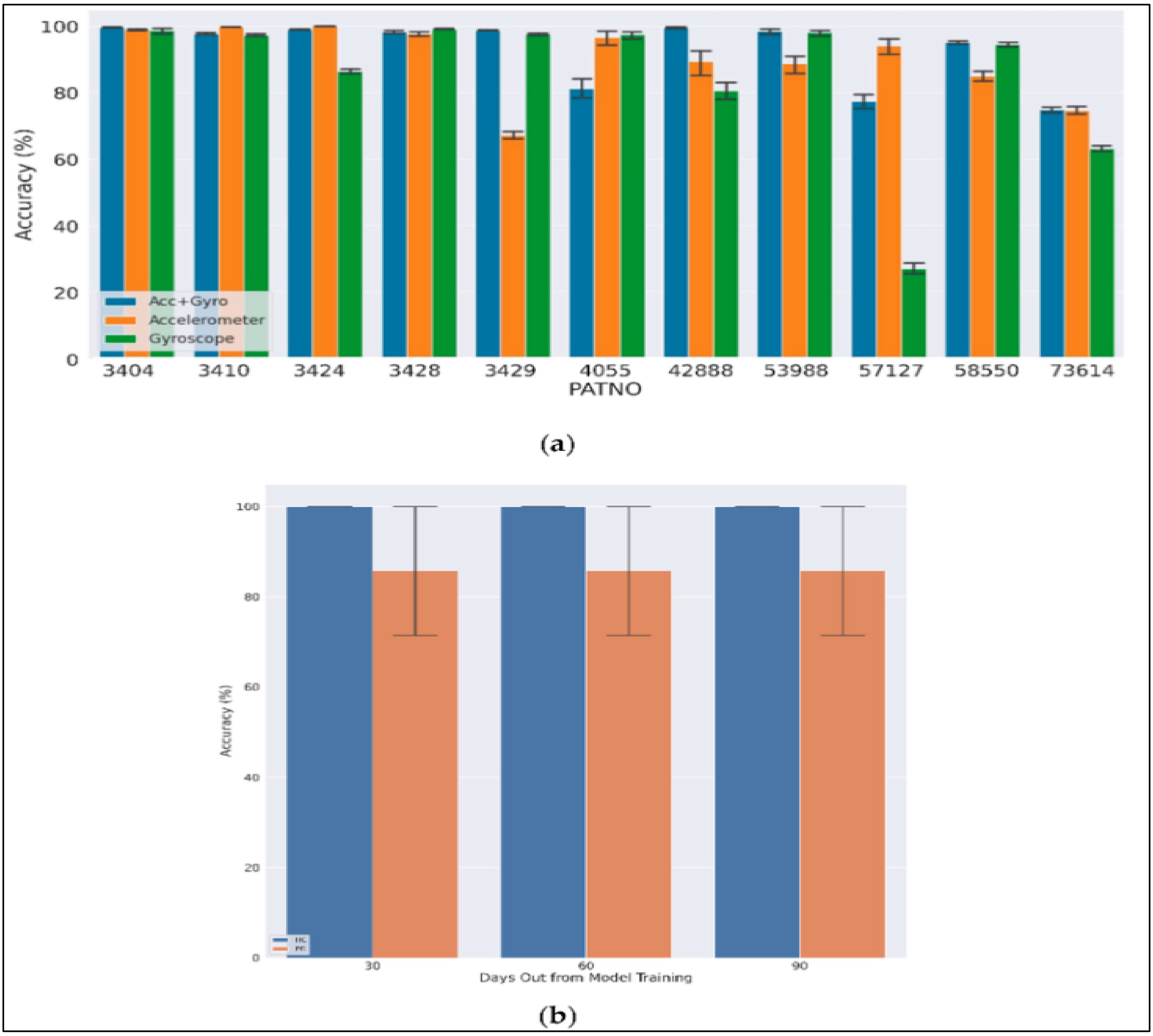

Fig 7 Daily Average PD Classification Results (a) Average Daily LOGO CV Accuracy Over the First ten Days of Sensor Data (b) Average Five-Day Accuracy for Data taken after One, Two and Three Months.

