

Tremor Monitoring and Severity Quantification Systems for Parkinson's Disease and Movement Disorders: A Literature Review

Mr. Ravikiran B A¹, Thanushree N², Ruchitha H³, Sharvya R⁴, Vinay S⁵

¹ Assistant Professor, Electronics and Communication Engineering, K. S. School of Engineering and Management

² Student, Electronics and Communication Engineering, K. S. School of Engineering and Management

³ Student, Electronics and Communication Engineering, K. S. School of Engineering and Management

⁴ Student, Electronics and Communication Engineering, K. S. School of Engineering and Management

⁵ Student, Electronics and Communication Engineering, K. S. School of Engineering and Management

Abstract - Tremor is one of the most noticeable and clinically important motor symptoms associated with Parkinson's disease (PD) other movement disorders, yet its assessment in routine practice still relies on visual clinical rating scales that are intermittent and can be influenced by differences between clinicians. Over the past decade, developments in affordable inertial sensors, surface electromyography, computer vision, and embedded technologies have opened new possibilities for objective and continuous monitoring of tremor. This paper presents a structured review of the literature sensor-based systems developed for tremor monitoring and severity assessment in PD and movement disorders, organized around four themes: sensing modalities (inertial measurement units, surface electromyography, smartphone and consumer devices, and vision-based systems), signal processing and feature extraction techniques, machine learning and deep learning approaches for tremor detection and severity estimation, and embedded or Internet of Things (IoT) architectures suited to continuous, real-world deployment. Fifty-seven studies published in peer-reviewed journals and conference proceedings are synthesized and compared on sensor configuration, algorithmic approach, sample size, and reported performance. The review identifies recurring methodological gaps, including small and imbalanced datasets, inconsistent frequency-band selection, limited external validation, and a scarcity of systems that combine embedded, low-power operation with clinically interpretable output. The paper concludes by outlining research directions that could help move tremor-quantification methods from controlled laboratory environments toward reliable, continuously operating clinical tools.

Key Words: Parkinson's disease, tremor quantification, wearable sensors, inertial measurement unit, surface electromyography machine learning, deep learning, Internet of Things, movement disorders.

1. INTRODUCTION

Parkinson's disease is a progressive neurodegenerative disorder characterized by the gradual loss of dopaminergic neurons in the substantia nigra, leading to a characteristic combination of motor and non-motor symptoms [1]. Among the cardinal motor features, tremor is often the symptom that first prompts a patient to seek medical evaluation, and it remains one of the more disabling aspects of the disease for many patients throughout its course [2]. Tremor does not occur alone; it interacts with bradykinesia, rigidity, and postural instability to reduce a patient's ability for daily tasks, and the resulting loss of independence has a measurable effect on quality of life even in the early and moderate stages of the disease [3]. Clinical practice continues to assess tremor primarily using the Movement

Disorder Society-sponsored Unified Parkinson's Disease Rating Scale (MDS-UPDRS) and the Hoehn and Yahr staging system [4], [5]. Both instruments rely on a trained examiner observing the patient during a brief clinic visit and rating symptom severity using an ordinal scale. However, this approach has several recognized limitations. A single clinic visit provides only a limited view of a patient's motor state and cannot capture the fluctuations that occur throughout a day, particularly when levodopa related motor complications develop during later disease stages. Ratings may also differ between examiners, and even for the same examiner across sessions, making comparisons of results across sites and over time more difficult. These limitations have encouraged a growing body of engineering research focused on measuring tremor objectively, continuously, and outside the clinic, using sensors ranging from basic accelerometers to advanced multi-camera vision systems [6], [8]. The rationale for sensor-based tremor monitoring is based on a simple observation: tremor appears as a rhythmic, nearly periodic movement occurring within a relatively identifiable frequency range, making it suitable for objective detection and measurement using signal-processing and pattern recognition techniques [9], [10]. After tremor-related signals are obtained from acceleration, angular velocity, muscle activity, or video recordings, several processing stages, including filtering, feature extraction, classification, and regression, can transform the collected data into an estimate of tremor severity that can potentially operate continuously without direct clinical supervision [11]. This approach is especially valuable in Parkinson's disease because motor symptoms can vary according to medication schedules, daily routines, and progression of the condition. Their work demonstrated the value of continuous monitoring beyond occasional measurements obtained during conventional clinical visits [12]. This review synthesizes 57 studies drawn from peer reviewed journals and conference proceedings that address tremor monitoring and severity quantification in Parkinson's disease and related movement disorders such as essential tremor.

2. BACKGROUND

A. Parkinson's Disease and Tremor

PD is diagnosed clinically according to criteria requiring bradykinesia together with rest tremor, rigidity, or both [1]. In PD, tremor is generally a rest tremor, referring to an involuntary rhythmic movement of a limb that occurs when the limb is completely supported against gravity and is usually reduced, either partly or entirely, during voluntary movement [10], [13], [14]. A re-emergent postural tremor, which develops

following a brief delay when a limb is maintained in a particular posture, is also regarded as a characteristic feature of the disease [10]. PD rest tremor commonly falls within the 3–6 Hz range, although its precise frequency differs among patients and may change throughout disease progression. Several other tremor disorders also overlap with this frequency range, making differential diagnosis more difficult [9], [13].

B. Tremor Classification

Table I presents the frequency ranges and common clinical characteristics of various movement disorders associated with tremor, based on consensus statements and review literature in clinical research and practice [10], [11], [13]. Essential tremor, which represents the most common alternative diagnosis to Parkinsonian tremor, is primarily an action tremor rather than a resting tremor, and it generally occurs at a higher frequency than PD tremor. Since distinguishing between these disorders can be challenging, especially during early disease stages, several studies discussed in Section 3 focus on differentiating PD tremor from essential tremor using sensor-based measurements rather than relying only on clinical assessment. Sensor-based methods have also been investigated for directly measuring essential-tremor severity, rather than using them solely to distinguish it from PD. For instance, Ali et al. combined wearable accelerometer and gyroscope measurements to estimate essential-tremor severity in comparison with clinical ratings over time [15].

