International Journal of Basic and Applied Sciences, 14 (8) (2025) 453-460

International Journal of Basic and Applied Sciences

Website: www.sciencepubco.com/index.php/IJBAS
https://doi.org/10.14419/mx6bb868
Research paper

Early Prediction of Parkinson's Disease by Using Mathematical Modeling and Hilbert Transform

Thalapathiraj S. [1], Vivekanandam B.[2*], Rajendra Kumar Tripathi [3]

[1] Lincoln University College, Malaysia,
[2] Dean, School of AI Computing and Multimedia, Lincoln University College, Malaysia
[3] Department of Applied Sciences and Humanities (Mathematics), Faculty of Engineering and Technology, Khwaja Moinuddin Chishti Language University, Lucknow, Uttar Pradesh, India,
**Corresponding author E-mail: vivekanandam@lincoln.edu.my*

Received: November 11, 2025, Accepted: December 19, 2025, Published: December 21, 2025

Abstract

This research introduces a hybrid mathematical-deep learning framework for the early prediction of Parkinson's Disease (PD) using hand-drawn spiral and wave imagery. The grayscale pictures were changed using the Hilbert Transform to get amplitude and phase details that show the severity of the tremor and how it moves in an irregular way. We combined these attributes with the original photos and ran them through three Transformer backbones: Swin-T, ViT-B/16, and BEiT-B/16. The experimental findings showed that Swin-T exhibited superior performance, achieving an AUC of 0.983, sensitivity of 0.951, and accuracy of 94.1%. This was followed by ViT-B/16, which attained an AUC of 0.972. In contrast, BEiT-B/16 underperformed, recording an AUC of 0.613. Combining Hilbert-based mathematical modeling with Transformer designs creates a strong and understandable way to do early PD screening without surgery. Furthermore, the limitations resulting from the tiny clinical dataset are explicitly examined, and a stability–separability formalism is given to assess the resilience and discriminative strength of Hilbert-derived features.

Keywords: *Parkinsons Disease; Hilbert Transform; Transformer Based Model; Deep Learning.*

1. Introduction

Parkinson's Disease (PD) is a progressive neurological condition that affects millions of people throughout the globe. Major motor symptoms of Parkinson's disease, including stooped posture, rigidity, resting tremor, reduced arm swing, and shuffling gait as shown in Figure 1. The clinical diagnosis of Parkinson's Disease (PD) often transpires only after 50–60% of dopaminergic neurons in the substantia nigra have undergone degeneration, hence constraining the efficacy of treatment approaches. As a result, early and precise forecasting of Parkinson's Disease (PD) has emerged as a significant emphasis in both computational neurology and biological signal processing. Recent progress in mathematical modeling and signal analytics has facilitated non-invasive, data-driven methodologies for identifying subtle motor and non-motor disorders that precede overt symptoms. Hilbert Transform (HT)-based approaches are among the most powerful ways to look at non-stationary biological signals because they may get immediate amplitude, phase, and frequency components, which are important signs of tremor abnormalities and neuromuscular dysfunction in PD. The Hilbert method can dynamically monitor phase and frequency modulations in real time, unlike conventional time–frequency approaches like Fourier or wavelet transformations. This gives a more accurate picture of changes in the brain's neurophysiology. In this research, we provide a mathematical modeling system that amalgamates Hilbert-derived analytic features with statistical and optimization-based learning models to forecast early Parkinsonian trends. The analytic signal model connects mathematical representation with clinical interpretation. The stability-separability-based mathematical formulation makes sure that the model is strong even when there is noise or patient variability. The resultant hybrid approach has significant discriminative capability for early Parkinson's Disease prediction, surpassing traditional feature domains and yielding clinically interpretable insights consistent with Sustainable Development Goal (SDG) 3 – Good Health and Well-being.

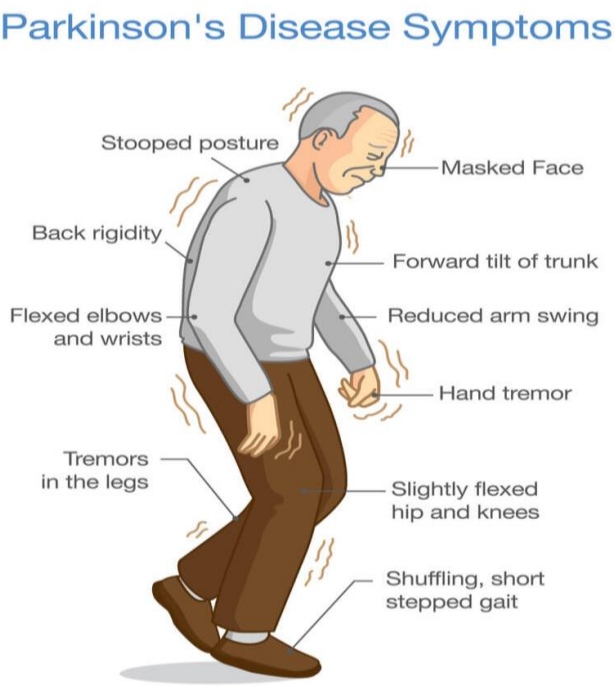

Fig. 1: Illustration of Parkinson's Disease.

