Detection of Parkinson's Disease Using Spiral Drawings and Machine Learning Approaches

Axell Albano Gutiérrez Ramírez*, Antonio Alarcón Paredes, and Cornelio Yáñez Márquez

Centro de Investigación en Computación del Instituto Politécnico Nacional (CIC-IPN). Av. Juan de Dios Bátiz, esq. Miguel Othón de Mendizábal, Col. Nueva Industrial Vallejo, Alcaldía Gustavo A. Madero, C.P. 07738, CDMX. Mexico

Email: agutierrezr1502@alumno.ipn.mx (A.A.G.R.); aalarcon@cic.ipn.mx (A.A.P.); cyanez@cic.ipn.mx (C.Y.M.)

*Corresponding author

***Abstract*—This paper presents the design of a pre-diagnostic system for Parkinson's Disease (PD) based on the analysis of hand-drawn spiral patterns using machine learning techniques. Parkinson's disease is a progressive neurodegenerative disorder whose early motor manifestations—such as micrographia and tremors—can be reflected in fine motor tasks like handwriting. Although handwriting disturbances are not part of the core diagnostic criteria, they are frequently observed in early stages and are recognized by the Movement Disorder Society as supportive markers. In this study, grayscale spiral drawings are preprocessed and binarized using Otsu's thresholding method. From each image, 100 equidistant pixel coordinates are extracted where the trace is present, forming structured feature vectors. These coordinates are then used to train and evaluate several machine learning classifiers, including Random Forest, k-Nearest Neighbors, and Support Vector Machines. The proposed method prioritizes simplicity, explainability, and computational efficiency, offering a lightweight yet effective tool for early Parkinson's detection. Experimental results demonstrate the model's potential as a clinically relevant and accessible diagnostic support system.**

***Keywords*—Parkinson disease, spiral drawing, random forest, Otsu's thresholding, feature extraction**

I. Introduction

A Parkinson's Disease (PD) is a progressive neurodegenerative disease of the motor system that is characterized by tremors, rigidity, and bradykinesia; which gradually deteriorates Early diagnosis of PD still is a major challenge in clinical routine, as the symptoms usually appear insidiously and may be attributed to physiological aging and other diseases. Early and accurate diagnosis is crucial to both delaying the progression of the disease and to enhancing quality of life for the patient via timely interventions [1–3].

One of the distinctive characteristics of PD is the usual positive response to dopaminergic treatment. Levodopa is still regarded as the most effective treatment for motor symptoms including bradykinesia and rigidity; and is featured on the WHO Model Lists of Essential Medicines [4]. This dopaminergic therapy responsiveness is clinically important as well as a useful supportive category in the differential diagnosis of PD [5].

Meanwhile, significant cost-effectiveness burdens have arisen from efforts to identify clinical, imaging, and molecular biomarkers for the differential diagnosis of Parkinson's Disease (PD). However, many of these approaches still fail to achieve sufficient diagnostic accuracy to reliably support early or definitive treatment decisions for PD [6, 7].

Physicians typically use standardized motor assessments in routine clinical care to assess the extent of disability in patients, for instance, using the Unified Parkinson's Disease Rating Scale (UPDRS) [8]. These tests are low-cost, simple-to-administer methods and can potentially be used as adjuncts in the diagnostic process, including in combination with digital, quantitative assessments such as drawing tests.

Although handwriting impairments such as micrographia and dysgraphia are not part of the core diagnostic criteria for PD, they are often among the earliest observable motor signs prompting clinical consultation. For this reason, the Movement Disorder Society (MDS) includes handwriting analysis among its recommended ancillary tests in the diagnostic evaluation of PD [2].

Over the past two decades, the quantitative analysis of handwritten tasks has emerged as a promising objective biomarker for early diagnosis, as handwriting integrates various motor and cognitive systems that are frequently affected in the early stages of Parkinsonian pathophysiology [9].

II. Literature Review

Over the past few years, Artificial Intelligence (AI) and pattern recognition-based Computer-Aided Diagnostic (CAD) systems have emerged as a highly promising tools to assist medical experts in diagnosing PD. Among various non-invasive monitoring techniques, spiral tracing tasks have proven effective in capturing subtle motor dysfunctions associated with Parkinsonian tremors.

Manuscript received June 20, 2025; revised July 15, 2025; accepted September 3, 2025; published 16, 2026.