TABLE I

CLASSIFICATION OF TREMOR TYPES BY FREQUENCY AND PRESENTATION [10], [11], [13]

Tremor Type	Body Distribution	Activation Condition	Freq. (Hz)
Parkinsonian tremor	Asymmetric upper limbs, legs, jaw	Rest, emergent on posture	3–6
Essential tremor	Bilateral upper limbs	Action (postural, kinetic)	4–12
Enhanced physiological tremor	Bilateral limbs	Postural, kinetic	8–12
Dystonic tremor	Asymmetric limbs / head	Action, occasionally rest	3–7
Cerebellar tremor	Limbs, trunk	Intention	<5
Orthostatic tremor	Bilateral lower limbs	Isometric, standing	13–18
Drug-induced tremor	Bilateral limbs	Action, occasionally rest	8–12

C. Clinical Rating Scales

The MDS-UPDRS continues to serve as the main reference standard for validating most sensor-based tremor assessment systems [5]. Part 3 of this scale evaluates rest tremor amplitude, postural and kinetic tremor, as well as tremor constancy, with each feature rated on a five-point ordinal scale ranging from zero (absent) to four (severe). In comparison, the Hoehn and Yahr scale offers a broader stagebased measure of overall disease progression [4]. Almost all studies examined in this review use one of these clinical scales, particularly MDS-UPDRS Part 3, as the ground-truth reference for training or

validating sensor-based measurements. Consequently, automated system performance is largely limited by the consistency and reliability of the clinical ratings used for training.

3. SENSING MODALITIES FOR TREMOR MONITORING

Three systematic reviews published between 2017 and 2022 have each surveyed the broader landscape of wearable sensing for PD, and their combined coverage gives a useful starting point for organizing the literature [6], [8]. Rovini et al. examined studies published between 2006 and 2016 and grouped applications into five categories: early diagnosis, tremor, body motion analysis, motor fluctuations, and long term monitoring [6]. Sica et al. restricted their review to studies with at least one unsupervised, free-living data capture period; that criterion alone tracks how far the field has moved from laboratory protocols toward naturalistic monitoring [7]. Ancona et al. reached a similar conclusion, noting that although wearable systems show consistent promise for home-based assessment, most studies to date still involve small, non-representative cohorts [8]. The subsections below examine the specific sensing modalities used across the reviewed literature.

A. Inertial Measurement Units

Tri axial accelerometers and gyroscopes, used alone or together as an inertial measurement unit (IMU), are widely used in tremor monitoring studies. Early work by Kim et al. used a wrist worn accelerometer and gyroscope with a convolutional neural network to measure PD tremor severity using MDS UPDRS labels, achieving 85% classification accuracy across four severity grades [16]. Thorp et al. monitored daily activities directly instead of using fixed tasks, reporting sensitivity and specificity above 99%. However, the study did not provide enough details about the model or extracted features for direct comparison with later studies [17]. Li et al. used a similar approach by placing inertial sensors on everyday objects such as cups and door handles, allowing motor symptoms to be recorded during normal daily activities instead of a scripted assessment [18]. Many studies have used IMU data to distinguish PD tremor from essential tremor, which is clinically important because the two conditions can appear similar. Loaiza Duque et al.

applied machine learning to accelerometer data recorded during rest and postural tasks. They found that angular velocity features combined with a K nearest neighbor classifier improved differential diagnosis accuracy compared with acceleration features alone [19], [20]. De Araújo et al. reported training and testing accuracies of 96.4% and 98.4%, respectively, for a hand resting tremor classifier based on combined time and frequency features [21]. Peres et al. studied a similar problem using inertial measurements of hand resting activity to distinguish PD patients from healthy controls [22]. Papadopoulos et al. used a different method by treating each patient's recordings as a labeled group under a multiple instance learning framework. They achieved 89.3% accuracy with an F1 score of 85.1% for detecting tremor from wrist worn IMU data collected during free living conditions [23], [24]. Distinguishing tremor from voluntary effort is another related machine learning problem. Aljihmani et al. trained classifiers to separate inertial patterns caused by tremor from those

produced during normal voluntary movement. This difference is important for severity grading systems, although it is rarely studied separately [25]. Several IMU based studies focus on severity grading instead of simple detection. Ranjan et al. converted one dimensional accelerometer signals into two dimensional image representations and used a convolutional neural network. They reported accuracies ranging from 100% for the mildest tremor grade to 66% for the most severe grade. This pattern is linked to the class imbalance found in many clinical tremor datasets because severe tremor is less common in typical patient groups [26]. AlMahadin et al. compared several classical classifiers for four level severity grading and reported 98% accuracy using a support vector machine trained with time, frequency, and statistical features. This was the highest accuracy among the classical machine learning studies reviewed here [27]. A related study by the same group on task based severity measurement [28]. Li et al. combined automatic label correction with a support vector machine using a radial basis function kernel and achieved 97.96% accuracy for grade classification, which was close to the highest result found for classical models in this review [29]. Among deep learning methods, Tong et al. reported the highest single accuracy identified in this review, 97.32%, using a convolutional neural network with inertial data for binary hand tremor detection [30].

Not all studies use a control group of participants without tremor, and this affects how their results should be understood. Wu et al. and Ibrahim et al. worked only with confirmed PD patients and focused on tremor subtyping or tremor onset detection within patients instead of distinguishing PD patients from healthy individuals [31], [32]. In contrast, San Segundo et al. tested tremor detection under free living conditions using data collected outside the laboratory. This approach was later extended by Zhang et al. and Sun et al. in similar real world settings [33] to [36]. Sigcha et al. combined a multitask convolutional neural network with smartwatch accelerometer data and reported 86.1% sensitivity and specificity for resting tremor detection, with an area under the receiver operating characteristic curve of 0.936 [37]. Zhao et al. focused on an earlier stage of the disease and used IMU based machine learning to detect PD before tremor severity reached the levels commonly studied in severity grading research [38]. More recently, Rodriguez et al. developed a supervised learning system trained and tested using a 67 hour, four limb dataset collected from patients performing normal daily activities in a hospital setting. This directly addressed the problem of monitoring patients during unrestricted activities, which most earlier studies could only partly represent [39]. Dattola et al. used an unsupervised method with k means clustering on motion intensity features obtained from accelerometer data. They reported 76% accuracy for tremor versus non tremor classification, but the accuracy dropped to 57.1% for multiclass severity grading. This difference shows that estimating severity is much more difficult than simply detecting tremor when labeled training data are not available [40].

