ACCSCIENCE PUBLISHING

Artificial Intelligence in Health

ORIGINAL RESEARCH ARTICLE

Decoding Parkinsonian tremor: An explainable framework integrating spatial and spectral dynamics of multi-revolution spiral drawings

Tharaka Wijethunge[1], Maheshi Dissanayake[2]*, and Sajitha Weerasinghe[3]

[1]Department of Electrical, Computer, and Systems Engineering, Rensselaer Polytechnic Institute, Troy, New York, United States of America

[2]Department of Electrical and Electronic Engineering, Faculty of Engineering, University of Peradeniya, Peradeniya, Sri Lanka

[3]Neurology Unit, Teaching Hospital Peradeniya, Peradeniya, Sri Lanka

Abstract

Parkinson's disease manifests with motor impairments that are detectable through digitized spiral drawings. This study introduces an explainable framework for Parkinson's disease screening using a novel radial-sampling feature fusion approach. We transform two-dimensional spiral images into one-dimensional revolution signals via a systematic ray-sampling technique to extract three distinct revolutions. We integrate spatial metrics, such as inter-revolution spacing variability and root mean square radial derivatives, with spectral descriptors derived from fast Fourier transform analysis across low-, mid-, and high-harmonic bands. A total of 20 features were utilized to train state-of-the-art machine learning models, including support vector machines, random forests, and light gradient boosting machines. Among these, the random-forest classifier demonstrated superior performance. Subsequent five-fold cross-validation stability analysis, along with feature importance analysis, identified the root mean square radial derivative of the outer revolution (r3_derive_rms) as the most critical biomarker. Stratified cross-validation demonstrates that combining spatial and frequency features significantly enhances detection accuracy compared to single-domain methods, facilitating effective clinical deployment even in data-scarce environments. This interpretable pipeline provides a robust, low-cost "white-box" screening tool, offering a practical alternative to opaque deep-learning models for early clinical intervention.

Keywords: Parkinson's disease; Explainable artificial intelligence; Feature fusion; Spectral analysis; Spiral drawings; Biomarker identification

***Corresponding author:**
Maheshi Dissanayake
(maheshid@eng.pdn.ac.lk)

Citation: Wijethunge T, Dissanayake M, Weerasinghe S. Decoding Parkinsonian tremor: An explainable framework integrating spatial and spectral dynamics of multi-revolution spiral drawings. *Artif Intell Health.* 2026;3(4):026060012. doi: 10.36922/AIH026060012

Received: February 6, 2026

Revised: March 20, 2026

Accepted: March 24, 2026

Published online: May 8, 2026

Copyright: © 2026 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

Publisher's Note: AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

Parkinson's disease (PD) is a progressive neurodegenerative disorder that impairs motor control, producing clinical signs such as tremor, rigidity, bradykinesia, postural instability, cognitive changes, mood disorders, sleep disturbances, and autonomic dysfunction.[1] Multiple studies have used standardized handwriting and drawing tasks to quantify PD motor deficits such as bradykinesia.[2-4] The earliest symptoms

are considered to include alterations in handwriting kinematics.[5] More broadly, handwriting abnormalities in PD extend beyond micrographia and include impairments in kinematic, pressure, and fluency-related parameters.[6] Prior work reports characteristic PD-related handwriting abnormalities, including reduced writing amplitude and impaired letter formation, as measured by standardized writing tasks[7,8], although Archimedean spiral tracing has been widely studied as a low-cost, non-invasive proxy for detecting this motor dysfunction.[9,10] At present, there is no definitive test for PD, and the rate of misdiagnosis, particularly by non-experts, remains significant, with misdiagnosis rates reported to reach up to 20%.[11] Digitized versions of Archimedean spirals provide a computer-based, objective approach to quantify drawing impairment and support automated classification in PD.[12-14] Even brief, routine handwriting tasks have been shown to differentiate PD patients from healthy controls, supporting the use of handwriting as a practical, objective biomarker.[15]

In recent years, deep learning (DL) has demonstrated impressive capabilities in image classification tasks. Specifically, convolutional neural networks (CNNs), a subset of DL models, have been notably effective in biomedical image classification tasks. While DL holds significant promise for image classification, it is essential to recognize that it requires considerable computational resources, large datasets, and specialized graphics processing units[16] for effective training. While deep CNNs can achieve high accuracy on large curated datasets[17], practical model training scenarios often operate with small samples, heterogeneous devices, and require interpretable decisions that clinicians and patients can trust, rather than "black-box" behaviors.[18]

Current clinical assessment of PD largely relies on semi-quantitative scales, such as the Movement Disorder Society–Unified PD Rating Scale (MDS-UPDRS), which are susceptible to significant inter-rater variability and subjective interpretation. Consequently, there is a pressing need for objective, digital biomarkers that can provide reproducible measures of motor dysfunction. The limitations in the practical applicability of existing automated approaches—often characterized by high data dependency or "black-box" architectures—have motivated the development of a simpler, more explainable framework for identifying PD through spiral drawings.

This study addresses gaps in the literature by providing an interpretable pipeline that translates complex spiral trajectories into clinically meaningful spatial and spectral indices. Rather than applying a "black-box" type model, such as CNNs, directly to raw images, our method evaluates the spiral through a sequence of image analysis and feature extraction steps designed to handle diverse patient drawings while maintaining structural validity. We integrate multi-revolution spatial features with detailed spectral dynamics to capture subtle motor irregularities, such as tremor-induced fluctuations and inter-revolution spacing variability.

The proposed framework provides substantial clinical significance by offering an objective tool for characterizing Parkinsonian tremor through a non-invasive, low-cost approach. It should be noted that the literature provides evidence supporting the clinical relevance of the digitized Archimedes spiral test while highlighting the importance of methodological standardization.[19] The explainable architecture proposed in this research enhances clinician trust by clearly demonstrating the biomechanical and frequency-based features influencing the model's decisions, aligning with the growing demand for transparent digital biomarkers in neurology. Because the approach is easy to administer, it can be seamlessly incorporated into routine clinical workflows, telemedicine, and home-based monitoring. This work enables reliable longitudinal tracking of disease progression and quantification of therapeutic response—specifically for medication adjustments or deep brain stimulation—ultimately supporting individualized, data-driven management of PD.

2. Related work

The majority of the literature in this domain primarily focuses on training large-scale DL models to identify PD. Specifically, transfer learning using architectures such as DenseNet and ResNet50 has demonstrated high performance; however, these models typically require an extensive number of spiral images for training.[1] Another approach, proposed by Li *et al.*[17], focuses on combining genetic data and structural magnetic resonance imaging for data acquisition. While effective, this method is not cost-effective and poses significant challenges for wider deployment. In addition to transfer learning techniques, the literature indicates that gradient-weighted class activation mapping has been utilized to provide model explainability. Nevertheless, similar to other transfer learning applications, the required volume of training images remains extensive.[20]

In the handcrafted feature analysis approach, movement-error quantification techniques calculate discrepancies based on the distance between the template spiral and the actual drawing. These methods emphasize strong point correspondence, such as equivalent-angle pairing. While these approaches are interpretable, they mainly concentrate on geometric errors and do not explicitly account for physiological tremor signatures in the

frequency domain.[9,21] Although smartphone deployment in practical settings has been covered in some publications, their focus is primarily on tremor and may therefore be limited in robustness.[22,23]

Moreover, several studies have enhanced the reliability and interpretability of spiral metrics by tackling practical measurement challenges, such as center detection and improved spiral utility.[24] Additionally, several researchers have investigated specific indicators, such as temporal irregularity in spiral drawings[25], whereas others have focused on validating graphical and drawing tasks to achieve clinically meaningful outcomes rather than simply categorizing results.[3,26] In addition, kinematic and pressure-based handwriting variables have been reported as useful for the differential diagnosis of PD.[27] These feature-driven approaches can be sensitive to the tasks performed and the preprocessing methods used, and it is important to note that effective features may not transfer well across different devices or populations without a fully integrated end-to-end framework. Prior handwriting-based decision-support studies have similarly shown that carefully engineered markers can support PD discrimination without relying exclusively on end-to-end deep architectures.[28] A broader feature taxonomy combining kinematic, geometrical, and nonlinear handwriting descriptors has also shown strong relevance for modeling PD-related motor deficits.[29]

Numerous studies support the use of spiral analysis as an objective indicator of motor impairment in PD. Research highlights that these measures provide valuable diagnostic information even in the early stages of the disease and are significantly correlated with changes in upper-limb function and pharmacological response.[26,30] Furthermore, evidence suggests that spiral irregularities may be influenced by factors beyond disease severity, such as medication-related complications; this underscores the necessity of accounting for clinical status when interpreting spiral-based metrics.[4,26]

Recent advancements have enhanced the accessibility of PD spiral data collection through graphics tablet-based interfaces and smart pens, facilitating low-cost screening outside of specialized laboratory settings.[31,32] While various learning-based approaches currently classify PD by extracting frequency-domain features from digitized patterns and temporal data, emerging research is shifting toward real-time, tablet-based artificial intelligence applications to evaluate feasibility in clinical environments.[31,33-36] Consequently, there is a critical need for small-scale inference and feature-processing machine learning (ML) frameworks capable of operating on edge devices. This need is reinforced by studies showing that biometric handwriting analysis can support not only PD detection but also disease grading in clinically meaningful settings.[37] Developing systems that can retrain on new samples with limited computational resources while providing model explainability offers a more robust and practical alternative to traditional, resource-intensive DL methods. Recent studies in medical image classification have further highlighted the value of hierarchical feature fusion, ensemble learning, and attention-enhanced CNN designs for improving robustness and generalization across heterogeneous biomedical datasets.[38,39]

3. Methodology

3.1. Dataset

We used two sources of spiral drawings: (i) a set collected in a clinical setting using a simple, fixed procedure, and (ii) a small, carefully selected group from public datasets to include different types of paper, pens, and lighting.[40,41]

The clinical data for the study were collected from the teaching hospital of the Central Province, which serves a large geographical catchment area. All patients were assessed and diagnosed by a single neurology team to reduce interobserver variability. To ensure diagnostic specificity, all participants underwent differential screening to exclude essential tremor, dystonic tremors, and metabolically induced tremors, focusing exclusively on idiopathic PD as per MDS-UPDRS criteria. All the patients were given A4-sized blank white paper and a standard ballpoint pen. The patients were asked to sit comfortably in the clinic with their arms resting on the table. Next, they were asked to study the given reference spiral image and draw a similar spiral on a blank sheet of paper without lifting their hand off the paper.