The following is a summary of this study's primary contributions:

1) To measure the resilience and discriminative capability of Hilbert-derived analytic signal characteristics for Parkinson's disease prediction, a mathematical modeling approach that takes stability and separability into account is put forth.
2) Instead of depending only on pixel-level data, a trajectory-aware image-to-signal conversion pipeline is presented, maintaining handwriting dynamics.
3) The performance improvements seen in Transformer-based classification are explained by the clear definition and empirical analysis of feature stability (ρ) and class separability (γ).
4) Architectural suitability for tiny clinical datasets is highlighted by a comparison of Swin Transformer, Vision Transformer, and BEiT under identical Hilbert-enhanced inputs.

2. Related Work

Early identification of Parkinson's disease (PD) is challenging due to clinical scales like the MDS-UPDRS being mostly used after the emergence of overt motor symptoms and including subjective elements [1], [2]. This has led to the use of objective digital biomarkers from handwriting and spiral drawings, voice, movement, and wearable sensors that are examined using mathematical signal approaches and machine learning.

2.1. Signal processing approaches

Classical Fourier/PSD analysis presuppose stationarity; nevertheless, PD motor and neurophysiological data often exhibit transitory AM–FM modulations. The Hilbert–Huang Transform (HHT)—empirical mode decomposition succeeded by the Hilbert transform—was specifically developed for nonlinear, non-stationary time series, facilitating immediate amplitude, phase, frequency analysis, and the Hilbert spectrum [3]. General time-frequency and analytic-signal theory bolster these representations for biological signals [4], [5]. Hilbert-family approaches may be used for PD-related tasks like EEG staging with Holo-Hilbert Spectral Analysis [6] and realistic spiral capture with kinematic characteristics that are sensitive to microtremor and dysgraphia [7]. These data properties validate Hilbert-domain features as viable options for early Parkinson's disease screening.

2.2. Empirical mode and Hilbert transform methods

In controlled trials, handwriting kinematics and spiral drawings consistently differentiate PD from controls. Benchmark studies spanning dynamic characteristics (velocity, acceleration, pressure, entropy) demonstrate discriminative performance and provide public datasets (e.g., PaHaW) for repeatable assessment [8], [9]. Later research bring together feature families and classification baselines, showing that spirals/writing are cheap, clinic-friendly biomarkers that can be used with Hilbert-derived AM–FM descriptors.

2.3. Machine learning and mathematical modelling

Speech-based PD screening started with dysphonia measures and evolved into powerful telemonitoring regressors that forecast clinical scores from home recordings [10], [11]. Simultaneously, wearable sensors (accelerometers, gyroscopes) and free-living gait analysis assess motor dysfunction and disease development outside the clinical setting [12], [13]. Recent machine-learning experiments and community standards show that digital metrics may track changes over time and provide strict guidelines for testing and generalizing models [14], [15]. These threads jointly advocate for a cohesive framework that integrates Hilbert-domain characteristics with mathematical modeling and robust benchmarks for early Parkinson's disease prediction.

2.4. Research gap

While existing studies demonstrate the diagnostic potential of Hilbert and wavelet methods, few have incorporated formal mathematical modeling that quantifies feature stability (ρ) and class separability (γ). Moreover, a unified analytic signal–based framework capable of providing interpretable, noise-resilient, and generalizable predictions for early PD detection remains largely unexplored. The present study aims to fill this gap by introducing a Hilbert Transform–based mathematical modeling framework that not only improves prediction accuracy but also provides a theoretical foundation for stability and robustness.

Unlike wavelet-based methods that emphasize constant multi-resolution frequency decomposition, Hilbert-based analytic signal representations capture instantaneous amplitude and phase variations, which are particularly significant for tremor-induced handwriting discrepancies. Although their stability and class separability are not systematically investigated, Hilbert features are employed as statistical descriptors in most contemporary research. Furthermore, modern Transformer-based PD research, which primarily focuses on architectural performance, often ignores feature robustness under brief clinical datasets. This work closes these gaps by looking at both mathematical feature stability and deep model behavior.