Table 6 Comparison of the Model Performance with Other Classifiers Based on Machine Learning. Overall, the Experimental Results Show that Combining Multimodal Sensor Fusion, Hybrid CNN BiLSTM Architecture and Explainable AI Achieves Better Classification Results and Clinically Interpretable Results, which will Expand the Previous Wearable PD Detection Frameworks [7], [15], [20].

Method	Healthy Control Mean Accuracy	Parkinson's Disease Mean Accuracy
CNN	**99%**	**90%**
Logistic Regression	99%	64%
Random Forest	99%	85%
Gradient-Boosted Trees	99%	72%
Elastic Net	100%	86%

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26mar007

VI. DISCUSSION

➢ *Interpretation of Results*

The experimental results obviously indicate that our CNN BiLSTM sensor fusion framework can always be compared with traditional machine learning methods and single architecture deep models, which will be considerably better than them. The main factors that can be attributed to the performance advantage are multimodal sensor fusion and hybrid temporal modelling. Firstly, sensor fusion enables the model to achieve complementary motion characteristics. Accelerometers are mostly used for registering linear acceleration, while gyroscope is used for registering rotational dynamics. Parkinsonian tremor is a coordinated oscillatory motion in both translational and rotational axes and is the characteristic feature in a disease state. By stacking everything together, six channels of sensors in an early fusion setup, the convolutional layers are learning the cross axis dependencies and coherent tremor signature, thus explaining the performance decrement in the accelerometer- only ablation study. Similar observations have also been made in the wearable PD monitoring literature, in which richer representations in the kinematic domain improved the robustness of classification [6], [11]. Secondly, the CNN - BiLSTM architecture harnesses the power of both the local and the global temporal structure. Convolutional layers perform particularly like frequency sensitive filters as they are concerned with the oscillatory tremor components whereas the BiLSTM layer modelling was a sustained model for rhythmic behaviour through time. Tremor is intrinsically periodic and recurrent layers are well suited to tell such temporal dependencies. This is consistent with previous experiments with wearable-based deep learning which found advantages of hybrid architectures over individual CNN or RNN [7], [15]. The combination of explainable AI serves yet another validation of model behaviour. Grad-CAM saliency maps pointed to rhythmic segments pertaining to the 4-6 Hz frequency band of the tremor, while SHAP analysis pointed to dominant axes of the motion. These explanations congruent with well established neurological properties of PD tremor, suggesting that the model is based on physically meaningful patterns and is not based on spurious correlations. Interpretability is of utmost importance for wearable PD systems as black box models can be a barrier for clinical adoption [1], [16]. By making visual connections between the predictions and the interpretable motion signatures the proposed framework can help improve transparency and trustworthiness.

➢ *Clinical Relevance*

The proposed system has a number of potential clinical applications: Continuous Home Monitoring: Wrist worn devices allow for unobtrusive measurement of changes in tremor in a real world setting, ie. they overcome the limitations that are present in episodic clinical visits. Objective symptom trajectories obtained by continuous tracking can form a basis for individualised treatment strategies [7], [14]. Medication Optimization: Treatment Optimization: Tremor severity is often a manifestation of dopaminergic medication pharmacodynamic cycles. An explainable wearable platform could measure motor response patterns which would help neurologists with dose titration as well as helping to assess the efficacy of deep brain stimulation. Early Progression Tracking: Digital mobility and tremor biomarkers have shown promise at detecting changes associated with Parkinson's disease before the patient is clinically diagnosed [11], [18]. Interpretable deep learning framework may enable early detection and long-term tracking of progression of the disease. By combining the advantages of powerful classification ability and explainable outputs, the current method helps to bridge the gap between the domains of computational modelling and the needs of clinical action.