However, their analysis requires the identification of characteristic patterns with undulatory morphology [10].

These sketches contain useful spatial and kinematic information for the purpose of an automatic separation between healthy subjects and Patients with Parkinson's Disease (PD) [11, 12].

Bange *et al.* [13] conducted a study to investigate differences in spiral drawing performance between 29 patients with PD and 31 healthy controls. Participants traced Archimedean spirals on predefined templates using a digital pen, while their brain activity was simultaneously recorded via Electroencephalography (EEG). Resting-state functional Magnetic Resonance Imaging (MRI) scans were conducted, and positional signals were exported to MATLAB for further processing. Although no significant differences were found in drawing duration, average velocity, or acceleration, PD patients exhibited higher sample entropy values and greater irregularities, particularly in the non-dominant hand. These findings suggest subtle impairments in fine motor control even in the absence of overt kinematic differences. An important limitation of this type of study is the need for specialized equipment, which may not be readily accessible.

At Mount Sinai Beth Israel Medical Center, 138 patients with PD including 50 with early-stage PD and 150 healthy controls were recruited. Each participant traced 10 spirals with each hand on a digitizing tablet, generating coordinate (x, y, z), pressure, and timing data. Several indices were derived to evaluate spiral performance, including overall execution severity, shape irregularity, kinematic disturbances, rigidity, mean velocity, and width variability. The groups were matched by age and handedness, and significant differences in spiral indices were found between PD patients and controls. A machine learning model trained on this data achieved strong discriminative validity (sensitivity = 0.86, specificity = 0.81), which remained consistent even among early-stage PD patients [14]. However, this approach requires a digitizing tablet and more comprehensive data to ensure broader generalization.

In another study, Varalakshmi *et al.* [15] used a dataset of 102 scanned hand-drawn spirals to diagnose PD using machine learning and deep learning models, including pretrained networks such as RESNET50, VGG16, AlexNet, and VGG19. Their hybrid RESNET50 + SVM model achieved the best performance metrics, reaching an accuracy of 98.45%, sensitivity of 99%, and specificity of 98%. Although the study achieved excellent results using only the images, it required greater computational resources due to the use of transfer learning and hybrid models, and it sacrificed interpretability.

More recently, Shanmugam and Arumugam [16] proposed a hybrid optimization-enabled deep learning approach that integrates voice signals and hand-drawn images for multimodal Parkinson's classification, taking into account stroke dynamics, pen pressure, and stroke-width variability which capture geometric smoothness, motor-control consistency, and force modulation, respectively. Their results highlight the promise of combining different input modalities with deep architectures, although such methods require greater computational resources and careful optimization strategies.

This work proposes a pre-diagnostic system for Parkinson's disease based on the analysis of hand-drawn spiral images. Unlike deep learning approaches that process the entire image, our method focuses on extracting a compact and informative set of spatial features. Nevertheless, integrating deep learning methods could potentially improve the generalization capability of the model [16, 17]. Each spiral image is first preprocessed and binarized using Otsu's thresholding method. From the resulting binary mask, the system identifies 100 equidistant pixel locations where the spiral trace (foreground) is present—specifically, where the pixel value is zero. These (x, y) coordinates are stored in a structured matrix, which serves as the input for traditional machine learning classifiers.

This approach emphasizes simplicity, interpretability, and computational efficiency while retaining relevant geometric information. The performance of several classifiers is evaluated under a consistent experimental framework, demonstrating the effectiveness of this feature-based strategy for early Parkinson's detection [18].

Traditional handcrafted methods remain relevant in low-data scenarios due to their simplicity, interpretability, and low computational cost, making them a useful complement to routine clinical assessments. Unlike prior work that uses high-dimensional kinematic data or raw pixel inputs, this study proposes a compact geometric descriptor. This enables efficient classification with standard machine learning models and provides insight into trace irregularities characteristic of Parkinsonian movement.

III. Materials and Methods

The proposed methodology involves several steps ranging from data acquisition to model evaluation. The approach emphasizes the extraction of spatial features from binarized images of hand-drawn spirals, followed by the application of supervised machine learning techniques.

A. Data Collection and Preprocessing

The dataset used in this study consists of grayscale spiral drawings captured from both, Parkinson's patients and healthy individuals. Each image is resized to a uniform resolution of 256×256 pixels.