B. Surface Electromyography

Surface electromyography (sEMG) provides a useful signal along with inertial sensing because it measures muscle activity rather than the movement produced by the muscles. PD tremor is linked to alternating bursts of activity in opposing muscle pairs, while essential tremor usually shows more simultaneous

muscle contraction. This physiological difference cannot be directly observed using inertial sensors. Farhani et al. used a bidirectional long short term memory network trained only on sEMG data from fifteen PD patients. They reported 84% accuracy for task classification and 88% accuracy for tremor type classification, both clearly above chance level [41]. Studies that combine sEMG with accelerometer data generally show better agreement with clinical labels than either signal alone, which is consistent with the wider use of multiple sensors discussed in Section 3 E.

C. Smartphone and Consumer Grade Devices

Another area of research examines whether tremor can be measured using devices that patients already have, without requiring a dedicated wearable device. Barrantes et al. placed a smartphone on the back of the hand and used frequency power spectral features to distinguish PD from essential tremor. They presented the results as receiver operating characteristic curves instead of one accuracy value [42]. Mousavi et al. used a simpler method with the smartphone's built in accelerometer and a support vector machine classifier to detect hand tremor without additional hardware [43]. Kuosmanen et al. developed the Sentient Tracking of Parkinson's (STOP) application to collect accelerometer data during natural, real world activities. They compared a calculated tremor intensity measure with self reported UPDRS II tremor scores and found significant differences in signal characteristics before and after medication intake [44]. Habets et al. used a different approach by developing a smartphone based ecological momentary assessment diary. The application collected motor and non motor symptom ratings several times each day and was designed to track changes in symptoms rather than directly measure tremor from raw sensor data [45]. Varghese et al. studied a wider smart device system for movement disorder assessment in a prospective clinical study and reported its diagnostic accuracy compared with standard clinical examination [46]. These studies show that consumer smartphones can be useful for long term symptom tracking. However, the position of a phone can change during use and its sampling is less consistent than a sensor fixed to a limb, which are practical limitations compared with dedicated wearable devices

D. Vision Based Systems

Camera based methods do not require any physical sensor to be attached to the patient. Instead, they use ordinary video to estimate body posture and movement. Liu et al. proposed a vision system that estimates MDS UPDRS tremor severity scores directly from video. The system used a global spectral module to extract frequency domain features and showed better severity prediction than earlier video based methods [47]. Vision based tremor measurement is useful because it does not require physical contact with the patient and can potentially work with a standard consumer camera. However, it requires stable lighting, a clear view of the limb, and pose estimation models that can require high computing power, making them less suitable for low power embedded systems.

E. Multimodal Sensor Fusion

Several studies discussed above already use more than one type of sensor. A common finding is that combining complementary signals, such as accelerometer data with gyroscope data or

inertial data with sEMG, often gives better results than using one signal alone. This is especially useful for the more difficult multiclass severity grading task compared with simple binary tremor detection. Gutowski et al. combined a forearm worn armband with smartphone sensors and found that models trained using the combined data consistently performed better than models using a single device. Tremor was the easiest of four PD symptoms to predict, while dyskinesia was the most difficult because there were very few high severity examples in the training data [48]. This finding is also seen in other studies. Tremor is relatively easy to detect from movement signals because it produces a strong and narrow frequency oscillation. In comparison, bradykinesia and dyskinesia involve more subtle changes in timing and movement strength and therefore benefit more from multiple types of sensor data.

4. SIGNAL PROCESSING AND FEATURE EXTRACTION

A. Preprocessing and Frequency Band Selection

Since tremor is mainly an oscillating movement, most systems first isolate the frequency range related to pathological tremor. High pass filtering is commonly used to remove gravity and very low frequency drift, followed by low pass filtering to reduce high frequency sensor noise [48]. The frequency range used for analysis differs across studies. Some studies use a narrow 3 to 6 Hz range that closely matches rest tremor, while others use a wider 1 to 20 Hz range that can also include postural and kinetic tremor [9], [10]. This difference is important because a frequency range that is too narrow may remove real tremor signals near the limits of the normal range. On the other hand, a range that is too wide may include noise and unwanted voluntary movement, which can reduce classifier performance. Therefore, accuracy values from studies using different frequency ranges should be compared carefully.

B. Feature Domains

Common time domain features extracted from accelerometer and gyroscope signals include the mean, standard deviation, root mean square amplitude, peak value, and zero crossing rate, along with the mean absolute value and its first difference [48]. Frequency domain features are usually derived from the power spectral density and include the dominant frequency, spectral centroid, power within the tremor frequency range, and harmonic indices that represent the energy of higher harmonics compared with the main tremor frequency. Statistical features such as skewness, kurtosis, and Shannon entropy are also often used in both domains to describe the irregularity that separates pathological tremor from smoother voluntary movement. Gutowski et al. found that frequency domain features from gyroscope data in the 3 to 9 Hz range were the most important for tremor severity prediction. In contrast, bradykinesia severity was better predicted using time domain features such as skewness and median amplitude. This shows that the best feature set depends on the symptom being studied rather than being the same for every condition [48].

C. Feature Selection and Dimensionality Reduction

A single sensor recording can generate a large number of possible features when multiple axes, frequency ranges, and feature types are combined. Therefore, most systems use dimensionality reduction before classification. Common methods include correlation-based pruning, principal

component analysis, and iterative wrapper methods that add features one at a time according to their cross validated performance. Wrapper based selection is generally reported to perform better than simpler filter methods when the features are strongly related to each other. This is common when the same physical quantity is measured across several axes and frequency ranges [48]. Legaria Santiago et al. studied this problem specifically for hand tremor and compared several computational models using different feature groups to determine which combinations of extracted features were most useful for evaluating tremor presence and severity [49].