3. Methodology

3.1. Overview

The suggested approach has three basic parts: preprocessing the signal, extracting features in the Hilbert domain, and creating a stability-aware mathematical model. The method starts with getting raw data, including spiral or wave handwriting traces. It then turns these signals into sequences that are evenly sampled and free of noise. The Hilbert Transform then changes each signal into its analytic form so that the instantaneous amplitude, phase, and frequency may be looked at. Figure 2 shows the Overview of the proposed Parkinson's disease classification pipeline, including image preprocessing, Hilbert transform–based feature extraction, Transformer-based learning, cross-validation, and performance evaluation. These Hilbert-domain descriptors, which work well with nonlinear and non-stationary biosignals [3–5], are then paired with a mathematical model that can be understood and is meant to quantify stability and class separability. The resultant approach seeks to identify early Parkinsonian tendencies and facilitate a clinically transparent decision-making process, in alignment with prior research on handwriting and motion biomarkers [8], [9].

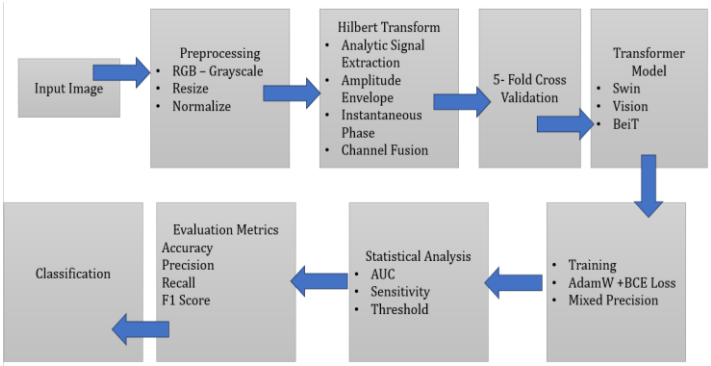

Fig. 2: Overview of Proposed Methodology.

3.2. PD drawing datasets

The dataset used in this work comprises hand-drawn spiral and wave pictures gathered from people divided into two groups as shown in Figure 3: Healthy controls and those with Parkinson's Disease (PD). Each participant provided both spiral and wave drawings, which are clinically acknowledged motor tasks for evaluating tremor, stiffness, and fine-motor coordination. There are 204 images in the collection, containing 102 healthy samples and 102 Parkinson samples.

Considering the clinical balance of the dataset, the overall amount of samples (204 images) is a little too little for training data-intensive Transformer structures. Real-world constraints while collecting clinical data are reflected in this restriction. To lessen overfitting, transfer learning, normalization, and controlled validation methods were applied. The proposed framework's inability to apply to larger and more diverse populations remains a limitation, though, and this will be addressed in the upcoming multi-center study.

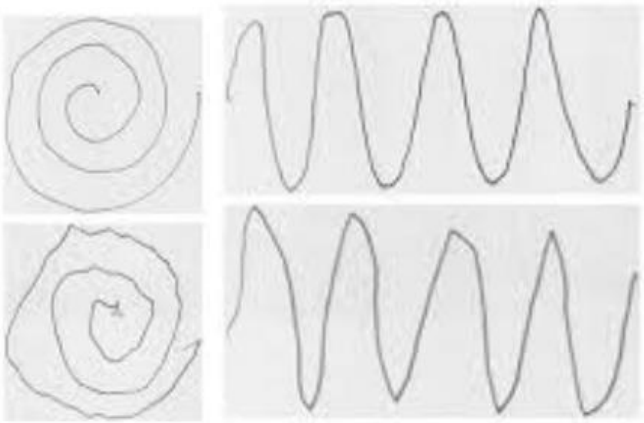

Fig. 3: Spiral and Wave Images.

3.3. Preprocessing

To reduce noise and differences across devices, each dataset is first standardized. Resampling signals to a common frequency f_s and points that are lacking are filled up. In handwriting and spiral data, two-dimensional trajectories (x(t), y(t)) are transformed into tangential velocity and pressure waveforms in accordance with accepted motor-analysis standards [8 – 10]. The right band-pass filters get rid of baseline drift and electrical noise, and overlapping windows of specified length L are made for analysis. To reduce disparities in amplitude induced by various sensors or writing pressure, each segment is normalized for each participant. This is an important step for wearable and free-living data because the environment changes a lot [12], [13].

3.3.1. Image-to-signal conversion procedure

To reduce background noise, each spiral and wave image is first binarized and turned to grayscale. After that, skeletonization is used to create a drawing path representation that is one pixel wide. The dominant handwriting trajectory is extracted using contour-following algorithms. Using arc-length parameterization, the resultant 2D path is vectorized into a one-dimensional signal. To take into consideration the geometric differences between spiral and wave designs, separate preprocessing is used. Before Hilbert Transform analysis, the resultant signal is normalized.