Fig. 1 shows a comparison between an original spiral image taken directly from the dataset and the same image after applying preprocessing steps such as grayscale conversion, image binarization, and median filtering. The original image, as provided in the dataset, exhibits noise and contrast variations that may hinder direct feature extraction. To address this, a 3×3 median filter was applied to enhance local homogeneity and preserve edge structures.

The data used in this study were obtained from Kaggle, from a dataset titled Parkinson's Drawings [19], which includes both spiral and wave images. For this experiment, only the spiral drawings were used. The training set (70%) consists of 72 images, 36 from patients with Parkinson's

disease and 36 from healthy individuals, while the testing set (30%) includes 30 images, with 15 from each class. The dataset was originally derived from the work by Zham *et al.* [20], in which spiral drawings were collected to differentiate stages of Parkinson's disease based on speed and pen-pressure metrics during sketching. For these experiments, we combined the dataset and generated *k* stratified folds for cross-validation.

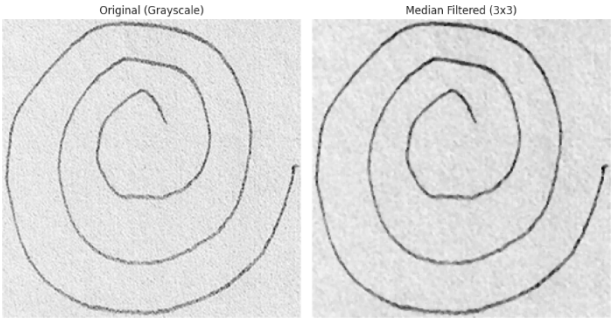

Fig. 1. Filter spiral image example.

Binarization is the process of converting a grayscale image $I(x,y) \in [0,255]$ into a binary image $B(x,y) \in \{0,1\}$ by separating foreground (the trace) from the background. A simple thresholding rule is defined as:

$$B(x,y) = \begin{cases} 1 & if\ l(x,y) > T \\ 0 & otherwise \end{cases} \quad (1)$$

where T is a threshold value.

In this work, Otsu's method is used to automatically determine the optimal threshold T^* that maximizes the between-class variance $\sigma_b^2(T)$, calculated as:

$$\sigma_b^2(T) = \omega_0(T)\omega_1(T)[\mu_0(T) - \mu_1(T)]^2 \quad (2)$$

here:

- $\omega_0(T)$ and $\omega_1(T)$ are the probabilities of the two-pixel classes (background and object),
- $\mu_0(T)$ and $\mu_1(T)$ are the mean intensities of those classes.

The optimal threshold is then selected as:

$$T^* = \arg\max_T \sigma_b^2(T) \quad (3)$$

This method is robust for images with bimodal histograms, as is common in digitized spiral drawings, as shown in Fig. 2.

A 3×3 median filter is used to remove salt-and-pepper noise while preserving edges in the image like shows Fig. 3. For each pixel $I(x,y)$, the value is replaced with the median of its 3×3 neighborhood $\mathcal{N}_{3X3}(x,y)$, defined as:

$$\mathcal{N}_{3X3}(x,y) = \{I(i,j) \mid i \in \{x-1, x, x+1\},\ j \in \{y-1, y, y+1\}\} \quad (4)$$

The output image is computed as:

$$I_{out}(x,y) = median\big(\mathcal{N}_{3X3}(x,y)\big) \quad (5)$$

This nonlinear operation helps to reduce noise without blurring critical structural information, such as the spiral trace.

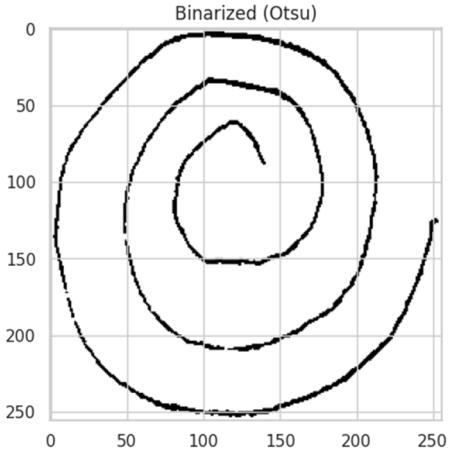

Fig. 2. Binarized image example.