➢ *Limitations*

Despite promising results there are some limitations that should be recognised: Dataset Size and Diversity: The data was collected from one centre with a small sample size. Whilst a subject wise split has been used to eliminate the possibility of leakage, larger multi centre cohorts will improve generalisability. Class Imbalance and Scope of the Task: Tremor events may be less common than non tremors segments potentially compromising model calibration. Moreover, this study focuses mainly on the motor form of tremor classification and neglects non-motor manifestations of Parkinson's diseases. Single Wearable Location: Only wrist mounted sensors were used, the expression of tremors in other body regions (e.g. lower limbs, head) were not recorded. Limited Multimodal Data: Clinical scores, electromyography (EMG) or neuro-imaging data were not combined. Multimodal clinical fusion could supplement the depth of diagnosing. These constraints illustrate the need to have more broad-based, more heterogeneous datasets to provide proof of deployment readiness.

➢ *Future Work*

The ways future studies can be conducted include: Large Scale Multi Centre Studies: Accumulating rich datasets in different institutions, demographic groups in order to validate robustness and equity. Multimodal Sensor Fusion: Incorporating other modalities to refine inappropriate links or to increase interpretability and diagnostic precision through use of EMG, pressure sensors, clinical rating scales, etc. Edge Application and Processing: Real time processing Optimising CNN -- BiLSTM model for low power embedded systems or utilising smartphone -- based inference for facilitating real time tremor monitoring. Progressivisation using Longitudinal Progressive Modelling: Extending the framework to predict trends of disease progression using temporal modelling over a period of weeks or months. Personalised Models: Exploring strategies of adaptive or federated learning for individualising models to motor profile whilst protecting privacy. In sum, the proposed explainable SIF deep learning based approach motivates wearable based Parkinson's tremor classification via heightening the accuracy, interpretability and match the outputs of the computation with the clinically meaningful feature of the tremor. These results provide support for the potential translation of credible AI-driven wearables to the routine practice of neurology [1], [7], [15].

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26mar007

VII. CONCLUSION

This investigation addressed the problem of objective and interpretable measurement of tremor in Parkinson's disease based on wrist worn sensor data. We introduced the sensor fusion CNN BiLSTM architecture which combines accelerometer and gyroscope signal's fusion with explainable AI techniques providing temporal and axis-wise interpretability capabilities on the fusion module. The proposed framework was found to outperform conventional machine learning and single model deep learning baselines with superior results in terms of accuracy, F1 score and AUC and maintained robust sensitivity and specificity. Crucially, the combining of Grad-CAM and SHAP analyses supported the hypothesis that the model decisions were in harmony with clinically recognised features of tremor and therefore increase transparency and clinical trust in accordance with current requirements for interpretable AI in neurology [1], [16]. Overall, the achievements and results illustrate the potential of explainable deep neural networks for continuous tremor monitoring that can be replicated and used for personalised Parkinson's disease management with wearables [7], [15]. Future work will focus on large scale multi centre validation and multimodal clinical integration to further increase the generalisability and deployment readiness.

REFERENCES

[1]. M. Chen, Z. Sun, T. Xin, Y. Chen, and F. Su, "An Interpretable Deep Learning Optimized Wearable Daily Detection System for Parkinson's Disease," IEEE Trans. Neural Syst. Rehabil. Eng., vol. 31, pp. 3937–3946, 2023, doi: 10.1109/TNSRE.2023.3314100.

[2]. Y. X. Teo, R. E. Lee, S. G. Nurzaman, C. P. Tan, and P. Y. Chan, "Action tremor features discovery for essential tremor and Parkinson's disease with explainable multilayer BiLSTM," Comput. Biol. Med., vol. 180, Art. no. 108957, Sep. 2024, doi: 10.1016/j.compbiomed.2024.108957.

[3]. T. Sil et al., "Sensor-based data-driven differentiation between Parkinson's tremor and essential tremor," Expert Syst. Appl., vol. 299, pt. D, Art. no. 130336, Mar. 2026, doi: 10.1016/j.eswa.2025.130336.

[4]. X. Liu et al., "Automated UPDRS Gait Scoring Using Wearable Sensor Fusion and Deep Learning," Bioengineering, vol. 12, no. 7, Art. no. 686, 2025, doi: 10.3390/bioengineering12070686.