Figure 1 displays typical sample drawings of a healthy individual and an individual with severe PD. These drawings were either scanned or photographed and subsequently processed using a consistent set of steps for each image: straightening the page, adjusting the lighting and background, converting to grayscale, binarizing to emphasize the ink, and tightly cropping around the spiral. We resized the images to a uniform square dimension while preserving their original shape and aspect ratio. We removed any drawings that were incomplete, had heavy shadows, or displayed other obvious issues. For the training and test sets, we retained images with clear strokes and sufficient contrast so that the full spiral is visible. This approach enables us to evaluate how effectively our methods work on everyday photos, rather than just on clinic scans. Moreover, if a subject had multiple spirals, we selected at most one for model training or testing to prevent overlap. When possible, we included simple notes to help us assess performance across different groups.

Following this careful screening, 30 spirals were collected in a clinical setting.

To enhance model generalizability, we combined our clinical dataset with spiral images from previously published, publicly available PD benchmark datasets[40,41] to create the final training dataset used for model development. Furthermore, a healthy control cohort was incorporated to serve as a baseline reference for comparative analysis. The final consolidated dataset comprised 116 high-resolution spiral images, providing a diverse foundation of both symptomatic and healthy signatures for robust feature extraction and classification. This balanced approach ensures that the model can effectively distinguish subtle motor impairments from normal physiological variations. We initially developed and optimized our models using public image datasets[40,41] and subsequently refined the model with clinical data. Unless otherwise specified, we use a 70/15/15 train/validation/test split, ensuring class balance and, when feasible, maintaining severity levels. Additionally, we conduct cross-source tests, where models trained on public images are directly evaluated on a clinical test set that has not been previously accessed.

3.2. Radial distance extraction from spiral drawing

3.2.1. Spiral segmentation

Each red–green–blue spiral image $\left(I \in R^{H\times W\times 3}\right)$ is first converted to a grayscale representation, G, where each pixel denotes a normalized intensity value. To enhance the signal-to-noise ratio, G is preprocessed using either a median or a bilateral filter for noise reduction. Subsequently, a binary mask $B \in [0,1]^{H\times W}$ is generated via global thresholding (**Equation 1**):

$$B(i,j)=\begin{cases}1, & G(i,j)\le \tau^{*}- \quad \text{(i.e., belongs to foreground)},\\ 0, & \text{otherwise}- \quad \text{(i.e., belongs to background)}.\end{cases} \tag{1}$$

where τ^{*} represents the optimal threshold selected to maximize the separability between the foreground (ink) and background classes. To ensure robustness against non-uniform illumination and scanning artifacts, we incorporate adaptive thresholding based on local brightness variations, followed by morphological opening and closing operations using a *3 × 3* structuring element to eliminate isolated artifacts and bridge small discontinuities in the binarized trace, thereby ensuring a continuous and representative spiral path.

3.2.2. Spiral center detection using circular statistics

To accurately characterize the spiral geometry, the center $\hat{c}=\left(\widehat{c_x},\widehat{c_y}\right)$ must be precisely identified. Let S denote the set of foreground pixel coordinates $\left(x_k,y_k\right)$ extracted from the binary mask *B*, such that $P\{(x,y)\,|\,B(x,y)=1\}$. We estimate the optimal spiral center by identifying the candidate point c for which the foreground pixels exhibit the most uniform angular coverage when expressed in polar coordinates. Unlike traditional centroid-based methods, this angular-uniformity approach provides robustness against image artifacts, scanning noise, and non-uniform stroke distributions that typically bias the center of mass. Its effectiveness was empirically evaluated against centroid- and distance-transform–based alternatives. Following the principles of circular statistics, the estimation is formulated as in **Equations 2–4**.

$$\hat{c}=\arg\max_{c} V(c) \tag{2}$$

$$V(c)=1-\left|\frac{1}{N_s}\sum\nolimits_{k=1}^{N_s} e^{j\theta_k(c)}\right| \tag{3}$$

$$\theta_k(c)=\angle\left((x_k,y_k)-c\right) \tag{4}$$

In this framework, $\theta_{\kappa}(c)$ represents the polar direction of the k-th foreground pixel relative to a candidate center c, where N_s is the total number of extracted foreground pixels $\left(N_s=|P|\right)$. The term within the absolute value bars in **Equation 3** represents the mean resultant vector length; a value approaching zero indicates that the pixel directions are distributed uniformly around the center, thereby maximizing the circular variance $V(c)$. Consequently, the candidate coordinate that maximizes $V(c)$ is selected as the most geometrically centered origin, providing a stable basis for subsequent feature extraction.

3.2.3. Coordinate transformation and scale normalization

To ensure that the extracted features are comparable across different drawings and varying acquisition scales, the raw spiral coordinates $\left(x_k,y_k\right)$ are transformed relative to the estimated center $\hat{c}$ and subsequently normalized by the maximum observed radius $\mathrm{R}_{\max}$. This process maps the entire spiral trace to a unit circle, as defined in **Equations 5** and **6**:

$$\overline{x_k}=\frac{x_k-\widehat{c_x}}{R_{max}}, \qquad \overline{y_k}=\frac{y_k-\widehat{c_y}}{R_{max}} \tag{5}$$

$$R_{max}=\max(k)\sqrt{\left(x_k-\widehat{c_x}\right)^2+\left(y_k-\widehat{c_y}\right)^2} \tag{6}$$

Following this transformation, the spiral's boundary maps to a radius of 1. All subsequent computations, including spectral analysis and irregularity measurements, utilize these normalized coordinates $\left(\overline{x_k},\overline{y_k}\right)$ to mitigate the influence of drawing size and focus strictly on the

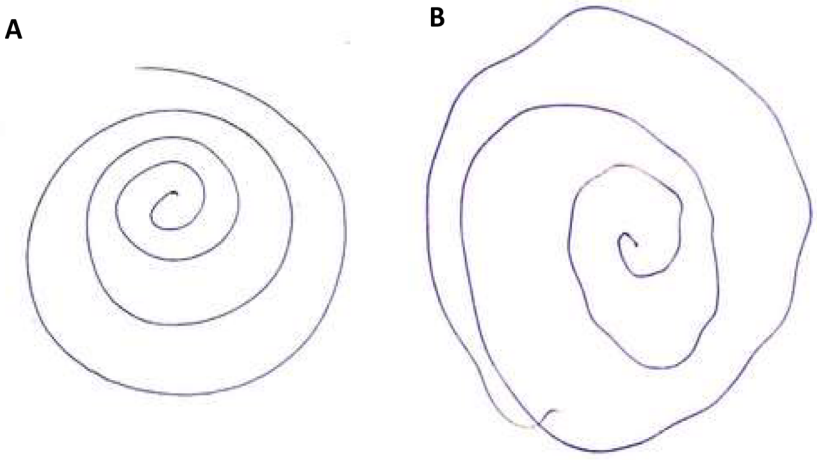


Figure 1. Spiral drawing by (A) a healthy control and (B) an individual with Parkinson's disease

morphological and kinematic patterns of the patient's handwriting.

3.3. Feature extraction and signal preprocessing

3.3.1. Radial feature extraction from spiral revolutions

To capture the fine-grained geometric properties of the hand-drawn spirals, we convert the two-dimensional binarized images into three distinct one-dimensional "revolution" (Rev) signals through a systematic ray-sampling technique. This process is initiated from the estimated center $(\widehat{c_x}, \widehat{c_y})$ and utilizes a uniform angular grid, where each sampling angle θ_j is defined as in **Equation 7**.

$$\theta_j = j\Delta\theta,\ j = 0,\ldots,N-1,\quad \Delta\theta = 5^\circ,\quad N = \frac{360^\circ}{\Delta\theta} \tag{7}$$

For each angular step, a radial ray is cast from the center. We detect consecutive foreground (ink) pixels along this ray at small radial increments, identifying the first three valid intersections that satisfy a minimum revolution length and are separated by a distinct background gap. These intersections correspond to the first (innermost), second, and third spiral revolutions. Figures 2 and 3 illustrate sample labeling of Rev 1, 2, and 3, distinguished by three different colors to visualize the extraction process. The radius for each revolution, $r_k[j]$, is recorded as the mean radial distance of all pixels within that specific intersection. This produces three discrete angular signals as in **Equation 8.**

$$r_k[j] = r_k(\theta_j),\quad k \in \{1,2,3\} \tag{8}$$

where $r_k[j]$ is measured in pixels. In instances where a ray intersection is not identified at a specific angle, the value is set to NaN, to preserve the temporal integrity of the signal for downstream analysis.

3.3.2. Signal conditioning and missing data handling

Extracted signals $r_k[j]$ often contain discontinuities due to intermittent pen contact or faint strokes. To ensure signal integrity for subsequent analysis, we implement a two-stage circular restoration process:

(i) Circular linear interpolation: Small angular gaps are filled via linear interpolation along the circular axis to maintain the periodic nature of the spiral trace.
(ii) Circular median filtering: A median filter with a window length W and wrap-around indexing is applied to mitigate high-frequency noise while preserving geometric trends.