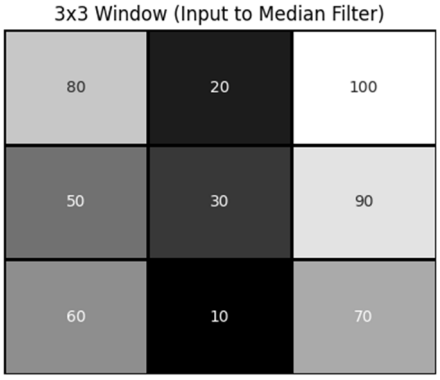

Fig. 3. Median filter.

B. Feature Extraction

Once the binary image is obtained, spatial coordinates of the foreground pixels (those with value 0) are extracted. From this set of pixels, 100 equidistant points are selected using linear sampling over the length of the array. These points represent the trace geometry in a highly compact form.

The (x,y) coordinates of the 100 selected points are then stored as feature vectors in a structured format: each image yields a 200-dimensional vector (100 for x, 100 for y). An additional column is appended to indicate the class label (1 for Parkinson's, 0 for healthy). Fig. 4 illustrates the binary spiral image with the 100 equidistant x-pixel positions (in black) that define the spatial trace used as input for classification.

C. Model Training Evaluation

Once the feature matrix is assembled, it is used to train and evaluate multiple machine learning classifiers including Random Forests, Support Vector Machines (SVM), k-Nearest Neighbors (k-NN), Multi-Layer Perceptron and Naive Bayes; with the hyperparameters defined in Table I. All models were trained and evaluated using the dataset's fixed scheme of 5-times repeated,

stratified 10-fold cross-validation with a random seed of 42. We averaged the scores across the 50 folds to obtain more stable estimates and ensure a fair comparison with the Kaggle dataset [19]. Performance metrics such as sensitivity, specificity, balanced accuracy and F1-score are calculated to assess the effectiveness of each classifier in detecting Parkinson's disease.

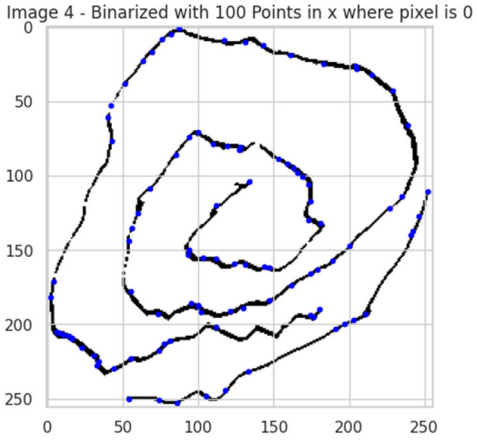

Fig. 4. Image binarized with 100 points in X where pixel is 0.

TABLE I. HYPERPARAMETERS

Algorithm	Hyperparameters
Random Forest	n_estimators: 100 Max depth: 10 Min samples split: 4 Random seed: 42
Multilayer Perceptron	Hidden layers: (50, 20) Activation: ReLU Optimizer: Adam Learning rate: 0.0001 Max iterations: 500 Random seed: 42
kNN	*k*: 5 Distance metric: Manhattan
Naïve Bayes	Type: Gaussian Naive Bayes Assumption: Feature independence
SVM (linear)	Kernel: linear Random seed: 42

This methodology prioritizes computational efficiency and explainability, making it suitable for integration into real-time pre-screening tools or as support for clinical diagnostics.

IV. RESULT AND DISCUSSION

The source code and experimental configurations are available at: https://github.com/Axell54/Parkinson_wave1.

A. Description and Configuration of the Models

1) SVM (linear)

Linear Support Vector Machines (SVMs) are supervised learning models that aim to find the optimal hyperplane that separates data into distinct classes with the maximum possible margin. They are particularly effective in high-dimensional spaces and have been widely applied in medical and pattern recognition tasks due to their robustness and generalization capacity [21]. A random seed equal to 42 was used in the experiment for reproducibility.

2) Naive bayes

The Naive Bayes is a probabilistic classifier based on Bayes' theorem, assuming independence among features. Despite its simplicity, it is particularly effective in high-dimensional datasets and is commonly used in text classification and early-stage medical diagnosis [22]. The experiment applies Bayes' theorem with the "naive" feature independence assumption.

3) K-NN

The k-Nearest Neighbors (k-NN) algorithm classifies a sample based on the majority label of its *k* closest neighbors in the feature space. Despite its simplicity, k-NN remains a powerful non-parametric method for pattern recognition and is often employed as a baseline in classification problems [23]. This model occupied the Manhattan distance and 5 nearest neighbors.