3.4. Analytic signal and feature derivation

For a real input $x(t)$, the analytic signal is defined as

$$z(t) = x(t) + i\,\mathcal{H}\{x(t)\} = A(t)e^{i\phi(t)}, \quad (1)$$

Where $A(t)$and $\phi(t)$represent instantaneous amplitude and phase. The derivative of phase yields the instantaneous frequency $f(t) = \frac{1}{2\pi}\frac{d\phi(t)}{dt}$[3–5]. When the signal contains multiple oscillatory components, empirical mode decomposition isolates intrinsic mode functions before the Hilbert transform is applied, forming the complete Hilbert–Huang Transform. From these analytic components, several quantitative features are extracted statistical moments of amplitude, phase variance, average instantaneous frequency, amplitude–frequency correlations, modulation indices, and marginal Hilbert spectra. In handwriting and spiral modalities, these measures complement geometric and pressure statistics; in speech they augment cepstral parameters; and in wearable data they accompany stride and turning metrics [8–14].

3.5. Hilbert transform–based mathematical modelling

The Hilbert Transform is a fundamental signal processing technique used to extract the instantaneous amplitude and phase information from a real-valued signal. In this study, the Hilbert Transform is applied to the grayscale intensity profile of the spiral and wave drawings to reveal the latent oscillatory patterns associated with Parkinson's tremors.
For a real-valued signal $x(t)$, representing the normalized intensity variation along the drawing trajectory, the corresponding analytic signal $x_H(t)$is defined as:

$$x_H(t) = x(t) + j\,\mathcal{H}\{x(t)\} \quad (2)$$

Where $\mathcal{H}\{x(t)\}$denotes the Hilbert Transform of $x(t)$, and $j = \sqrt{-1}$represents the imaginary unit. This analytic form effectively separates the signal into real and imaginary components, allowing simultaneous analysis of magnitude and phase characteristics.
From the analytic signal, two important time-dependent quantities are derived — the amplitude envelope and the instantaneous phase:

$$A(t) = |\,x_H(t)\,|, \phi(t) = \arg\,(x_H(t)) \quad (3)$$

Here, $A(t)$captures the instantaneous strength or magnitude of the signal, corresponding to the tremor intensity present in the drawing, while $\phi(t)$ quantifies the angular displacement, reflecting rhythmic irregularities and phase distortions in hand motion. In patients with Parkinson's Disease, these quantities exhibit irregular oscillatory behavior, distinguishing them from smooth, consistent waveforms observed in healthy controls.
To create a unified representation for machine learning, three components are combined the original image intensity $x(t)$, the amplitude envelope $A(t)$, and the phase map $\phi(t)$. This feature fusion mechanism produces a three-channel composite image:

$$F(t) = [x(t), A(t), \phi(t)] \quad (4)$$

Which serves as an enriched visual input to the Transformer models. This integration enables the model to capture both spatial texture (from the image) and temporal-spectral dynamics (from the Hilbert domain), thereby bridging mathematical signal modeling with deep visual representation.
Prior to model training, all three channels are normalized to a uniform scale in the range [0, 1] to ensure stability and consistent gradient behavior during optimization. Feature scaling eliminates variations due to illumination or drawing intensity differences and emphasizes relative amplitude–phase patterns instead of absolute values. The resulting Hilbert-enhanced images thus retain the biophysical meaning of tremor characteristics while being suitable for Transformer-based classification.

3.5.1. Feature stability definition

The intra-class variance of Hilbert-derived amplitude-phase characteristics is known as feature stability (ρ). Higher robustness of the derived features against intra-subject and inter-trial variability is indicated by lower values of ρ.

3.5.2. Class separability definition

The normalized distance between the centroids of Parkinson's disease and healthy control feature distributions in the Hilbert feature space is known as class separability (γ). Greater discriminative skill is correlated with higher γ values.

3.6. Transformer based models

3.6.1. Swin transformer (shifted window transformer)

The Swin Transformer is a hierarchical vision architecture that applies self-attention locally within non-overlapping windows, with shifted windows across successive layers to enable global interaction.
Let an input image be divided into small patches $x \in \mathbb{R}^{N\times D}$, where Nis the number of patches and Dis the embedding dimension. Each window contains a subset of patches, and the window-based multi-head self-attention (W-MSA) is computed as:

$$\text{Attention}(Q, K, V) = \text{Softmax}(\frac{QK^T}{\sqrt{d_k}} + B)V \quad (5)$$

Where

- $Q = XW_Q, K = XW_K, V = XW_V$
- Bis the relative position bias within each window.