While large gaps are excluded from general statistical analyses, they are restored via circular interpolation specifically for fast Fourier transform computations to ensure signal continuity required for spectral analysis.

3.3.3. Residuals and inter-revolution spacing

To isolate local irregularities, we derive two additional one-dimensional signal sets from the Rev radii. First, each revolution is centered by its median to obtain the residual signal $e_k[j]$ as in **Equation 9**, which captures deviations from the typical radius while remaining robust to outliers:

$$e_k[j] = r_k[j] - \text{median}\{r_k[j]\} \tag{9}$$

Second, we quantify the consistency of the spiral's expansion by measuring the spacing $s_{ab}[j]$ between adjacent revolutions for $(a,b) \in \{(1,2),(2,3)\}$ as in **Equation 10**.

$$s_{ab}[j] = r_b[j] - r_a[j] \tag{10}$$

These signals summarize how the gap between successive turns fluctuates with the angle, providing a set of critical indicators of motor dysfunction.

3.4. Spatial feature extraction

Spatial features are derived from the raw revolution signals $r_k[j]$, their robust residuals $e_k[j]$, and the inter-revolution spacing $s_{ab}[j]$, as defined in **Equations 8–10**, respectively. All statistics are calculated over the subset of angles where the required signal values are finite.

3.4.1. Kinematic and roughness descriptors

These features quantify the high-frequency fluctuations and micro-irregularities characteristic of Parkinsonian tremors.

(i) Root mean square (RMS) radial derivative ($r_{3_\mathrm{rms}_\mathrm{deriv}}$): Measures rapid angular fluctuations in the outer revolution by computing the RMS of the residual derivative, as in **Equation 11**:

$$\mathrm{RMS}_{deriv} = \sqrt{\frac{1}{N}\sum_j \left(e_3[j] - e_3[j-1]\right)^2} \tag{11}$$

(ii) Zero-crossing rate (r_{1_zcr}): Quantifies the frequency at which small angular fluctuations in the innermost revolution switch direction. Let $\mathbb{1}(\cdot)$ be the indicator function; the r_{1_zcr} rate is given by **Equation 12**:

$$\text{Zero-crossing rate} = \frac{1}{N-1}\sum_{j=1}^{N-1} \mathbb{1}\left(e_1[j]\cdot e_1[j-1] < 0\right) \tag{12}$$

A higher zero-crossing rate indicates increased oscillatory, tremor-like variations.

(iii) Robust roughness ($r_{2_\mathrm{res}_\mathrm{mad}}$): A measure of instability for the middle revolution based on the median absolute deviation of the residuals is given by **Equation 13**:

$$\textit{Median absolute deviation} = \textit{median}\left(\left|e_2[j] - \textit{median}(e_2)\right|\right) \tag{13}$$

Since e_k is centered by its median, the second term is typically near zero, isolating the raw dispersion of the signal.

3.4.2. Geometric and morphological features

These metrics describe the global shape and continuity of the spiral:

(i) Outer radius statistics ($r_{3_\mathrm{mean}}, r_{3_\mathrm{median}}$): Summarize the typical scale of the spiral's outer boundary, as in **Equation 14**:

$$\mu_{r3} = \frac{1}{N}\sum r_3[j], \quad \tilde{r}_3 = \mathrm{median}(r_3) \tag{14}$$

(ii) Maximum missing segment ($r_{1_\mathrm{max}_\mathrm{gap}_\mathrm{deg}}$): Identifies the largest contiguous angular gap in the first revolution. If M is the maximum number of consecutive missing bins, the gap in degrees is given by **Equation 15**:

$$\mathrm{Gap}_{deg} = M \cdot \Delta\theta \tag{15}$$

3.4.3. Inter-revolution consistency and overlap

Inter-revolution features characterize the stability of the spiral's expansion.

(i) Spacing statistics (s_{12_mean}, s_{12_cv}, s_{12_cv}): For the gap between the first and second revolutions, $s_{12}[j]$, we compute the mean (μ_{s12}), standard deviation (σ_{s12}), and coefficient of variation (CV_s), as in **Equation 16**:

$$CV_s = \frac{\sigma_{s12}}{\mu_{s12}} \tag{16}$$

(ii) Revolution crossing fraction (*s*_{12_crossings}): Measures the frequency of irregular overlaps (where the spiral "crosses" itself) by calculating the fraction of angles where spacing becomes negative, as in **Equation 17**.

$$Crossing \cdot rate = \frac{1}{N}\sum_j \mathbb{1}(s_{12}[j] < 0) \tag{17}$$

(iii) Outer spacing consistency (s_{23_std}): Captures the variability in the gap between the second and third revolutions to assess turn-to-turn uniformity.

3.5. Frequency-domain feature extraction

Frequency-domain features used in this study are derived from the discrete Fourier transform of each revolution signal. To ensure spectral continuity, missing samples in the radial signals are first restored via circular interpolation. Let $\widetilde{r_k}[j]$ represent the interpolated version of $r_k[j]$. To isolate oscillatory content from the global spiral trend, the direct current component is optionally removed following **Equation 18**:

$$\widehat{r_k}[j] = \widetilde{r_k}[j] - \mathrm{mean}\left(\widetilde{r_k}\right) \tag{18}$$

The discrete Fourier transform magnitude is then computed for each harmonic index m using **Equation 19**:

$$M_k[m] = \left|\sum_{j=0}^{N-1} \hat{r}_k[j] e^{-j\frac{2\pi}{N}mj}\right| \tag{19}$$

Next, we categorize spectral energy into three harmonic bands to differentiate between geometric structure and pathological tremors: low band (m = 2–4, which captures the fundamental spiral geometry), mid band (m = 5–8, which represents intermediate fluctuations), and high band (m = 9–12, which isolates high-frequency micro-tremors).

3.5.1. Spectral band energies

We quantify the spectral distribution by calculating the cumulative energy within specific harmonic bands (low, mid, and high), providing a profile of the tremor's intensity across different scales:

(i) Band energies ($E_{low}, E_{mid}, E_{high}$): Computed as the

sum of squared magnitudes within defined harmonic ranges as in **Equation 20**:

$$E_{band} = \sum_{m \in \text{band}} \left(M_k [m] \right)^2 \quad (20)$$

For example, the band energies for Rev 1 (i.e., $k = 1$) are represented as $E_{low} = \sum_{h=2}^{4} M_1[m]^2$, $E_{mid} = \sum_{h=5}^{8} M_1[m]^2$, and $E_{high} = \sum_{h=9}^{12} M_1[m]^2$. Similarly, the band energies for Rev 2 and Rev 3 can be expressed following **Equation 20**.

(ii) High-to-low spectral ratio ($r_{1_hilo_ratio}$): This ratio characterizes the dominance of high-frequency micro-tremors relative to lower-frequency geometric components as in **Equation 21**:

$$Ratio_{HL} = \frac{E_{high}}{E_{low}} \quad (21)$$

3.5.2. Dominant oscillatory and decay components

To further distinguish pathological movements from healthy handwriting, we analyze the peak intensity and the rate of spectral roll-off:

(i) Dominant amplitude ($r_{1_dom_amp}$): Captures the maximum non-direct current magnitude, representing the most prominent oscillatory component in the first revolution, as in **Equation 22**.

$$A_{dom} = \max_{m>0} M_k [m] \quad \text{k} = \{1,2,3\} \quad (22)$$

(ii) Spectral magnitude decay ($r_{k_spec_decay}$): Quantifies the rate at which spectral energy decreases as the harmonic index increases. This is determined by fitting a linear regression to the log-magnitude spectrum as in **Equation 23**.

$$\log(M_k[m]) \approx \alpha \cdot m + \beta \quad (23)$$

A more negative decay constant (α) indicates a steeper drop-off, signifying a lack of high-frequency content typical of healthy, smooth handwriting.

3.5.3. Spectral cross-correlation

To assess the consistency of motor dysfunction across the drawing task, we measure the spectral similarity between the first and second revolutions using the Pearson correlation coefficient over shared harmonic indices as in **Equation 24**.

$$\rho_{12} = \text{corr}(M_1[m], M_2[m]) \quad (24)$$

3.6. Classification and performance evaluation

In this study, a total of 20 features—comprising the spatial and frequency-domain descriptors detailed in Sections 3.4 and 3.5—were utilized to train and evaluate state-of-the-art ML models to detect the presence of PD. The final preprocessed output, illustrating the clear morphological and spectral distinctions between a healthy control and a patient with PD, is depicted in Figures 2 and 3. These figures visualize the extraction process by highlighting three spiral revolutions (Revs 1, 2, and 3) in distinct colors. As shown in the provided sample images, the first (innermost), second, and third intersections with the radial ray are clearly identified. Furthermore, each individual sampled pixel position is marked to emphasize the differences in stroke consistency and geometric regularity between the two groups.

4. Machine learning model training

We treat PD screening from spiral drawings as a supervised binary classification problem using the extracted spiral-deformation feature vector . The target labels are encoded as $y \in \{0, 1\}$ using a standard label encoder. The class assignments are defined such that the healthy class is represented by 0 (negative class) and the PD class is represented by 1 (positive class).

4.1. Cross-validation protocol and metrics

To ensure the robustness and reproducibility of our results, we implemented a stratified five-fold cross-validation scheme. Data shuffling was enabled with a fixed random seed to strictly preserve the healthy-to-PD class proportions across all folds. For each model architecture, hyperparameter optimization was conducted via an exhaustive grid search, utilizing the area under the receiver operating characteristic (ROC–AUC) curve as the primary objective function for model selection. Upon identifying the optimal configurations, we computed cross-validated predictions to assess performance. We report a comprehensive suite of metrics, including accuracy, ROC–AUC, and the confusion matrix, alongside per-class precision, recall, and F1-score (pooled across folds). Additionally, ROC curves were generated using the cross-validated predicted probabilities to visualize the diagnostic trade-offs of each classifier.