4) Multilayer Perceptron (MLP)

The Multilayer Perceptron (MLP) is a feedforward artificial neural network consisting of an input layer, hidden layers, and output layer. It uses backpropagation for training and can model fairly complicated nonlinear relationships between input features and output class. MLPs are widely used in bioinformatics and image classification [24]. The Multilayer Perceptron (MLP) model adopted in this experiment was set with two hidden layers with 50 and 20 neurons, respectively.

We used ReLU (Rectified Linear Unit) as an activation function, because of its ability to manage non-linearities, prevent vanishing gradients, and its efficiency. The model was trained using the Adam optimizer with an adaptive learning rate that was initialized to 0.0001; the stopping criterion was 500 iterations. Weights were initialized using a random seed of 42 for reproducibility.

5) Random forest

Random Forest is an ensemble learning method that constructs a multitude of decision trees during training and outputs the class that is the mode of the classes predicted by individual trees. It offers high accuracy, robustness to overfitting, and the ability to handle both categorical and numerical data, making it suitable for classification tasks in noisy biomedical datasets [25]. In this study, the Random Forest classifier was configured with 100 trees, a maximum depth of 10, a minimum split of 4 samples, and a random seed of 42 to ensure reproducibility.

B. Comparison of Results

1) Performance metrics summary

In this study, the selected performance metrics revealed strong results for the Random Forest classifier, which consistently outperformed the other models. The Multilayer Perceptron, 5-NN, and Naive Bayes exhibited moderate performance across the evaluated datasets. In contrast, the Linear SVM model showed poor performance, as evidenced by the results presented in Table I and illustrated in Fig. 5.

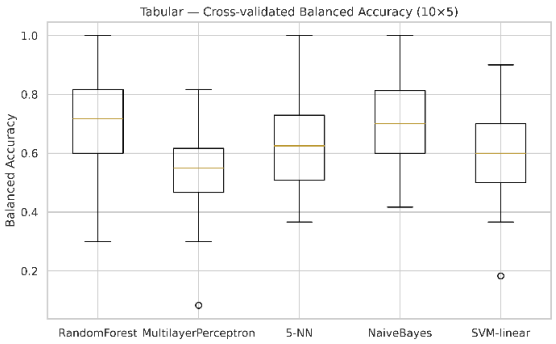

Fig. 5. Balanced accuracy comparison between models.

The metrics in Table II correspond to the Parkinson class only. Values are reported as percentages and represent the mean over the 5×10 stratified cross-validation; standard deviations are shown in parentheses. This approach provides a balanced assessment of model behavior.

TABLE II. PERFORMANCE METRICS

Algorithm	Specificity	Sensitivity	F1-Score	Balanced Accuracy
Random Forest	66.67(21.90)	73.93(18.18)	71.93(15.40)	70.30(15.10)
Multilayer Perceptron	45.13(24.65)	62.47(29.68)	55.86(20.33)	53.80(14.67)
5NN	73.27(18.97)	51.87(20.77)	57.43(18.29)	62.57(14.72)
Naïve Bayes	85.40(17.48)	54.13(24.92)	61.61(22.33)	69.77(13.26)
SVM (linear)	62.40(19.73)	56.53(19.66)	58.35(16.71)	59.47(14.48)

Paired Wilcoxon signed-rank tests (Table III) were conducted comparing the Random Forest (top balanced-accuracy) against each baseline across the 5×10 stratified CV folds; p-values were adjusted using Holm's step-down procedure ($\alpha = 0.05$) to assess significance.

TABLE III. INVARIANT FEATURE EXTRACTION

Algorithm vs Random Forest	Precision (%)	Recall (%)
Multilayer Perceptron	0.000003177343817	0.00001270937527
SVM (linear)	0.001463100538	0.002926201076
5-NN	0.001034547711	0.003103643132
Naïve Bayes	0.7207078301	0.7207078301

After applying Holm's correction, Random Forest remained significantly better than all baselines except Naive Bayes, for which the difference was not significant (adjusted $p > 0.05$).

The results obtained in this study indicate that the Random Forest classifier outperformed other machine learning models across all performance metrics, as shown in Table II and Fig. 5. This superior performance can be attributed to Random Forest's ensemble nature and its robustness in handling high-dimensional, noisy, and non-linearly separable data—properties often observed in image-derived features from hand-drawn spirals. Unlike models such as Linear SVM or Naive Bayes, which assume linear boundaries or feature independence, Random Forest is capable of modeling complex decision surfaces and feature interactions without strong parametric assumptions [22].