5. MACHINE LEARNING AND DEEP LEARNING APPROACHES

A. Classical Machine Learning

Support vector machines (SVMs) are among the most frequently used classifier in the reviewed literature, generally achieving accuracies at or above 76% and reaching as high as 98% when trained on a well-chosen combination of temporal, frequency, and statistical features [27]. Random forest and Knearest-neighbor classifiers appear almost as frequently, often as a baseline against which a proposed deep-learning model is compared. AlMahadin et al. found that Bayesian hyperparameter optimization, which uses the outcome of previous evaluations to select promising parameter combinations more efficiently than an exhaustive grid search, produced the best-performing SVM in their comparison [27], while several other groups relied on conventional grid search combined with precision- or F1-score-based model selection [26], [29].

B. Deep Learning

Convolutional neural networks (CNNs) are the most widely used deep architecture in this literature, frequently applied after transforming a one-dimensional inertial signal into a two-dimensional time-frequency image, an approach that allows standard image-classification network designs to be reused for tremor analysis [26], [30], [35]. Tong et al.'s CNN-based hand-tremor detector achieved the highest reported single accuracy among all deep-learning studies surveyed, 97.32% [30]. Recurrent architectures, particularly long short-term memory (LSTM) and gated recurrent unit (GRU) networks, appear less often but are a natural fit for the sequential nature of tremor signals; Hathaliya et al. combined a GRU-LSTM architecture with blockchain-based data provenance for essential and Parkinsonian tremor classification, reporting 80.4% training and 74.1% validation accuracy [50], while Farhani et al.'s bidirectional LSTM applied to sEMG achieved comparable performance using a substantially smaller and more specialized dataset [41]. Attention-based and multiple-instance learning architectures, exemplified by Papadopoulos et al.'s work, represent a further refinement aimed specifically at handling weakly labeled, patient-level data collected outside the laboratory [23], [24].

C. Unsupervised and Weakly Supervised Learning

Since expert-labeled tremor data are expensive to collect, several studies have explored whether unsupervised methods can reduce the field's dependence on clinician-assigned labels. Dattola et al.'s k-means clustering approach achieved reasonable accuracy for binary tremor detection but showed a

marked drop in performance for multiclass severity grading. Presence detection and severity grading are meaningfully different problems, and a modeling strategy that works well for one does not necessarily carry over to the other [40]. Multiple-instance learning, used by Papadopoulos et al., offers a related but distinct solution by allowing patient-level rather than sample-level labels to guide training, which reduces the labeling burden without discarding supervision altogether [23], [24].

D. Model Interpretability

Clinical adoption of any automated tremor-quantification system depends in part on the ability of a clinician to understand why the model produced a given output, a requirement that has motivated growing interest in explainable machine learning for this application. Permutation-based feature-importance analysis, as used by Gutowski et al., and SHapley Additive exPlanations (SHAP), used in more recent embedded systems discussed in Section VI, both aim to expose which input features drove a given prediction. That work moves the field away from purely black-box classification and toward outputs a clinician could, in principle, check against known tremor physiology [48].

6. EMBEDDED AND IOT-BASED TREMOR MONITORING SYSTEMS

A subset of the reviewed literature moves beyond offline algorithm development toward complete, deployable hardware systems built around low-cost microcontrollers, a direction of particular relevance to resource-constrained development environments. Saez et al. developed the ELENA system, which pairs an MPU6050 six-axis inertial sensor with an ESP32 microcontroller to capture wrist-worn tremor data, transmits the data over Wi-Fi to a cloud platform for storage and MATLAB-based frequency analysis, and includes a microSD card for local buffering during connectivity loss [51]. The system was validated against an Apple Watch as a reference device and achieved frequency deviations under 0.5% and an average percentage error of -0.37%. For this specific task, a low-cost sensor and microcontroller combination can match the frequency-measurement accuracy of a commercial smartwatch [51]. A related IoT-based system, built on the same MPU6050 and ESP32 combination, addressed the class imbalance inherent to continuous tremor monitoring, where tremor events are comparatively rare relative to non-tremor periods, using the synthetic minority oversampling technique before training a gradient-boosted decision-tree (LightGBM) classifier, and paired the resulting model with SHAP-based explanations delivered through a web interface, explicitly targeting the interpretability gap discussed in Section 5.

Two design decisions distinguish this embedded and IoT-oriented branch of the literature from the primarily offline studies discussed in Sections 3 through 5. First, these systems must budget computation and communication carefully, since a battery-powered wrist-worn device cannot support the same model complexity as a workstation-trained deep network without unacceptable power draw; this pushes the field toward lightweight classical models, such as gradient-boosted trees or shallow SVMs, rather than large CNNs, at least for on-device inference. Second, because these systems are explicitly built for continuous, unsupervised operation rather than a single scripted clinic visit, they must handle intermittent connectivity, local

data buffering, and calibration drift over long deployment periods, concerns that are largely absent from studies conducted entirely within a single laboratory session. Given that low-power inertial sensing combined with an ESP32-class microcontroller now demonstrably approaches commercial-smartwatch accuracy [51], embedded and IoT architectures represent one of the more direct paths from the algorithmic advances surveyed in Sections 3 through 5 toward a system a patient could plausibly wear throughout an ordinary day. Several earlier embedded systems anticipated this direction. Channa et al. developed the A-WEAR bracelet, a wrist-worn device built specifically to detect hand tremor and bradykinesia in PD patients using onboard inertial sensing [52], and later extended this work to combine deep learning with cloud infrastructure for continuous severity estimation rather than on-device classification alone [53]. Ehsan et al. took the cloud-offloading approach further, building an automated cloud-based screening tool intended for low-resource clinical settings where dedicated neurological expertise may not be locally available [54]. Babayan et al. addressed the body-sensor-network side of the same problem, developing a signal-processing pipeline for everyday-life tremor monitoring using a network of on-body sensor nodes rather than a single wrist-worn device [55]. Together with the ELENA system, these studies illustrate a consistent trajectory in the embedded literature: moving tremor quantification from single-purpose laboratory hardware toward networked, cloud-connected, or multi-node systems intended for sustained real-world use.