4.2. Model selection

Leveraging the proposed set of 20 spiral-based features, we conducted a systematic comparative analysis of six representative ML classifiers: support vector machine, logistic regression, random forest (RF), gradient boosting, XGBoost, LightGBM, and Vanilla CNN. To ensure an unbiased performance assessment, all models were rigorously evaluated under a uniform stratified cross-validation protocol. Furthermore, to address the inherent class imbalance between healthy and PD samples, we implemented class weighting or minority-class

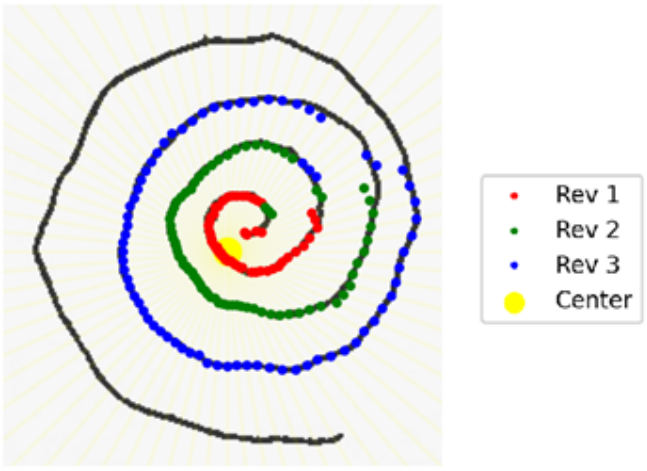


Figure 2. Fully preprocessed output used for feature extraction for the healthy sample
Abbreviation: Rev: Revolution.

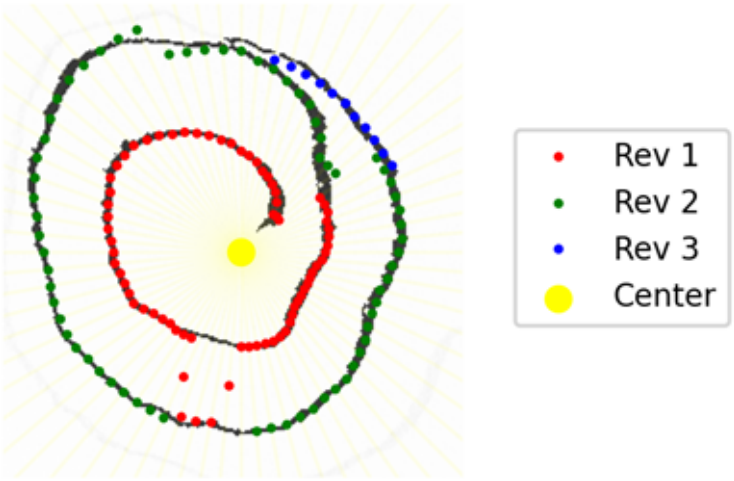


Figure 3. Fully preprocessed output used for feature extraction for the Parkinson's disease sample
Abbreviation: Rev: Revolution.

upweighting strategies across all architectures. The specific hyperparameter configurations and the grid search ranges optimized for each model are summarized in Table 1.

5. Results: Observations, validations, and analysis

5.1. Validation of center detection strategy

To empirically validate the center-detection stage, we conducted an ablation study comparing three alternatives: foreground centroid (centroid), distance-transform peak (DT peak), and the proposed hybrid center estimator. For each method, the full radial sampling, feature extraction, and RF classification pipeline was rerun using identical five-fold stratified cross-validation.

As shown in Table 2, the proposed method achieved the highest ROC–AUC (0.751), and the best feature extraction completeness, quantified by mean angular coverage (0.791) and the lowest mean maximum gap (59.24°). Although the centroid baseline yielded slightly higher classification accuracy (0.690 vs. 0.655), it produced lower coverage and substantially larger missing angular regions. These results indicate that the proposed center detector provides a stronger basis for robust spiral feature extraction and improves discrimination performance when assessed by ROC–AUC.

5.2. Sensitivity analysis of harmonic band definitions

To assess whether the frequency-domain representation

Table 1. Hyperparameter configurations and grid search ranges

Classifier	Implementation details	Hyperparameters and search space
Support vector machines with RBF	RBF kernel; class-balanced weights	Regularization strength (C) ∈ {0.3, 1, 3, 10}; kernel width (γ) ∈ {0.001, 0.003, 0.01, 0.03}
Logistic regression	ℓ2 regularization; class-balanced weights	The inverse regularization strength (C^{-1}) ∈ {0.1, 1, 10}
Random forest	400 trees; class-balanced weights	Maximum depth ∈ {none, 10, 20}; minimum samples split ∈ {2, 5}
Gradient boosting	Decision tree base; deviance loss	n_estimators ∈ {200, 400}; learning rate ∈ {0.05, 0.1}
XGBoost	Logistic objective; ratio-based upweighting to address class imbalance	n_trees ∈ {200, 400}; maximum depth ∈ {3, 6}; learning rate ∈ {0.05, 0.1, 0.2}; subsample ∈ {0.8, 1.0}
LightGBM	Histogram-based; class-balanced weights	n_trees ∈ {200, 400}; maximum leaves ∈ {15, 31, 63}; learning rate ∈ {0.05, 0.1, 0.2}
Vanilla convolutional neural network	Three convolutional layers with dropout and data augmentation	Optimizer = adam; batch_size = 8; loss = binary_crossentropy; input size = 128 × 128 × 3

Abbreviation: RBF: Radial basis function.

was overly dependent on the manually selected harmonic bands, we performed a sensitivity analysis using several nearby alternative partitions within harmonics 2–12. Specifically, the original band definition (2–4, 5–8, 9–12) was compared against three alternatives: (2–3, 4–7, 8–12), (2–5, 6–8, 9–12), and (2–4, 5–7, 8–12). For each configuration, the feature extraction and RF classification pipeline was evaluated over repeated random subsampling runs. As shown in Table 3, performance varied only modestly across the tested band definitions. The original band partition achieved the highest mean ROC–AUC (0.735), while the remaining alternatives produced comparable results. These findings indicate that the frequency-domain representation is not critically dependent on a single arbitrary band boundary and that the original partition is a reasonable and stable choice.

5.3. Primary classifier selection

Figures 4 and 5 characterize the spatial-domain feature extraction, specifically depicting the variations in radial distance from the centroid across successive revolutions for the PD samples. To complement this spatial analysis, Figures 6 and 7 provide the corresponding fast Fourier transform spectra. These visualizations demonstrate that the inter-radial distance fluctuations and their associated frequency responses exhibit discernible patterns, providing a robust basis for differentiating between classes.

The cross-validated performance and optimal hyperparameter configurations for each model are summarized in Table 4. The results indicate that the RF classifier achieves the highest diagnostic performance, yielding superior results in both accuracy and ROC–AUC. The discriminative power of the RF model is further illustrated by the ROC curve in Figure 8. Consequently, based on its robust performance across all evaluation metrics, RF was selected as the primary classifier for all subsequent analyses.

To further evaluate the diagnostic reliability of the selected RF classifier, the confusion matrix is presented in Figure 9. An analysis of the classification results reveals balanced performance across both cohorts, with the model achieving an F1-score of 0.7477 for the healthy class and 0.7840 for the PD class. This consistency across classes highlights the model's ability to minimize both false positives and false negatives, ensuring a robust diagnostic

Table 2. Empirical comparison of center-detection methods for spiral feature extraction and classification

Center method	Mean coverage fraction	Mean number of gaps	Mean maximum gap (°)	RF accuracy	RF ROC–AUC
Proposed hybrid method	0.791	1.198	59.24	0.655	0.751
Centroid	0.696	1.399	116.75	0.690	0.743
DT peak	0.258	1.072	277.67	0.629	0.746

Abbreviations: DT: Distance transform; RF: Random forest; ROC–AUC: Area under the receiver operating characteristic.

Table 3. Sensitivity analysis of harmonic band definitions

Scheme	Low band	Mid band	High band	Mean RF accuracy	Mean RF ROC–AUC
Original	2–4	5–8	9–12	0.692	0.735
Scheme A	2–3	4–7	8–12	0.679	0.727
Scheme B	2–5	6–8	9–12	0.693	0.728
Scheme C	2–4	5–7	8–12	0.693	0.731

Abbreviations: RF: Random forest; ROC–AUC: Area under the receiver operating characteristic curve.

framework for biomedical applications.

5.4. Explainability: Feature importance and interpretation

Given that the framework utilizes explicitly engineered radial and spectral biomarkers, we evaluate model explainability through global feature importance rankings. Based on Figure 10, the top-ranked attributes for the RF classifier were identified, highlighting the dominance of deformation-centric measures over static morphological traits.

To validate the robustness of these importance rankings, a five-fold cross-validation stability analysis was conducted. The results of this analysis are presented in Figure 11. By calculating the coefficient of variation (CV) across the five folds (Table 5), we observed that the primary discriminative features, specifically those related to radial derivatives and residual peak-to-peak fluctuations, exhibited high stability (CV < 20%). This consistency confirms that the model is capturing reproducible pathological signals rather than overfitting to noise within the cohort (n = 116). These findings reinforce the framework's utility as a stable and interpretable clinical decision-support tool.

Furthermore, the model's predictive power is primarily driven by roughness and jerkiness metrics, spacing variability, and band power-derived descriptors. These results underscore the model's reliance on capturing subtle high-frequency irregularities and altered stroke dynamics (micro-tremors) characteristic of Parkinsonian motor dysfunction. Unlike "black-box" DL architectures, the features utilized here are inherently interpretable; for instance, increased spectral energy in specific bands corresponds directly to physical oscillatory patterns associated with resting or action tremor.