Previous studies using sketch-based or handwriting datasets for Parkinson's disease detection have shown thatmanually extracted geometric and kinematic features can be as effective as deep learning approaches in scenarios with relatively small, domain-specific datasets. Notably, Zham *et al.* [20] proposed using composite indices of speed and pen pressure during spiral drawing tasks to distinguish between Parkinson's disease stages, while Al-Kasasbeh *et al.* [22], systematically reviewed various machine learning approaches for handwriting-based detection. These studies suggest that conventional feature engineering methods often outperform deep learning when data is limited. While powerful, deep learning models such as Convolutional Neural Networks (CNNs) typically require large, labeled datasets and substantial computational resources, with an increased risk of overfitting on small samples [23].

Moreover, traditional clinical approaches for Parkinson's diagnosis such as the UPDRS motor score or neuroimaging techniques are either subjective, resource-intensive, or invasive. In contrast, the proposed method is low-cost, non-invasive, and automatable, relying only on simple drawing tasks and lightweight preprocessing. The advantage of Random Forest in this context lies in its ability to generalize well without intensive hyperparameter tuning and its interpretability, as feature importance scores can provide insights into the underlying structure of the drawing data.

In summary, the Random Forest model demonstrates strong potential for practical deployment in early Parkinson's screening tools, particularly in low-resource settings or as a complementary system to clinical evaluation. However, further validation on larger and more diverse datasets remains necessary.

2) Geometric invariant feature extraction

Additional invariant features were incorporated into the best-performing model. First, the spiral trace was centered by subtracting the centroid of the extracted coordinates. Furthermore, Hu Moments were computed from the binarized image to capture global shape descriptors that are invariant to rotation, translation, and scale. The original spatial vectors were replaced by these invariant features.

TABLE IV. INVARIANT FEATURE EXTRACTION

Algorithm	Specificity	Sensitivity	F1-Score	Balanced Accuracy
Random Forest	66.67(21.90)	73.93(18.18)	71.93(15.40)	70.30(15.10)
Centroid normalization	72.47(19.93)	77.13(17.88)	75.89(14.03)	74.80(14.03)
Centroid normalization + Hu moments	71.73(21.99)	77.00(16.60)	75.71(12.14)	74.37(13.33)

However, as shown in Table IV, this integration led to an increase in the model's performance metrics when

using the same hyperparameters as in previous experiments.

3) Comparison with Convolutional Neural Networks (CNN)

For this experiment, we adopted a technique similar to that employed by the best-performing models reported in the literature, particularly in the work of Varalakshmi *et al.* [16]. Specifically, we used pretrained convolutional neural networks ResNet50, VGG16, and VGG19 to extract image features, which were then classified using a Support Vector Machine (SVM) with a linear kernel and a regularization parameter C = 1. A batch size of 16 was used during feature extraction. Although the performance metrics obtained were comparable to those of our proposed Random Forest model (as shown in Table V and Fig. 6), the execution time for the deep learning-based approach exceeded 40 s per model, whereas the Random Forest model required less than 5 s to run.

TABLE V. COMPARISON WITH CNN

Algorithm	Specificity	Sensitivity	F1-Score	Balanced Accuracy
Random Forest	66.67(21.90)	73.93(18.18)	71.93(15.40)	70.30(15.10)
Centroid normalization + Hu moments	71.73(21.99)	77.00(16.60)	75.71(12.14)	74.37(13.33)
ResNet50 + SVM	91.20(10.81)	78.53(18.54)	83.30(13.58)	84.87(10.87)
VGG16+SVM	83.13(15.63)	80.80(18.87)	81.52(14.12)	81.97(12.88)
VGG19+SVM	79.07(13.70)	72.40(22.08)	74.01(16.03)	75.73(12.40)

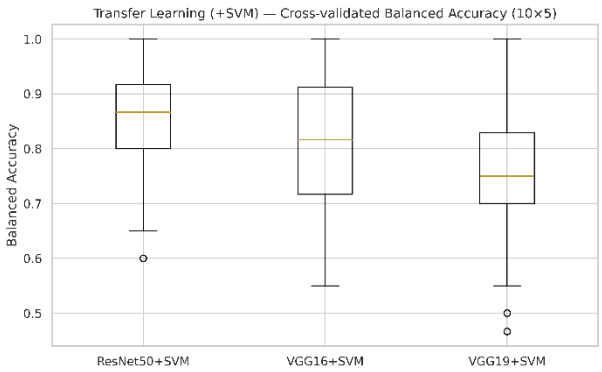

Fig. 6. Balanced accuracy comparison between CNN models.