7. RESEARCH GAPS AND CHALLENGES

A. Small and Imbalanced Datasets

Most of the studies reviewed rely on small patient cohorts of fewer than fifty participants, and severe-tremor examples are consistently underrepresented relative to mild or absent tremor, since severe tremor is comparatively uncommon in a typical clinical population [26], [40], [48]. This imbalance directly explains the accuracy drop-off observed at the highest severity grades in several studies [26], and it limits the extent to which any single reported accuracy figure can be generalized to an independent patient population. Although synthetic oversampling and class-weighted loss functions, have been used in some studies of the embedded systems discussed in Section VI, partially address the problem but do not substitute for larger, more balanced data-collection.

B. Inconsistent Frequency-Band and Protocol Selection

As discussed in Section 4, reviewed studies differ in the frequency band retained for tremor analysis, in the specific tasks patients are asked to perform, and in whether recordings are collected under scripted laboratory conditions or freelifing activity. These inconsistencies make direct comparison of reported accuracies across studies unreliable and complicate the identification of which methodological choices genuinely drive performance differences, rather than merely reflecting differences in task difficulty or dataset composition. Shawen et al. examined this problem directly, showing that data measurement characteristics, such as sensor placement, sampling rate, and recording duration, materially affect the accuracy with which wearable sensors detect PD symptoms, which reinforces the case for standardized reporting practices [57].

C. Limited External and Cross-Cohort Validation

Almost none of the studies surveyed validate a trained model on an independent dataset collected at a different site or with different equipment; nearly all report only within study cross-validation, whether stratified, leave-one-patient out, or a simple train-test split. Rodriguez et al.'s 67-hour free living dataset represents an unusually large and ecologically valid single-site collection, but even this study did not evaluate the trained model against data from another clinical center [39]. Without external validation, it remains unclear how well any given model would perform on a new device, a different patient population, or a different clinical protocol, which is precisely the setting in which a deployed clinical tool would need to operate.

D. Patient-Centric Data Partitioning

Several studies partition training and test data at the level of individual recording segments rather than individual patients, which risks placing segments from the same patient in both the training and test sets and can inflate reported accuracy relative to a genuinely patient-independent evaluation. Leave-one-patient-out cross-validation, used by some of the more recent studies surveyed here, addresses this concern directly, and its wider adoption would improve the comparability and credibility of reported results across the field.

E. The Gap Between Detection and Clinically Useful Severity Grading

Binary tremor-versus-no-tremor detection is now a comparatively mature problem, with several studies reporting accuracies above 90%. Multiclass severity grading that reproduces the full ordinal MDS-UPDRS scale remains substantially harder, both because of class imbalance at the severe end of the scale and because the boundary between adjacent severity grades is itself somewhat subjective in the underlying clinical rating. Closing this gap matters clinically, since a system that can only distinguish tremor from its absence offers limited value for tracking disease progression or titrating medication, both of which require finer-grained severity information.

F. Power, Interpretability, and Deployment Constraints for Embedded Systems

The embedded and IoT-based systems discussed in Section VI demonstrate that low-cost inertial sensing can approach commercial-device accuracy for frequency measurement, but relatively few studies in this review report sustained battery life, on-device inference latency, or long-term calibration stability under continuous, multi-day operation, all of which matter for a system intended to be worn throughout ordinary daily life rather than during a single laboratory session. Interpretability tools such as SHAP are beginning to appear in this branch of the literature but remain the exception rather than the norm.

8. FUTURE DIRECTIONS

Several directions follow naturally from the gaps identified above. Larger, multi-site data-sharing efforts, potentially built on the handful of public datasets already used in the field, would directly address the small-sample-size and external-validation problems that recur throughout this review. Standardization of frequency bands, task protocols, and

reporting practices, such as consistently specifying sample size, class distribution, and cross-validation strategy, would make results across studies more directly comparable. On the modeling side, continued development of lightweight architectures suitable for on-device inference, paired with interpretable output, would help close the gap between the algorithmic advances described in Sections 3 through 5 and the practical, power-constrained requirements of the embedded systems described in Section 6. Longitudinal studies that track individual patients over months rather than hours would allow future work to move beyond severity classification at a single moment toward genuine tracking of disease and medication-response trajectories, which is arguably the more clinically valuable long-term goal of this research area.

9. CONCLUSION

Tremor monitoring for Parkinson's disease and related movement disorders has progressed from early accelerometer-based classification studies to a diverse ecosystem of inertial, electromyographic, smartphone, visionbased, and embedded IoT systems, several of which now report classification accuracies above 90% for tremor detection and above 95% for severity grading under laboratory conditions. The literature reviewed here shows consistent evidence that combining complementary sensing modalities, and combining time-domain, frequency-domain, and statistical features, improves performance relative to any single modality or feature domain in isolation. At the same time, small and imbalanced datasets, inconsistent methodology, and a near-total absence of external validation remain the field's most persistent limitations, and these limitations, rather than any deficiency in the underlying sensing or modeling techniques, are what currently separate laboratory-validated tremor-quantification algorithms from dependable clinical tools. Embedded systems built around low-cost microcontrollers and inertial sensors have already demonstrated frequency-measurement accuracy comparable to commercial smartwatches. The hardware barrier to continuous, real-world tremor monitoring therefore appears lower than the methodological one. Closing that methodological gap, through larger and better-validated datasets, standardized protocols, and interpretable, power aware models, is the clearest path from the body of work reviewed in this paper to tools that clinicians and patients can actually use.