To ensure clinical relevance, these importance rankings were qualitatively cross-referenced with the expert observations of the study's consulting neurologist (SW). The high importance assigned to frequency-domain markers provides empirical alignment with known pathological handwriting characteristics, such as diminished motor consistency and rhythmic instability. While this analysis establishes global interpretability, confirming that the model prioritizes clinically sound biomarkers, we acknowledge that "case-level" interpretability (e.g., patient-specific decision visualizations) remains an area for future development. Nonetheless, the current feature-driven approach ensures that the model's logic remains

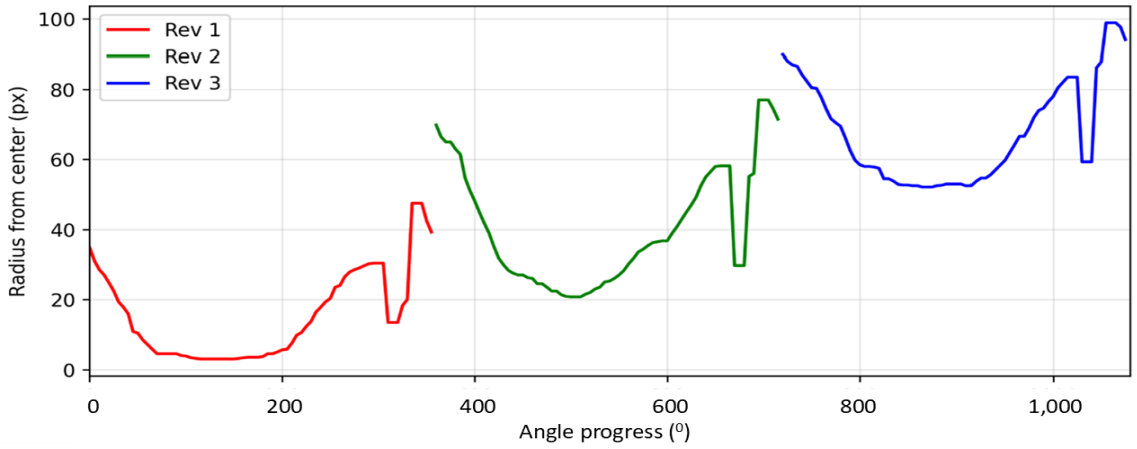


Figure 4. Radial distance for each revolution (Rev) for healthy control samples

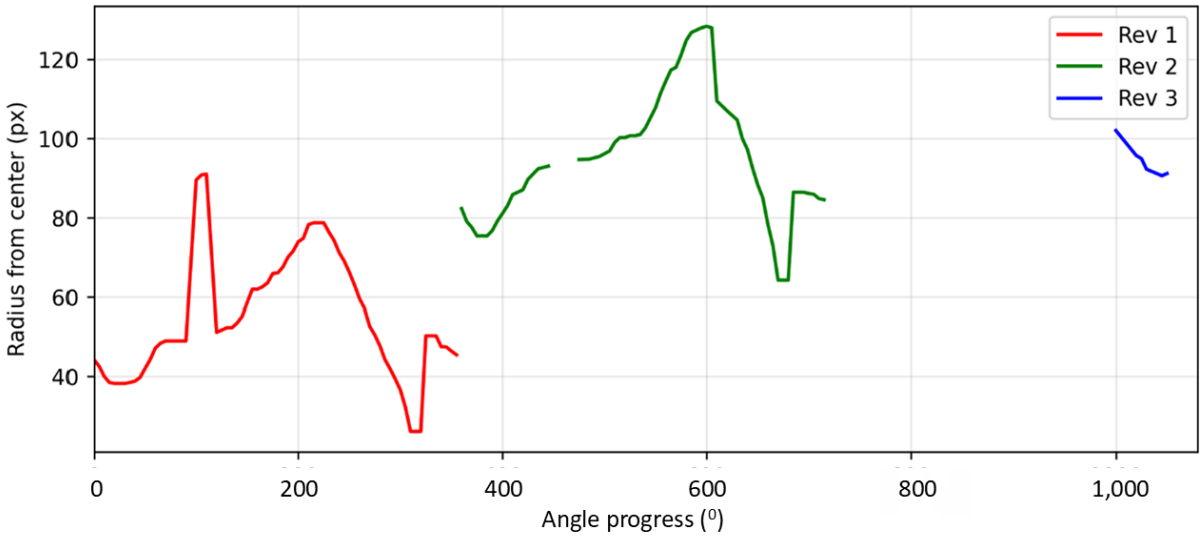


Figure 5. Radial distance for each revolution (Rev) for Parkinson's disease samples

transparent and auditable by medical professionals, a key requirement for clinical decision-support systems.

6. Discussion

6.1. Clinical significance of radial and spectral biomarkers

The identification of the RMS radial derivative, $r3_deriv_rms$, and $r3_mean$ as the most important features for diagnosis provides clinicians with a quantifiable measure of motor fluctuations. This aligns with prior work showing that digital spiral analysis can serve as a quantitative outcome measure for tremor severity in movement disorders.[42] In clinical practice, these metrics can serve as an automated "tremor signature." For instance, a high $r_{1_hilo_ratio}$ value directly correlates with the intensity of micro-tremors, allowing for objective monitoring of medication efficacy. Furthermore, the three-revolution radial sampling approach ensures that patient fatigue, often visible in the transition from the inner to the outer revolution, is captured and measured, providing insights into motor endurance that are often missed in traditional visual assessments.

6.2. Generalizability and robustness

A key strength of this study is the high degree of source heterogeneity within the dataset. By integrating expert-verified clinical samples collected under medical

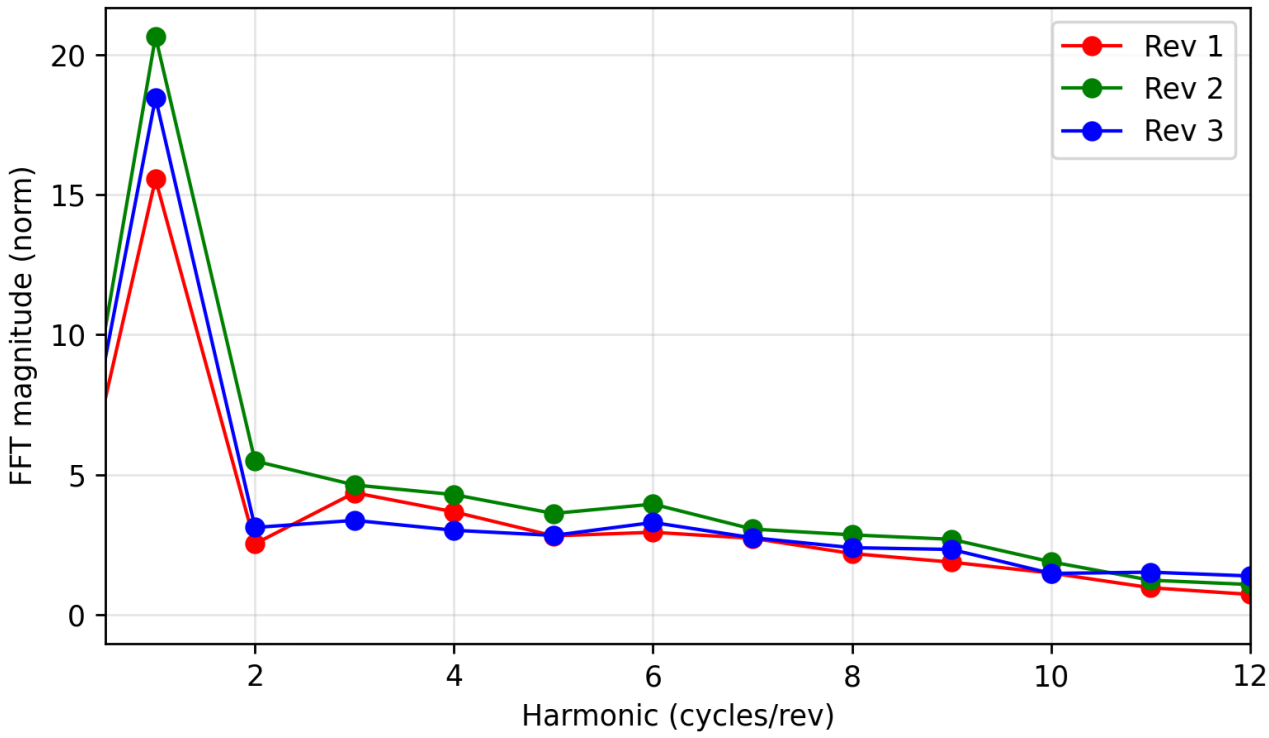


Figure 6. Fast Fourier transform (FFT) magnitude plot for each revolution (Rev) for healthy control samples