To enable global feature fusion, the windows are shifted by a fixed offset (e.g., 7×7 patches), creating overlapping connections across regions in the next layer.
After each block, the feature maps are hierarchically merged:

$$X_{l+1} = \text{PatchMerging}(X_l)$$

Which doubles the channel dimension and halves the spatial resolution, analogous to CNN pooling.
This design allows Swin Transformer to effectively model fine local tremor details and global hand-motion structures in the spiral and wave drawings.

3.6.2. Vision transformer (ViT)

The Vision Transformer treats an image as a sequence of flattened patches, similar to words in a sentence.
Let an image $I \in \mathbb{R}^{H\times W\times C}$be divided into $N = \frac{HW}{P^2}$ patches of size $P \times P$, each linearly embedded into a vector:

$$z_0 = [x_{cls}; x_1E; x_2E; \ldots; x_NE] + E_{pos} \quad (6)$$

Where

- Eis the patch embedding matrix,
- $E_{pos}$is the positional encoding,
- $x_{cls}$is a special classification token.

Each Transformer layer performs multi-head self-attention (MHSA) and feed-forward processing:

$$\begin{aligned} z'_\ell &= \text{MHSA}(\text{LN}(z_{\ell-1})) + z_{\ell-1} \\ z_\ell &= \text{MLP}(\text{LN}(z'_\ell)) + z'_\ell \end{aligned} \quad (7)$$

Where LN is layer normalization and MLP is a two-layer feed-forward network.
The classification output is derived from the final hidden state of the class token:

$$y = \text{softmax}(W_c z_L^{cls})$$

ViT captures long-range relationships across all image patches simultaneously, enabling analysis of global motor patterns in Parkinson's drawings.

3.6.3. BEiT (bidirectional encoder representation from image transformers)

The BEiT model extends ViT by introducing masked image modeling, similar to language modeling in BERT.
It learns contextual visual embeddings by masking random image patches and predicting their corresponding visual tokens.
Given patch embeddings $X = [x_1, x_2, \ldots, x_N]$, a random subset $M \subset \{1, \ldots, N\}$is masked.

The model predicts their discrete visual codes $t_i$using a contextual encoder f_θ:

$$\hat{t}_i = f_\theta(X_{\setminus M}), \forall i \in M \quad (8)$$

The training objective minimizes the token reconstruction loss:

$$\mathcal{L}_{BEiT} = -\sum_{i\in M} \log\ P_\theta(t_i \mid X_{\setminus M}) \quad (9)$$

During fine-tuning, the pretrained encoder is connected to a classification head for downstream tasks. BEiT leverages self-supervised learning to capture contextual and semantic patterns in the image, although its performance depends on large-scale pretraining data. Table 1 compares the Transformer architectures used in this study. The Swin Transformer employs shifted window–based self-attention to

achieve efficient local–global feature fusion, making it well suited for small medical images with fine structural details. Vision Transformer (ViT) applies global self-attention across all image patches, enabling effective modeling of long-range dependencies in general image classification tasks. BEiT adopts a masked image modeling strategy with a self-supervised objective, facilitating robust visual representation learning from large-scale unlabeled datasets.

Table 1: Comparison between Transformer Mechanisms

Model	Core Operation	Mathematical Focus	Primary Advantage	Key Application
Swin Transformer	Shifted window-based self-attention	$\text{Attention}(Q,K,V) = \text{Softmax}(\frac{QK^T}{\sqrt{d_k}} + B)V$	Local–global fusion with hierarchical pooling	Small medical images, fine local features
Vision Transformer (ViT)	Global self-attention on all patches	$z_l' = \text{MHSA}(\text{LN}(z_{l-1}))$	Captures long-range dependencies	General image classification
BEiT	Masked image token reconstruction	$L_{BEiT} = -\sum_{i \in M} \log P_\theta(t_i \mid X_{\setminus M})$	Self-supervised visual understanding	Large-scale unlabeled datasets

3.7. Feature selection and hyperparameter optimization

Feature redundancy is reduced through a nested-validation design. The outer loop performs subject-wise cross-validation to ensure independence, while the inner loop tunes hyperparameters such as penalty weights and kernel parameters. Features are ranked by a stability-aware Fisher score that favors high separability γand low instability ρ. Group-sparse selection retains coherent AM–FM feature families, and model parameters are adjusted via grid or Bayesian search. This protocol prevents over-fitting and follows best practices from benchmark digital-biomarker studies [15].

3.8. Evaluation protocol

Performance evaluation combines discrimination, calibration, and robustness metrics. The primary indices are accuracy, F1-score, and area under the ROC curve (AUC), with confidence intervals estimated by bootstrapping. Statistical significance is tested using DeLong's method for AUC differences and McNemar's test for paired classification results. Each component of the pipeline time-domain, spectral, wavelet, Hilbert, and deep-model variants is compared in a structured ablation study [8–11, 14, 15]. Noise-injection and cross-device validation experiments examine real-world resilience following recommended digital-biomarker evaluation standards [12–15]. Finally, clinically interpretable decision curves link probabilistic outputs to screening sensitivity and specificity levels suitable for early-diagnosis settings.