REFERENCE

- [1] R. B. Postuma, D. Berg, M. Stern, W. Poewe, C. W. Olanow, W. Oertel, J. Obeso, K. Marek, I. Litvan, A. E. Lang et al., "MDS clinical diagnostic criteria for Parkinson's disease," *Movement Disorders*, vol. 30, no. 12, pp. 1591–1601, 2015.
- [2] P. J. Garcia Ruiz, M. J. Catalan, and J. M. Fernández Carril, "Initial motor symptoms of Parkinson disease," *The Neurologist*, vol. 17, pp. S18–S20, 2011.
- [3] D. Muslimović, B. Post, J. D. Speelman, B. Schmand, R. J. de Haan et al., "Determinants of disability and quality of life in mild to moderate Parkinson disease," *Neurology*, vol. 70, no. 23, pp. 2241–2247, 2008.
- [4] C. G. Goetz, W. Poewe, O. Rascol, C. Sampaio, G. T. Stebbins, C. Counsell, N. Giladi, R. G. Holloway, C. G. Moore, G. K. Wenning et al., "Movement Disorder Society task force report on the Hoehn and Yahr staging scale: status and

recommendations," *Movement Disorders*, vol. 19, no. 9, pp. 1020–1028, 2004.

[5] C. G. Goetz, B. C. Tilley, S. R. Shaftman, G. T. Stebbins, S. Fahn, P. Martinez-Martin, W. Poewe, C. Sampaio, M. B. Stern, R. Dodel et al., "Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results," *Movement Disorders*, vol. 23, no. 15, pp. 2129–2170, 2008.

[6] E. Rovini, C. Maremmani, and F. Cavallo, "How wearable sensors can support Parkinson's disease diagnosis and treatment: a systematic review," *Frontiers in Neuroscience*, vol. 11, p. 555, 2017.

[7] M. Sica, S. Tedesco, C. Crowe, L. Kenny, K. Moore, S. Timmons, J. Barton, B. O'Flynn, and D.-S. Komaris, "Continuous home monitoring of Parkinson's disease using inertial sensors: a systematic review," *PLoS ONE*, vol. 16, no. 2, p. e0246528, 2021.

[8] S. Ancona, F. D. Faraci, E. Khatab, L. Fiorillo, O. Gnarra, T. Nef, C. L. A. Bassetti, and P. Bargiotas, "Wearables in the home-based assessment of abnormal movements in Parkinson's disease: a systematic review of the literature," *Journal of Neurology*, vol. 269, no. 1, pp. 100–110, 2022.

[9] M. F. Dirkx and M. Bologna, "The pathophysiology of Parkinson's disease tremor," *Journal of the Neurological Sciences*, vol. 435, p. 120196, 2022.

[10] T. A. Saifee, "Tremor," *British Medical Bulletin*, vol. 130, no. 1, pp. 51–63, 2019.

[11] K. P. Bhatia, P. Bain, N. Bajaj, R. J. Elble, M. Hallett, E. D. Louis, J. Raethjen, M. Stamelou, C. M. Testa, G. Deuschl et al., "Consensus statement on the classification of tremors, from the task force on tremor of the International Parkinson and Movement Disorder Society," *Movement Disorders*, vol. 33, no. 1, pp. 75–87, 2018.

[12] R. Powers, M. Etezadi-Amoli, E. M. Arnold, S. Kianian, I. Mance, M. Gibiansky, D. Trietsch, A. Singh Alvarado, J. D. Kretlow, T. M. Herrington et al., "Smartwatch inertial sensors continuously monitor real-world motor fluctuations in Parkinson's disease," *Science Translational Medicine*, vol. 13, no. 579, p. eabd7865, 2021.

[13] J. Raethjen, R. B. Govindan, M. Muthuraman, F. Kopper, J. Volkmann, and G. Deuschl, "Cortical correlates of the basic and first harmonic frequency of parkinsonian tremor," *Clinical Neurophysiology*, vol. 120, no. 10, pp. 1866–1872, 2009.

[14] G. Deuschl, J. Raethjen, R. Baron, M. Lindemann, H. Wilms, and P. Krack, "The pathophysiology of parkinsonian tremor: a review," *Journal of Neurology*, vol. 247, no. 5, pp. V33–V48, 2000.

[15] S. M. Ali, S. Poosapadi Arjunan, J. Peter, L. Perju-Dumbrava, C. Ding, M. Eller, S. Raghav, P. Kempster, M. A. Motin, P. J. Radcliffe, and D. K. Kumar, "Wearable accelerometer and gyroscope sensors for estimating the severity of essential tremor," *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 12, pp. 194–203, 2024.

[16] H. B. Kim, W. W. Lee, A. Kim, H. J. Lee, H. Y. Park, H. S. Jeon, S. K. Kim, B. Jeon, and K. S. Park, "Wrist sensor-based tremor severity quantification in Parkinson's disease using convolutional neural network," *Computers in Biology and Medicine*, vol. 95, pp. 140–146, 2018.

[17] J. E. Thorp, P. G. Adamczyk, H.-L. Ploeg, and K. A. Pickett, "Monitoring motor symptoms during activities of

daily living in individuals with Parkinson's disease," *Frontiers in Neurology*, vol. 9, p. 1036, 2018.

[18] N. Li, F. Tian, X. Fan, Y. Zhu, H. Wang, and G. Dai, "Monitoring motor symptoms in Parkinson's disease via instrumenting daily artifacts with inertia sensors," *CCF Transactions on Pervasive Computing and Interaction*, vol. 1, pp. 100–113, 2019.

[19] J. D. Loaiza Duque, A. M. González-Vargas, A. J. Sánchez Egea, and H. A. González Rojas, "Using machine learning and accelerometry data for differential diagnosis of Parkinson's disease and essential tremor," in *Proc. Workshop on Engineering Applications*, Santa Marta, Colombia, 2019, pp. 368–378.

[20] J. D. Loaiza Duque, A. J. Sánchez Egea, T. Reeb, H. A. González Rojas, and A. M. Gonzalez-Vargas, "Angular velocity analysis boosted by machine learning for helping in the differential diagnosis of Parkinson's disease and essential tremor," *IEEE Access*, vol. 8, pp. 88,866–88,875, 2020.