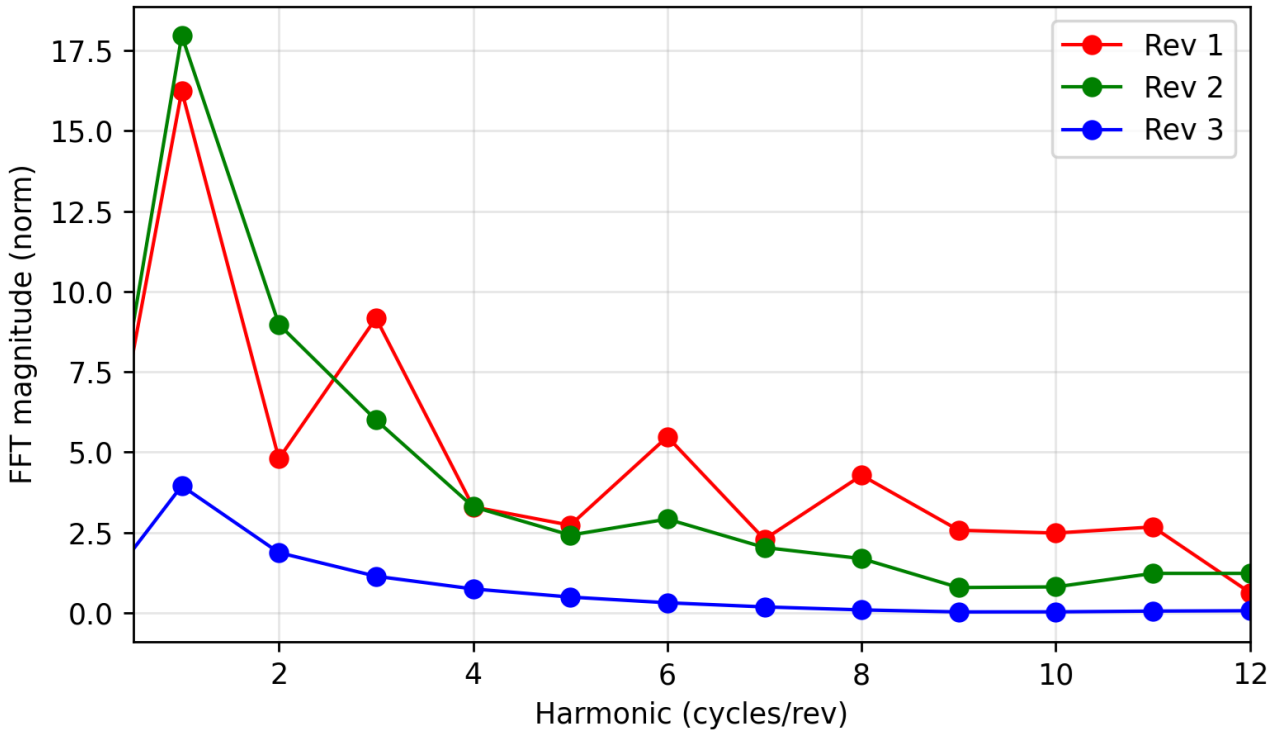


Figure 7. Fast Fourier transform (FFT) magnitude plot for each revolution (Rev) for Parkinson's disease samples

Table 4. Five-fold stratified cross-validation

Model	Best hyperparameters	Accuracy	ROC–AUC
Support vector machines with radial basis function	C = 1; γ = 0.03	0.698	0.729
Logistic regression	$C^{-1} = 0.1$	0.698	0.678
Random forest	Maximum depth = unlim; Minimum samples split = 2	0.767	0.832
Gradient boosting	n_estimators = 200; lr = 0.05	0.724	0.780
XGBoost	n_trees = 200; maximum depth = 6; lr = 0.2; subsample = 0.8	0.733	0.821
LightGBM	n_trees = 400; maximum leaves = 15; lr = 0.2	0.759	0.831
Vanilla convolutional neural network	Optimizer = adam; ConvLayers = 3; loss = binary_crossentropy; metrics = accuracy; input = 128 × 128 × 3	0.50	0.50

Abbreviation: ROC–AUC: Area under the receiver operating characteristic curve.

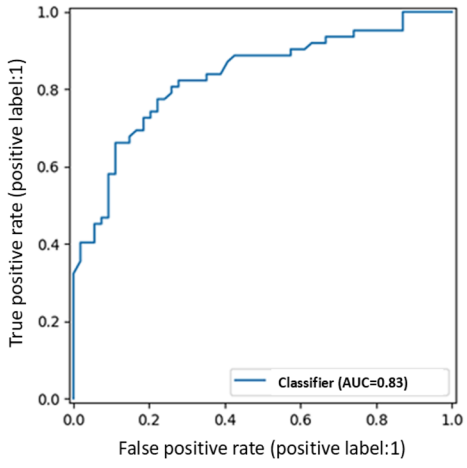


Figure 8. ROC curve for the selected primary random forest classifier model
Abbreviations: AUC: Area under the curve; ROC: Receiver operating characteristic.

supervision across the Central Province of Sri Lanka with diverse global public databases, the framework was exposed to digitized spiral features that remain invariant across different hardware, collection environments, and demographic variations. This multi-source training acts as a natural stress test, mitigating the risk of cohort-specific overfitting and suggesting that the reported performance is grounded in fundamental pathological dynamics rather than institutional- or device-specific artifacts.

This robustness is further validated by the necessity of the feature-engineering approach over automated feature learning. To verify this, a vanilla CNN was trained and evaluated on the same raw two-dimensional image data to assess whether an end-to-end architecture could autonomously handle the data heterogeneity and limited sample size. However, the DL model did not achieve performance beyond chance level (50% accuracy), indicating that, without a sufficiently large training dataset,

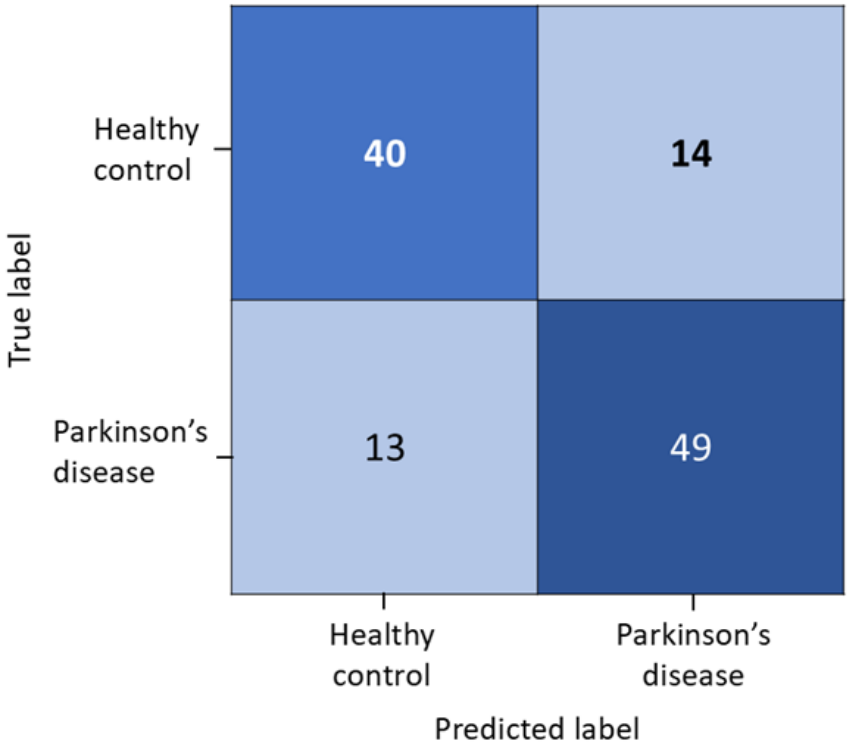


Figure 9. Confusion matrix for the selected primary random forest classifier model

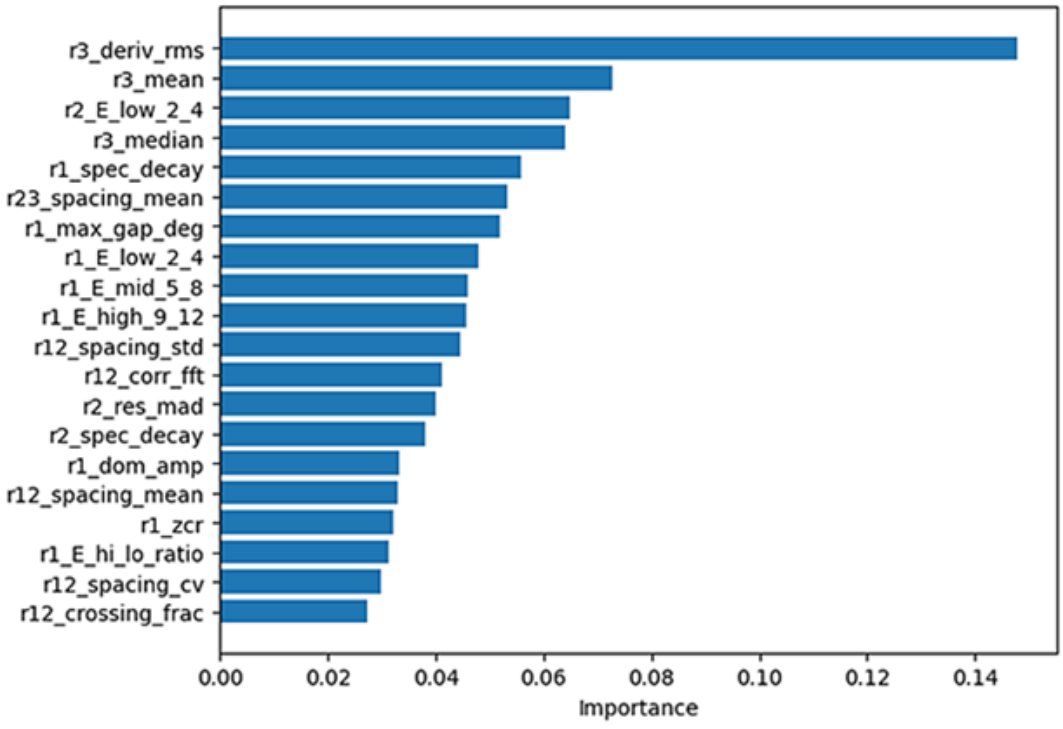


Figure 10. Feature importance for the selected random forest classifier model

Table 5. Stability of the top five discriminative features

Feature name	Mean importance	SD	Stability rating (CV)
r3_deriv_rms	0.0496	0.0136	Acceptable (27.4%)
r3_dom_amp	0.0238	0.0047	High (19.7%)
r3_std	0.0220	0.0038	High (17.5%)
r1_spec_decay	0.0217	0.0042	High (19.5%)
r3_res_ptp	0.0194	0.0017	High (8.7%)

Abbreviations: CV: Coefficient of variation; SD: Standard deviation.