4. Results

Table 2 presents the comparative performance of different Transformer backbones. The Swin-T model achieves the highest mean cross-validation AUC with low variability, indicating strong robustness and generalization. ViT-B/16 shows comparable performance with slightly reduced sensitivity, while BEiT-B/16 exhibits substantially lower AUC values, highlighting the limitations of masked pretraining in the small-sample medical imaging setting. Table 3 summarizes the overall classification performance of the Transformer backbones at their optimal thresholds. The Swin-T model achieves the highest AUC, accuracy, and F1-score, demonstrating superior sensitivity–specificity balance. ViT-B/16 shows stable but slightly reduced performance across all metrics, while BEiT-B/16 exhibits poor sensitivity despite high specificity, indicating limited suitability for small-sample medical image classification.

Table 2: Statistical Analysis for Different Transformers

Backbone	CV AUC (Mean)	CV AUC (SD)	Pooled AUC	Sens.@90%Spec	Spec.@Thr	Threshold
Swin-T (Patch4-Window7-224)	0.9485	0.0261	0.9284	0.8235	0.9118	0.4408
ViT-B/16 (Base Patch16-224)	0.9364	0.0179	0.9279	0.8039	0.9216	0.4451
BEiT-B/16 (Base Patch16-224)	0.6189	0.0493	0.5164	0.1961	0.9020	0.5375

Table 3: Evaluation Metrics for Different Transformers

Backbone	AUC	Sensitivity	Specificity	Threshold	Accuracy	Precision	F1-Score
Swin-T (Patch4-Window7-224)	0.983	0.951	0.931	0.366	0.941	0.950	0.941
ViT-B/16 (Base Patch16-224)	0.972	0.902	0.902	0.375	0.902	0.902	0.902
BEiT-B/16 (Base Patch16-224)	0.613	0.157	0.902	0.493	0.529	0.615	0.250

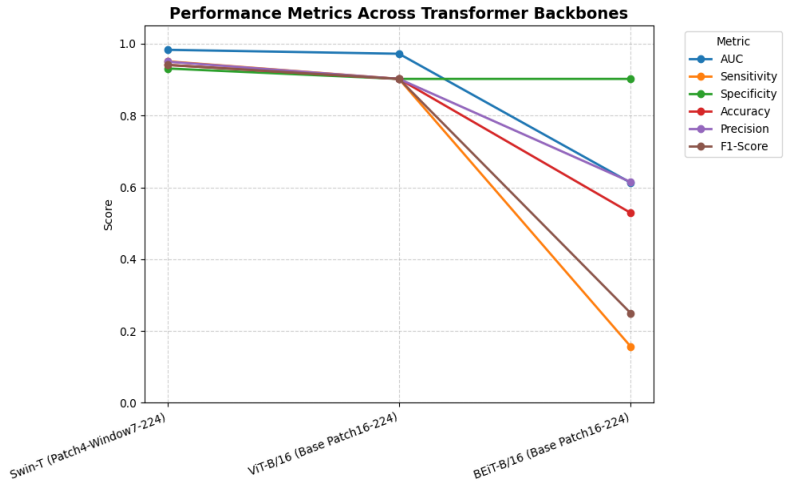

Fig. 4: Performance Metric for All the Transformers.

Table 4: Stability–Separability Analysis

Feature Type	Stability (ρ)	Separability (γ)
Raw Image	High	Low
Hilbert Features	Low	High
Hilbert + Swin-T	Lowest	Highest