[21] A. C. Alves de Araújo, E. G. Da Rocha Santos, K. S. Guedes De Sà, V. K. Teixeira Furtado, F. A. Santos, R. C. De Lima, L. V. Krejcova, B. L. Santos-Lobato, G. H. Lima Pinto, A. dos Santos Cabral et al., "Hand resting tremor assessment of healthy and patients with Parkinson's disease: an exploratory machine learning study," *Frontiers in Bioengineering and Biotechnology*, vol. 8, p. 778, 2020.

[22] L. B. Peres, B. C. Calil, A. P. S. Paixão Barroso da Silva, V. C. Dionísio, M. F. Vieira, A. de Oliveira Andrade, and A. A. Pereira, "Discrimination between healthy and patients with Parkinson's disease from hand resting activity using inertial measurement unit," *BioMedical Engineering OnLine*, vol. 20, p. 50, 2021.

[23] A. Papadopoulos, K. Kyritsis, L. Klingelhofer, S. Bostanjopoulou, K. R. Chaudhuri, and A. Delopoulos, "Detecting parkinsonian tremor from IMU data collected in-the-wild using deep multiple-instance learning," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 9, pp. 2559–2569, 2019.

[24] A. Papadopoulos, K. Kyritsis, S. Bostanjopoulou, L. Klingelhofer, K. R. Chaudhuri, and A. Delopoulos, "Multiple-instance learning for in-the-wild parkinsonian tremor detection," in *Proc. 41st Annu. Int. Conf. IEEE Engineering in Medicine and Biology Society (EMBC)*, Berlin, Germany, 2019, pp. 6188–6191.

[25] L. Aljihmani, O. Kerdjidi, Y. Zhu, R. K. Mehta, and K. Qaraqe, "Detection of tremor associated with rest and effort activity using machine learning," in *Proc. Int. Conf. Automatics and Informatics (ICAI)*, Varna, Bulgaria, 2020, pp. 1–6.

[26] R. Ranjan, B. Bhushan, M. Palaniswami, and A. Verma, "A convolutional neural network approach for quantification of tremor severity in neurological movement disorders," in *Intelligent Systems and Applications: Proc. Intelligent Systems Conference (IntelliSys)*, vol. 3, London, U.K., 2021, pp. 416–429.

[27] G. AlMahadin, A. Lotfi, M. Mc Carthy, and P. Breedon, "Parkinson's disease tremor severity classification—a comparison between on and off medication state," in *Proc. Int. Conf. Innovative Techniques and Applications of Artificial Intelligence*, Cambridge, U.K., 2021, pp. 364–370.

[28] —, "Task-oriented intelligent solution to measure Parkinson's disease tremor severity," *Journal of Healthcare Engineering*, vol. 2021, pp. 1–15, 2021, retracted.