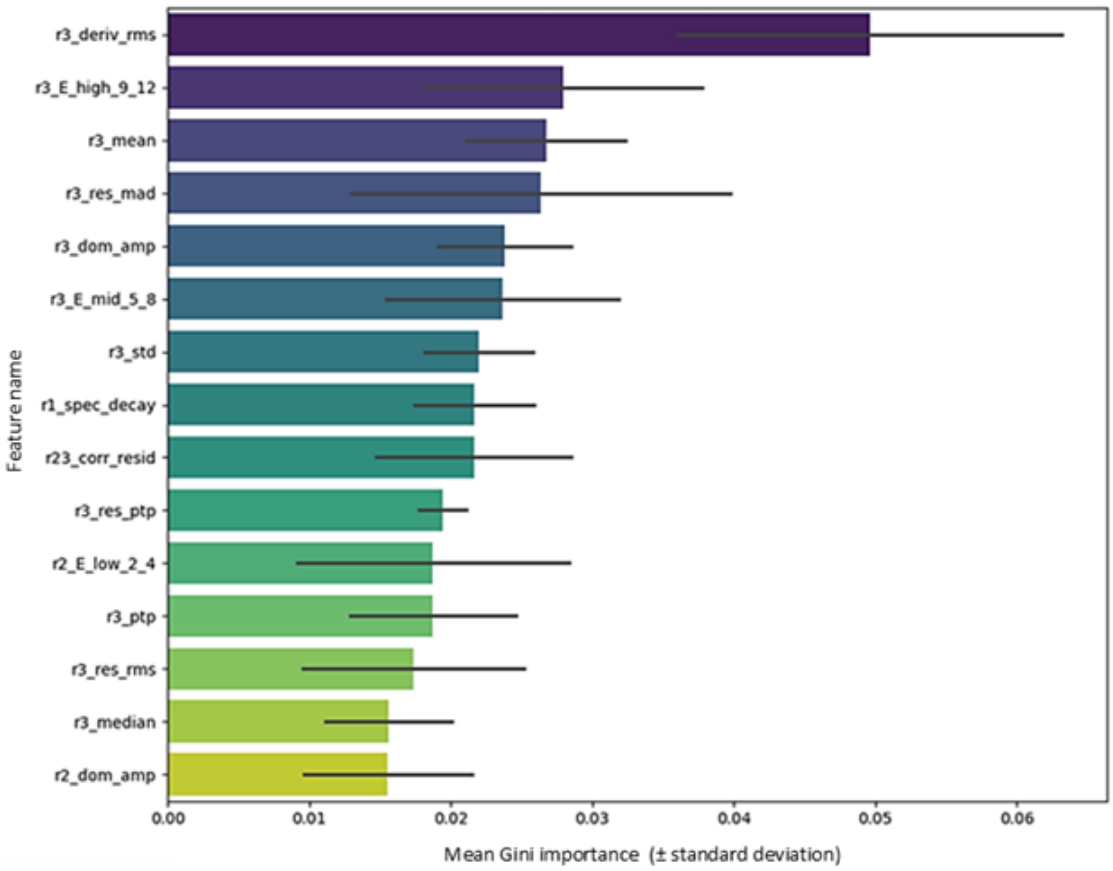


Figure 11. Analysis of feature importance stability

end-to-end architectures are not well suited for this data scale. In contrast, by utilizing the RF model alongside the targeted radial and spectral biomarkers, the approach effectively mitigated environmental noise that affected the CNN, enabling a highly discriminative and generalizable diagnostic performance.

6.3. Limitations and future work

While the proposed framework demonstrates high diagnostic potential, several limitations remain. First, although the total sample size ($n = 116$) and the integration of diverse public data along with locally collected data at a centralized location enhance robustness, the cohort lacks detailed longitudinal characterization. Future iterations should involve larger multi-center cohorts to validate these findings across broader demographics.

Second, this study was designed as a binary classification between patients with PD and healthy controls. While this establishes a proof of concept, the current model does not yet provide a differential diagnosis against "look-alike" conditions such as essential tremor or dystonia, which also

manifest as spiral drawing abnormalities. To maintain high specificity in this pilot stage, our clinical team performed a strict differential diagnosis; however, the absence of a disease control group remains a limitation that we intend to address in future work.

Third, while our cohort consisted exclusively of treatment-naïve patients to ensure a baseline of motor symptoms unaffected by pharmacological intervention, we acknowledge that the lack of MDS-UPDRS scores and other clinical scales limits the depth of the analysis. At this stage, we have not provided a correlation analysis between the extracted computational features and established clinical severity scales. Bridging this gap is a priority for future research to validate these digital features as clinically meaningful biomarkers.

Finally, the current study suggests that the system is best suited as a clinical decision-support tool rather than a standalone diagnostic platform, particularly for screening in low-resource environments. Future research should focus on expanding the dataset to include disease controls and longitudinal tracking to evaluate the model's sensitivity to disease progression and treatment efficacy.

7. Conclusion

This study demonstrates that the spatial and frequency-domain analysis of spiral trajectories constitutes a robust, lightweight, and computationally efficient framework for the automated detection of PD. By characterizing spiral morphology through radial fluctuations, kinematic smoothness, and turn-to-turn consistency, the proposed method achieves high diagnostic accuracy with clinically relevant biomarkers without relying on the extreme data dependency or "black-box" nature of traditional DL architectures. Our findings indicate that features derived from the outer revolution's radial derivative ($r3_deriv_rms$), spacing variability, and high-frequency spectral energy are the most discriminative factors in distinguishing PD subjects from healthy controls. Specifically, the analysis of irregular revolution overlaps and the spectral balance between high- and low-frequency content provides strong diagnostic performance.

The inherent transparency of this feature-driven approach ensures that the model's outputs are directly aligned with the neurological understanding of motor dysfunction, making it an ideal candidate for integration into clinical decision-support systems where explainability is paramount. To identify the most effective classifier, we evaluated several distinct ML architectures, including linear (logistic regression), kernel-based (support vector machine), and ensemble methods (e.g., gradient boosting, XGBoost, and LightGBM). RF emerged as the most robust model for this small-sample regime. Specifically, our framework with RF achieves robust performance (AUC = 0.832) by training on a heterogeneous dataset that integrates expert-verified clinical samples from across the Central Province of Sri Lanka with global public data. This approach ensures feature invariance across different hardware and demographic variations, effectively mitigating the risks of cohort-specific overfitting and reinforcing the model's potential for real-world clinical deployment.

While the current results are promising, several avenues exist for extending this research. Longitudinal monitoring of patients could track disease progression and assess treatment response over time. Furthermore, we aim to enhance performance by fusing frequency-domain attributes with high-resolution kinematic data, such as pen velocity, acceleration, and axial pressure. Finally, conducting broader clinical trials across more diverse, large-scale datasets will be essential to ensure the generalizability and reliability required for real-world healthcare deployment.

Acknowledgments

None.

Funding

None.

Conflict of interest

The authors declare they have no competing interests.

Author contributions

Conceptualization: Maheshi Dissanayake, Sajitha Weerasinghe
Formal analysis: All authors
Investigation: All authors
Methodology: All authors
Software: Tharaka Wijethunge
Supervision: Maheshi Dissanayake, Sajitha Weerasinghe
Validation: Sajitha Weerasinghe
Writing–original draft: Tharaka Wijethunge
Writing–review & editing: All authors

Ethics approval and consent to participate

Ethical clearance was obtained (approval no.: THP/PLANNING/ECR/19/2024; date: September 2, 2025) and granted by the Ethical Review Committee of the Teaching Hospital Peradeniya (University Hospital of the University of Peradeniya), Sri Lanka.

Consent for publication

Not applicable, as no direct clinical data are disclosed in this publication.

Availability of data

The sprial image data that support the findings of this study are available on request from the corresponding author. Custom codes and scripts supporting this reserach are publicly available at https://github.com/tharaka-w/PD/blob/main/Full_Framework_new.ipynb.

Further disclosure

Gemini (Google) and Grammarly were used during manuscript drafting to improve sentence structure, readability, and language. The content was subsequently reviewed and edited by the authors, and the authors are fully responsible for the content of the publication.