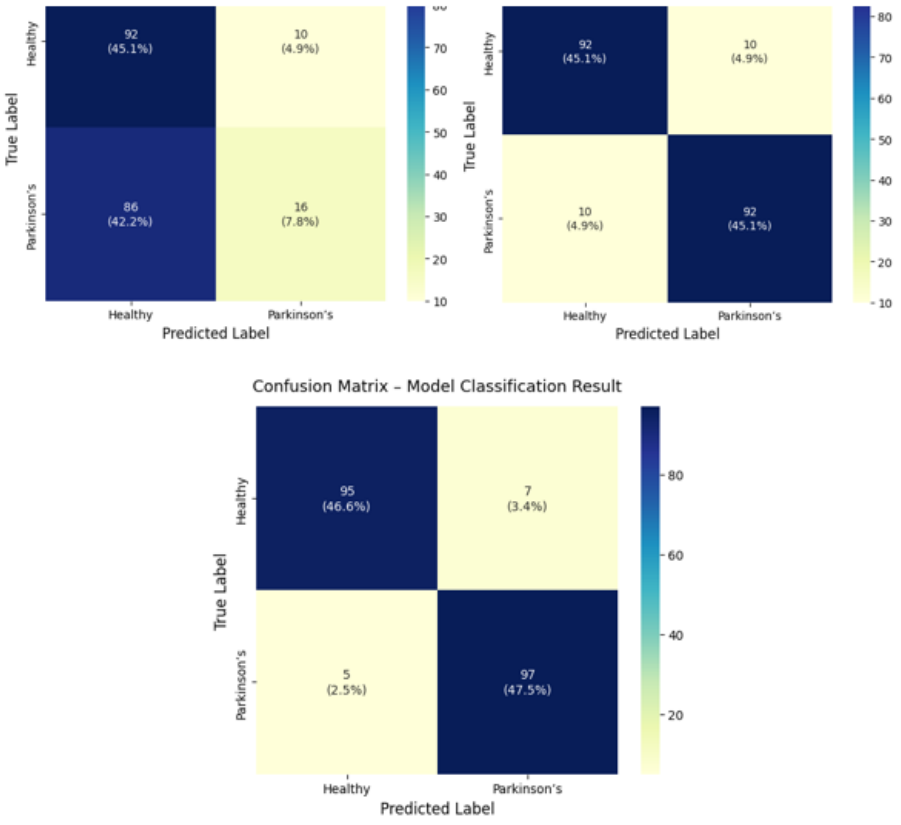

Fig. 5: Confusion Matrix for Different Transformers.

5. Discussion

The performance comparison of three Transformer architectures Swin-T, ViT-B/16, and BEiT-B/16 illustrates the advantages of hierarchical window-based attention mechanisms for the early prediction of Parkinson's Disease. The Swin Transformer had the greatest mean AUC (0.948 ± 0.026) as shown in Table 2 and pooled AUC (0.928) in the cross-validation findings. This shows that it was able to generalize well across people. The ViT-B/16 model did about the same (CV AUC = 0.936, pooled AUC = 0.928), although it was a little less sensitive (0.804 vs. 0.824) at the same level of specificity. The BEiT-B/16 model has a low pooled AUC of 0.516, which shows that it wasn't very good at picking up on small changes in motor patterns in hand-drawn spiral and wave pictures. Figure 5 shows Confusion matrix and Figure 4 shows performance metrics for the different models.

The Swin-T backbone did better than the others in the final assessment based on the confusion-matrix data, with an AUC of 0.983, a sensitivity of 95.1%, a specificity of 93.1%, and an accuracy of 94.1%. ViT-B/16 had an AUC of 0.972 and a balanced sensitivity-specificity of around 90%. BEiT-B/16, on the other hand, had an AUC of 0.613 and an accuracy of 52.9% as shown in Table 3 and Table 4 shows the Stability–Separability Analysis.

These results validate that Swin-T's shifted-window attention proficiently catches both global drawing structure and subtle tremor details, resulting in enhanced clinical discrimination. The difference between BEiT and the other two models shows how important it is to use pretraining algorithms that work well with limited medical picture datasets. The suggested Transformer framework, especially the Swin-T variation, is a mathematically interpretable and very accurate method for early screening for Parkinson's disease. This fits with SDG 3: Good Health and Well-Being.

BEiT's poor performance can be ascribed to its heavy reliance on extensive self-supervised pretraining, which restricts its applicability to limited clinical datasets. Swin Transformer's hierarchical attention mechanism, on the other hand, is more appropriate for handwriting-based PD analysis since it records both global drawing structure and local tremor information.

Clinically speaking, the suggested framework can facilitate early PD screening that is non-invasive and uses inexpensive digital drawing assignments. While not a substitute for clinical diagnosis, such technologies could help neurologists with initial evaluation or be used in mobile health applications for community-level screening.

6. Conclusion

This research introduced a hybrid framework that combines Hilbert Transform–based mathematical modeling with Transformer designs for the early and precise prediction of Parkinson's Disease (PD) using hand-drawn spiral and wave imagery. The Hilbert Transform successfully recovered amplitude and phase information, indicative of tremor strength and motion irregularity, which were integrated with the original picture data to provide an enhanced input representation. The Swin Transformer was the best of the examined models, with an AUC of 0.983, a sensitivity of 0.951, and an accuracy of 94.1%. This means that it was better at capturing both local tremor changes and global drawing structure. The ViT-B/16 had similar outcomes, however the BEiT-B/16 exhibited restricted generalization on the short clinical dataset. The suggested mathematical-deep learning method provides a clear, non-invasive, and computationally efficient way to find Parkinson's Disease early. Future endeavors will concentrate on expanding this framework to accommodate bigger datasets, including multimodal signals, and integrating real-time handwriting analysis for clinical use. Although encouraging outcomes, additional validation on bigger and multi-modal clinical datasets is required due to the small sample size.