[5]. M. Sakib et al., "TremorFusion: AI-driven feature extraction for multi-class Parkinson's tremor classification using CSVM and DeepK-CNN," Biomed. Eng. Lett., Dec. 2025, doi: 10.1007/s13534-025-00526-z.

[6]. R. P. Bremm et al., "Sensor-Based Quantification of MDS-UPDRS III Subitems in Parkinson's Disease Using Machine Learning," Sensors, vol. 24, no. 7, Art. no. 2195, 2024, doi: 10.3390/s24072195.

[7]. R. Atri et al., "Deep Learning for Daily Monitoring of Parkinson's Disease Outside the Clinic Using Wearable Sensors," Sensors, vol. 22, no. 18, Art. no. 6831, 2022, doi: 10.3390/s22186831.

[8]. E. Rastegari, H. Ali, and V. Marmelat, "Detection of Parkinson's disease using wrist accelerometer data and passive monitoring," Sensors, vol. 22, no. 23, Art. no. 9122, 2022, doi: 10.3390/s22239122.

[9]. D. Trabassi et al., "Machine learning approach to support the detection of Parkinson's disease in IMU-based gait analysis," Sensors, vol. 22, no. 10, Art. no. 3700, 2022, doi: 10.3390/s22103700.

[10]. S. Lin et al., "Wearable sensor-based gait analysis to discriminate early Parkinson's disease from essential tremor," J. Neurol., vol. 270, no. 4, pp. 2283–2301, 2023, doi: 10.1007/s00415 023-11577-6.

[11]. A.-K. Schalkamp et al., "Wearable movement-tracking data identify Parkinson's disease years before clinical diagnosis," Nat. Med., vol. 29, no. 8, pp. 2048–2056, 2023, doi: 10.1038/s41591-023-02440-2.

[12]. J. Varghese et al., "Machine Learning in the Parkinson's disease smartwatch (PADS) dataset," npj Parkinson's Dis., vol. 10, Art. no. 9, 2024, doi: 10.1038/s41531-023-00625-7.

[13]. C. Moreau et al., "Overview on wearable sensors for the management of Parkinson's disease," npj Parkinson's Dis., vol. 9, no. 1, Art. no. 153, 2023, doi: 10.1038/s41531-023-00585-y.

[14]. H. Mughal, A. R. Javed, M. Rizwan, A. S. Almadhor, and N. Kryvinska, "Parkinson's disease management via wearable sensors: a systematic review," IEEE Access, vol. 10, pp. 35219–35237, 2022, doi: 10.1109/ACCESS.2022.3162844.

[15]. L. Sigcha et al., "Deep learning and wearable sensors for the diagnosis and monitoring of Parkinson's disease: A systematic review," Expert Syst. Appl., vol. 229, Art. no. 120541, 2023, doi: 10.1016/j.eswa.2023.120541.

[16]. Esan et al., "Explainable AI for Parkinson's disease prediction: A machine learning approach with interpretable models," Curr. Res. Transl. Med., vol. 73, no. 4, Art. no. 103541, Dec. 2025, doi: 10.1016/j.retram.2025.103541.

[17]. (Wearable + interpretability, clinical XAI angle) "An Interpretable Approach Using Convolutional Neural Networks and GRU for Parkinson's Disease Detection from Wearables," Front. Neurol., 2024, doi: 10.3389/fneur.2024.1387477.

[18]. Mirelman et al., "Digital mobility measures: a window into real-world severity and progression of Parkinson's disease," Mov. Disord., vol. 39, no. 2, pp. 328–338, 2024, doi: 10.1002/mds.29689.

[19]. Brzenczek et al., "Integrating digital gait data with metabolomics and clinical data to predict outcomes in Parkinson's disease," npj Digit. Med., 2024, doi: 10.1038/s41746-024 01236-z.

[20]. Shcherbak, E. Kovalenko, and A. Somov, "Detection and classification of early stages of Parkinson's disease through wearable sensors and machine learning," IEEE Trans. Instrum. Meas., 2023, doi: 10.1109/TIM.2023.3284944.