References

1. Mercaldo F, Brunese L, Cesarelli M, Martinelli F, Santone A. Spiral Drawing Test and Explainable Convolutional Neural Networks for Parkinson's Disease Detection. In: Proceedings of the 16th International Conference on Agents and Artificial Intelligence *(ICAART 2024)*. 2024:443-452.

   doi: 10.5220/0012407100003636

2. Smits EJ, Tolonen AJ, Cluitmans L, *et al.* Standardized handwriting to assess bradykinesia, micrographia and tremor in Parkinson's disease. *PLoS ONE.* 2014;9(5):e97614.

   doi: 10.1371/journal.pone.0097614

3. Wang H, Wang L, Pullman SL, Louis ED. Spiral analysis—Improved clinical utility with center detection. *J Neurosci Methods.* 2008;171(2):264-270.

   doi: 10.1016/j.jneumeth.2008.03.009

4. Liu X, Carroll CB, Wang SY, Zajicek J, Bain PG. Quantifying drug-induced dyskinesias in the arms using digitised spiral-drawing tasks. *J Neurosci Methods.* 2005;144(1):47-52.

   doi: 10.1016/j.jneumeth.2004.10.005

5. Letanneux A, Danna J, Velay JL, Viallet F, Pinto S. From micrographia to Parkinson's disease dysgraphia. *Mov Disord.* 2014;29(12):1467-1475.

   doi: 10.1002/mds.25990

6. Thomas M, Lenka A, Pal PK. Handwriting analysis in Parkinson's disease: Current status and future directions. *Mov Disord Clin Pract.* 2017;4(6):806-818.

   doi: 10.1002/mdc3.12552

7. Ma HI, Hwang WJ, Chang SH, Wang TY. Progressive micrographia shown in horizontal, but not vertical, writing in Parkinson's disease. *Behav Neurol.* 2013;27(2):169-174.

   doi: 10.1155/2013/212675

8. Zham P, Arjunan SP, Raghav S, Kumar DK. A kinematic study of progressive micrographia in Parkinson's disease. *Front Neurol.* 2019;10:403.

   doi: 10.3389/fneur.2019.00403

9. San Luciano M, Wang C, Ortega RA, *et al.* Digitized spiral drawing: A possible biomarker for early Parkinson's disease. *PLoS ONE.* 2016; 11(10): e0162799.

   doi: 10.1371/journal.pone.0162799

10. Pullman SL. Spiral analysis: A new technique for measuring tremor with a digitizing tablet. *Mov Disord.* 1998;13(s3):85-89.

    doi: 10.1002/mds.870131315

11. Rizzo G, Copetti M, Arcuti S, Martino D, Fontana A, Logroscino G. Accuracy of clinical diagnosis of Parkinson disease: A systematic review and meta-analysis. *Neurology.* 2016;86(6):566-576.

    doi: 10.1212/WNL.0000000000002350

12. Westin J, Ghiamati S, Memedi M, *et al.* A new computer method for assessing drawing impairment in Parkinson's disease. *J Neurosci Methods.* 2010;190(1):143-148.

    doi: 10.1016/j.jneumeth.2010.04.027

13. Elble RJ, Pullman S, Matsumoto JY, Raethjen J, Deuschl G, Tintner R. Digitizing tablet and Fahn-Tolosa-Marín ratings of Archimedes spirals have comparable minimum detectable change in essential tremor. *Tremor Other Hyperkinet Mov.* 2017.

    doi: 10.5334/tohm.344

14. Kamble M, Shrivastava P, Jain M. Digitized spiral drawing classification for Parkinson's disease diagnosis. *Meas : Sens.* 2021;16:100047.

    doi: 10.1016/j.measen.2021.100047

15. Rosenblum S, Samuel M, Zlotnik S, Erikh I, Schlesinger I. Handwriting as an objective tool for Parkinson's disease diagnosis. *J Neurol.* 2013;260(9):2357-2361.

    doi: 10.1007/s00415-013-6996-x

16. Huang P, He P, Tian S, *et al.* A ViT-AMC Network With Adaptive Model Fusion and Multiobjective Optimization for Interpretable Laryngeal Tumor Grading From Histopathological Images. *IEEE Trans Med Imaging.* 2023;42(1):15-28.

    doi: 10.1109/TMI.2022.3202248

17. Li W, Rao Q, Dong S, *et al.* PIDGN: An explainable multimodal deep learning framework for early prediction of Parkinson's disease. *J Neurosci Methods.* 2025;415:110363.

    doi: 10.1016/j.jneumeth.2025.110363

18. Tonekaboni S, Joshi S, McCradden MD, Goldenberg A. What Clinicians Want: Contextualizing Explainable Machine

Learning for Clinical End Use. In: *Proceedings of Machine Learning Research.* Vol 106. 2019:359-380. Available from: https://proceedings.mlr.press/v106/tonekaboni19a.html [Last accessed on April 25, 2026].

19. Wang S, Schwirtlich L, Pullman SL, *et al.* Clinical applications and measurement properties of the digitized Archimedes spiral drawing test: A scoping review. *Mov Disord Clin Pract.* 2025;12(11):1742-1755.

    doi: 10.1002/mdc3.70278

20. Peng Y, Han S, Wu D, *et al.* A deep learning approach to remotely assessing essential tremor with handwritten images. *Sci Rep.* 2025;15(1):10783.

    doi: 10.1038/s41598-025-94729-0

21. Yoon H, Ahn M. Quantification of Movement Error from Spiral Drawing Test. *Sensors.* 2023;23(6):3043.

    doi: 10.3390/s23063043

22. Ishii N, Mochizuki Y, Shiomi K, Nakazato M, Mochizuki H. Spiral drawing: Quantitative analysis and artificial-intelligence-based diagnosis using a smartphone. *J Neurol Sci.* 2020;411:116723.

    doi: 10.1016/j.jns.2020.116723

23. Louis ED. The roles of age and aging in essential tremor: An epidemiological perspective. *Neuroepidemiology.* 2019;52(1-2):111-118.

    doi: 10.1159/000492831

24. Danna J, Velay JL, de Gabory I, *et al.* Digitalized spiral drawing in Parkinson's disease: A tool for evaluating beyond the written trace. *Hum Mov Sci.* 2019;65:80-88.

    doi: 10.1016/j.humov.2018.08.003

25. Aghanavesi S, Memedi M, Westin J. Verification of a method for measuring Parkinson's disease related temporal irregularity in spiral drawings. *Sensors.* 2017;17(10):2341.

    doi: 10.3390/s17102341

26. Smits EJ, Tolonen AJ, de Vos CC, *et al.* Graphical tasks to measure upper limb function in patients with Parkinson's disease: Validity and response to dopaminergic medication. *IEEE J Biomed Health Inform.* 2017;21(1):283-289.

    doi: 10.1109/JBHI.2015.2503802

27. Drotár P, Mekyska J, Rektorová I, Masarová L, Smékal Z, Faundez-Zanuy M. Evaluation of handwriting kinematics and pressure for differential diagnosis of Parkinson's disease. *Artif Intell Med.* 2016;67:39-46.

    doi: 10.1016/j.artmed.2016.01.004

28. Drotár P, Mekyska J, Rektorová I, Masarová L, Smékal Z, Faundez-Zanuy M. Decision support framework for Parkinson's disease based on novel handwriting markers. *IEEE Trans Neural Syst Rehabil Eng.* 2015;23(3):508-516.

    doi: 10.1109/TNSRE.2014.2359997

29. Rios-Urrego CD, Vásquez-Correa JC, Vargas-Bonilla JF, Nöth E, Lopera F, Orozco-Arroyave JR. Analysis and evaluation of handwriting in patients with Parkinson's disease using kinematic, geometrical, and non-linear features. *Comput Methods Programs Biomed.* 2019;173:43-52.

    doi: 10.1016/j.cmpb.2019.03.005

30. Saunders-Pullman R, Derby C, Stanley K, *et al.* Validity of spiral analysis in early Parkinson's disease. *Mov Disord.* 2008;23(4):531-537.

    doi: 10.1002/mds.21874

31. Toffoli S, Gazzola G, Bonvento B, *et al.* Spiral drawing analysis with a smart ink pen to identify Parkinson's disease fine motor deficits. *Front Neurol.* 2023;14.

    doi: 10.3389/fneur.2023.1093690

32. Sisti JA, Christophe B, Seville AR, *et al.* Computerized spiral analysis using the iPad. *J Neurosci Methods.* 2017;275:50-54.

    doi: 10.1016/j.jneumeth.2016.11.004

33. Zham P, Arjunan SP, Raghav S, Kumar DK. Efficacy of guided spiral drawing in the classification of Parkinson's disease. *IEEE J Biomed Health Inform.* 2018;22(5):1648-1652.

    doi: 10.1109/JBHI.2017.2762008

34. Gil-Martín M, Montero JM, San-Segundo R. Parkinson's disease detection from drawing movements using convolutional neural networks. *Electronics.* 2019;8(8):907.

    doi: 10.3390/electronics8080907

35. Senatore R, Marcelli A, De Micco R, Tessitore A, Teulings HL. Distinctive handwriting signs in early Parkinson's disease. *Appl Sci.* 2022;12(23):12338.

    doi: 10.3390/app122312338

36. Purk M, *et al.* Utilizing a tablet-based artificial intelligence system to assess movement disorders in a prospective study. *Sci Rep.* 2023;13.

    doi: 10.1038/s41598-023-37388-3

37. Cascarano GD, Loconsole C, Brunetti A, *et al.* Biometric handwriting analysis to support Parkinson's disease assessment and grading. *BMC Med Inform Decis Mak.* 2019;19(s9):252.

    doi: 10.1186/s12911-019-0989-3

38. Ullah Z, Kim J. Hierarchical Deep Feature Fusion and Ensemble Learning for Enhanced Brain Tumor MRI Classification. *Mathematics.* 2025;13(17):2787.

    doi: 10.3390/math13172787

39. Ullah Z, Hong M, Mahmood T, Kim J. Systematic Integration of Attention Modules into CNNs for Accurate and Generalizable Medical Image Classification. *Mathematics.* 2025;13(22):3728.

doi: 10.3390/math13223728

40. Kumar BA, Bansal M. A transfer learning approach with MobileNetV2 for Parkinson's Disease Detection using hand-drawings. In: Proceedings of the 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT). IEEE; 2023:1-7.

doi: 10.1109/ICCCNT56998.2023.10307641

41. Zham P, Kumar DK, Dabnichki P, Poosapadi Arjunan S, Raghav S. Distinguishing Different Stages of Parkinson's Disease Using Composite Index of Speed and Pen-Pressure of Sketching a Spiral. *Front Neurol*. 2017;8.

doi: 10.3389/fneur.2017.00435

42. Haubenberger D, Kalowitz D, Nahab FB, *et al.* Validation of digital spiral analysis as outcome parameter for clinical trials in essential tremor. *Mov Disord.* 2011;26(11):2073-2080.

doi: 10.1002/mds.23808