- [29] Y. Li, Z. Wang, and H. Dai, "Improved parkinsonian tremor quantification based on automatic label modification and SVM with RBF kernel," *Physiological Measurement*, vol. 44, no. 2, p. 025003, 2023.
- [30] L. Tong, J. He, and L. Peng, "CNN-based PD hand tremor detection using inertial sensors," *IEEE Sensors Letters*, vol. 5, no. 7, pp. 1–4, 2021.
- [31] H. Wu, Y. Zhang, X. Wu, and F. Yang, "Assessment of upper limb tremors in patients with Parkinson's disease based on displacement and acceleration information," in *Proc. 5th Int. Conf. Automation, Control and Robotics Engineering (CACRE)*, Dalian, China, 2020, pp. 177–182.
- [32] A. Ibrahim, Y. Zhou, M. E. Jenkins, M. D. Naish, and A. L. Trejos, "Parkinson's tremor onset detection and active tremor classification using a multilayer perceptron," in *Proc. IEEE Canadian Conf. Electrical and Computer Engineering (CCECE)*, London, ON, Canada, 2020, pp. 1–4.
- [33] A. Zhang, F. De la Torre, and J. Hodgins, "Comparing laboratory and in-the-wild data for continuous Parkinson's disease tremor detection," in *Proc. 42nd Annu. Int. Conf. IEEE Engineering in Medicine and Biology Society (EMBC)*, Montreal, QC, Canada, 2020, pp. 5436–5441.
- [34] R. San-Segundo, A. Zhang, A. Cebulla, S. Panev, G. Tabor, K. Stebbins, R. E. Massa, A. Whitford, F. De la Torre, and J. Hodgins, "Parkinson's disease tremor detection in the wild using wearable accelerometers," *Sensors*, vol. 20, no. 20, p. 5817, 2020.
- [35] M. Sun, A. Watson, G. Blackwell, W. Jung, S. Wang, K. Koltermann, N. Helm, G. Zhou, L. Cloud, and I. Pretzer-Aboff, "TremorSense: tremor detection for Parkinson's disease using convolutional neural network," in *Proc. IEEE Conf. Connected Health: Applications, Systems and Engineering Technologies (CHASE)*, Washington, DC, USA, 2021, pp. 1–10.
- [36] M. Sun, W. Jung, K. Koltermann, G. Zhou, A. Watson, G. Blackwell, N. Helm, L. Cloud, and I. Pretzer-Aboff, "Parkinson's disease action tremor detection with supervised-learning models," in *Proc. IEEE Conf. Connected Health: Applications, Systems and Engineering Technologies (CHASE)*, Orlando, FL, USA, 2023, pp. 1–10.
- [37] L. Sigcha, I. Pavón, N. Costa, S. Costa, M. Gago, P. Arezes, J. M. López, and G. De Arcas, "Automatic resting tremor assessment in Parkinson's disease using smartwatches and multitask convolutional neural networks," *Sensors*, vol. 21, no. 1, p. 291, 2021.
- [38] Z. Zhao, H. Yu, F. Wang, and H. Zheng, "Early detection of Parkinson's disease based on IMU and machine learning," in *Proc. SPIE Int. Conf. Computer Graphics, Artificial Intelligence, and Data Processing (ICCAID)*, vol. 12604, Guangzhou, China, 2023, pp. 1123–1129.
- [39] F. Rodriguez, P. Krauss, J. Kluckert, F. Ryser, L. Stieglitz, C. Baumann, R. Gassert, L. Imbach, and O. Bichsel, "Continuous and unconstrained tremor monitoring in Parkinson's disease using supervised machine learning and wearable sensors," *Parkinson's Disease*, vol. 2024, p. 5787563, 2024.
- [40] S. Dattola, A. Ielo, A. Quartarone, and M. C. De Cola, "Integrating wearable sensor signal processing with unsupervised learning methods for tremor classification in Parkinson's disease," *Bioengineering*, vol. 12, no. 1, p. 37, 2025.
- [41] G. Farhani, Y. Zhou, M. E. Jenkins, M. D. Naish, and A. L. Trejos, "Using deep learning for task and tremor type classification in people with Parkinson's disease," *Sensors*, vol. 22, no. 19, p. 7322, 2022.
- [42] S. Barrantes, A. J. Sánchez Egea, H. A. González Rojas, M. J. Martí, Y. Compta, F. Valldeoriola, E. Simo Mezquita, E. Tolosa, and J. Valls-Solè, "Differential diagnosis between Parkinson's disease and essential tremor using the smartphone's accelerometer," *PLoS ONE*, vol. 12, no. 8, p. e0183843, 2017.
- [43] S. A. Mousavi, M. H. Abdulrazzaq, M. A. Hasan, and M. Naghavizadeh, "Diagnosis of hand tremor using a smart phone accelerometer and SVM," in *Proc. 4th Int. Symp. Multidisciplinary Studies and Innovative Technologies (ISMSIT)*, Istanbul, Turkey, 2020, pp. 1–4.
- [44] E. Kuosmanen, F. Wolling, J. Vega, V. Kan, Y. Nishiyama, S. Harper, K. Van Laerhoven, S. Hosio, and D. Ferreira, "Smartphone-based monitoring of Parkinson disease: quasi-experimental study to quantify hand tremor severity and medication effectiveness," *JMIR mHealth and uHealth*, vol. 8, no. 11, p. e21543, 2020.
- [45] J. Habets, M. Heijmans, C. Herff, C. Simons, A. F. G. Leentjens, Y. Temel, M. Kuijf, and P. Kubben, "Mobile health daily life monitoring for Parkinson disease: development and validation of ecological momentary assessments," *JMIR mHealth and uHealth*, vol. 8, no. 5, p. e15628, 2020.
- [46] J. Varghese, M. Fujarski, T. Hahn, M. Dugas, and T. Warnecke, "The smart device system for movement disorders: preliminary evaluation of diagnostic accuracy in a prospective study," *Digital Personalized Health and Medicine*, vol. 270, pp. 889–893, 2020.
- [47] W. Liu, X. Lin, X. Chen, Q. Wang, X. Wang, B. Yang, N. Cai, R. Chen, G. Chen, and Y. Lin, "Vision-based estimation of MDS-UPDRS scores for quantifying Parkinson's disease tremor severity," *Medical Image Analysis*, vol. 85, p. 102754, 2023.
- [48] T. Gutowski, O. Stodulska, A. Ćwiklińska, K. Gutowska, K. Kopeć, M. Betka, R. Antkiewicz, D. Koziorowski, and S. Szlufik, "Machine learning-based assessment of Parkinson's disease symptoms using wearable and smartphone sensors," *Sensors*, vol. 25, no. 16, p. 4924, 2025.
- [49] V. K. Legaria-Santiago, L. P. Sánchez-Fernández, L. A. Sánchez-Pérez, and A. Garza-Rodríguez, "Computer models evaluating hand tremors in Parkinson's disease patients," *Computers in Biology and Medicine*, vol. 140, p. 105059, 2022.
- [50] J. J. Hathaliya, H. Modi, R. Gupta, and S. Tanwar, "Deep learning and blockchain-based essential and Parkinson tremor classification scheme," in *Proc. IEEE Conf. Computer Communications Workshops (INFOCOM WKSHPS)*, New York, NY, USA, 2022, pp. 1–6.
- [51] Y. Saez, C. Ureña, J. Valenzuela, A. García, and E. Collado, "A wearable Internet of Things-based device for the quantitative assessment of hand tremors in Parkinson's disease: the ELENA project," *Sensors*, vol. 25, no. 9, p. 2763, 2025.
- [52] A. Channa, R.-C. Ifrim, D. Popescu, and N. Popescu, "A-WEAR bracelet for detection of hand tremor and bradykinesia in Parkinson's patients," *Sensors*, vol. 21, no. 3, p. 981, 2021.
- [53] A. Channa, G. Ruggeri, N. Mammone, R.-C. Ifrim, A. Iera, and N. Popescu, "Parkinson's disease severity estimation using deep learning and cloud technology," in *Proc. Int. Conf. Omni-layer Intelligent Systems (COINS)*, Barcelona, Spain, 2022, pp. 1–7.
- [54] M. T. Ehsan, M. S. R. Sajal, and K. A. Mamun, "An automated cloudbased tool for screening of Parkinson's disease in Bangladesh," in *Proc. 2nd Int. Conf. Robotics, Electrical and*

Signal Processing Techniques (ICREST), Dhaka, Bangladesh, 2021, pp. 664–668.

[55] J. Babayan, M. Lueken, A. Berking, A. Pickartz, K. Reetz, F. Holtbernd, S. Leonhardt, and C. Ngo, ``Everyday life tremor signal processing in PD patients using BSN," in Proc. 6th Int. Conf. Advances in Biomedical Engineering (ICABME), Werdanyeh, Lebanon, 2021, pp. 184–188.

[56] L. Igene, A. Alim, M. H. Imtiaz, and S. Schuckers, ``A machine learning model for early prediction of Parkinson's disease from wearable sensors," in Proc. 13th Annu. Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, USA, 2023, pp. 734–737.

[57] N. Shawen, M. K. O'Brien, S. Venkatesan, L. Lonini, T. Simuni, J. L. Hamilton, R. Ghaffari, J. A. Rogers, and A. Jayaraman, ``Role of data measurement characteristics in the accurate detection of Parkinson's disease symptoms using wearable sensors," *Journal of NeuroEngineering and Rehabilitation*, vol. 17, p. 52, 2020.