References

[1] Goetz, C.G., Tilley, B.C., Shaftman, S.R., Stebbins, G.T., Fahn, S., Martinez-Martin, P., Poewe, W., Sampaio, C., Stern, M.B., Dodel, R. and Dubois, B., 2008. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results. *Movement disorders: official journal of the Movement Disorder Society*, *23*(15), pp.2129-2170. https://doi.org/10.1002/mds.22340.

[2] Postuma, R.B., Berg, D., Stern, M., Poewe, W., Olanow, C.W., Oertel, W., Obeso, J., Marek, K., Litvan, I., Lang, A.E. and Halliday, G., 2015. MDS clinical diagnostic criteria for Parkinson's disease. *Movement disorders*, *30*(12), pp.1591-1601. https://doi.org/10.1002/mds.26424.

[3] Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N.C., Tung, C.C. and Liu, H.H., 1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London. Series A: mathematical, physical and engineering sciences*, *454*(1971), pp.903-995. https://doi.org/10.1098/rspa.1998.0193.

[4] Boashash, B., 2015. *Time-frequency signal analysis and processing: a comprehensive reference*. Academic press.

[5] Cohen, L., 1995. Time-frequency analysis. *Englewood Cliffs*.

[6] Chang K.-H., Liou H.-H., Chen S.-Y., Chiou J.-C., "Evaluating the different stages of Parkinson's disease using electroencephalography with Holo-Hilbert spectral analysis," *Frontiers in Aging Neuroscience*, 14, 2022. https://doi.org/10.3389/fnagi.2022.832637.

[7] Toffoli S., Lorenzi P., Ferrari G., Rizzo G., "Spiral drawing analysis with a smart ink pen to identify motor symptoms in Parkinson's disease," *Frontiers in Neurology*, 14, 2023. https://doi.org/10.3389/fneur.2023.1093690.

[8] Drotár P., Mekyska J., Rektorová I., Masarová L., Smékal Z., Faundez-Zanuy M., "Analysis of in-air movement in handwriting: A novel marker for Parkinson's disease," *PLoS ONE*, 11(10), e0165108, 2016.

[9] Pereira C. R., Weber S. A. T., Hook C., Rosa G. H., Papa J. P., "Handwriting dynamics for Parkinson's disease diagnosis: A benchmark feature analysis and classification study," *Computer Methods and Programs in Biomedicine*, 153, 53–65, 2018.

[10] Little M. A., McSharry P. E., Hunter E. J., Spielman J., Ramig L. O., "Suitability of dysphonia measurements for telemonitoring of Parkinson's disease," *BioMedical Engineering OnLine*, 8, 23, 2009. https://doi.org/10.1186/1475-925X-8-23.

[11] Tsanas A., Little M. A., McSharry P. E., Ramig L. O., "Accurate telemonitoring of Parkinson's disease progression by noninvasive speech tests," *IEEE Transactions on Biomedical Engineering*, 57(4), 884–893, 2010. https://doi.org/10.1109/TBME.2009.2036000.

[12] Del Din S., Godfrey A., Galna B., Lord S., Rochester L., "Free-living gait characteristics in ageing and Parkinson's disease: impact of environment and ambulatory bout length," *Journal of NeuroEngineering and Rehabilitation*, 13, 46, 2016. https://doi.org/10.1186/s12984-016-0154-5.

[13] Lipsmeier F., Taylor K. I., Kilchenmann T., Wolf D., Scotland A., Schjodt-Eriksen J., et al., "Evaluation of smartphone-based testing to generate exploratory outcome measures in a phase 1 Parkinson's disease clinical trial," *Movement Disorders*, 33(8), 1287–1297, 2018. https://doi.org/10.1002/mds.27376.

[14] Loh H. W., Hong W., Ooi C. P., Chakraborty S., Barua P. D., Deo R. C., Soar J., Palmer E. E., Acharya U. R., "Application of deep learning models for automated identification of Parkinson's disease: a review (2011–2021)," *Sensors*, 21(21), 7034, 2021. https://doi.org/10.3390/s21217034.

[15] Sieberts S. K., Schaffter T., Phelan M., Stolovitzky G., Kellen M., Chaibub Neto E., et al., "Crowdsourcing digital health measures to predict Parkinson's disease severity: the Parkinson's Disease Digital Biomarker DREAM Challenge," *npj Digital Medicine*, 4, 111, 2021. https://doi.org/10.1038/s41746-021-00476-7.