V. CONCLUSION

A. Main Features

This study shows good performance in approaching Parkinson's disease diagnosis using geometric features extracted from binarized spiral drawings. The proposed methodology demonstrated promising results, particularly when evaluated with Random Forest classifier.

Among the models tested, Random Forest achieved the highest overall performance across accuracy, precision, recall, and F1-score, highlighting its robustness and suitability for biomedical classification tasks involving hand-drawn patterns.

Our approach offers a complementary, lightweight, and interpretable alternative suited to low-resource settings and early screening. Moreover, integrating deep-learning function approximation could further enhance accuracy and robustness.

B. Limitation and Future Work

Future work will focus on validating the proposed method with larger and more diverse datasets, as well as exploring hybrid architectures that combine handcrafted feature extraction with deep learning components to further enhance performance and generalizability.

A key limitation of the proposed approach is the risk of overfitting given the relatively small dataset. Although repeated stratified cross-validation was used to mitigate this, models trained on limited samples may still learn noise or dataset-specific artifacts rather than generalizable Parkinson's signatures. This concern is especially salient in medical applications, where robust generalization is essential for safe deployment. Future work should validate the method on larger and more diverse cohorts ideally with external (multi-site) test sets and enforce patient-level separation to prevent information leakage.

A practical avenue is to extend the current coordinate-based representation with additional descriptors: (i) curvature (e.g., local turning curvature κ along the trace), (ii) stroke-width variability (estimated from the binarized mask via distance transform or skeleton-to-boundary distances), and (iii) pen-pressure (when captured by a digital tablet) can provide complementary information.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Axell Albano Gutiérrez Ramírez conceived and designed the study, implemented the data preprocessing and feature extraction methods, and conducted the machine learning experiments. Antonio Alarcón Paredes and Cornelio Yáñez Márquez supervised the methodology, guided the interpretation of results, and reviewed and edited the manuscript. Authors approved the final version of the manuscript.

ACKNOWLEDGMENT

The authors would like to thank the Centro de Investigación en Computación (CIC-IPN) for their academic support, and the organizers of the dataset

"Distinguishing Different Stages of Parkinson's Disease Using Composite Index of Speed and Pen-Pressure of Sketching a Spiral" for making the data publicly available.

REFERENCES

[1] R. D. Dorsey, T. Sherer, M. S. Okun, and B. R. Bloem, "The rising prevalence of Parkinson's disease," *Lancet Neurology*, vol. 17, no. 11, pp. 939–940, 2018.

[2] R. B. Postuma *et al.*, "MDS clinical diagnostic criteria for Parkinson's disease," *Movement Disorders*, vol. 30, no. 12, pp. 1591–1601, 2015.

[3] World Health Organization. (Aug. 2023). Parkinson disease. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/parkinson-disease

[4] World Health Organization. (Jul. 2023). WHO model lists of essential medicines. [Online]. Available: https://www.who.int/publications/i/item/WHO-MHP-HPS-EML-2023.02

[5] W. C. Koller, "When does Parkinson's disease begin?" *Neurology*, vol. 42, no. 4, pp. 27–31, 1992.

[6] B. R. Bloem, M. Okun, and R. D. Dorsey, "The challenges of predicting Parkinson's disease," *Lancet Neurology*, vol. 20, no. 2, pp. 89–90, 2021.

[7] A. E. Lang and A. J. Espay, "Disease modification in Parkinson's disease: Current approaches, challenges, and future considerations," *Movement Disorders*, vol. 33, no. 5, pp. 660–677, 2018.

[8] S. Fahn and R. L. Elton, "Unified Parkinson's disease rating scale," in *Recent Developments in Parkinson's Disease,* vol. 2, S. Fahn, C. D. Marsden, D. B. Calne, and M. Goldstein, Eds. Florham Park, NJ: Macmillan Health Care Information, 1987, pp. 153–163.

[9] E. Díaz, J. Vargas, A. Teulings, and M. Serranová, "Handwriting analysis in Parkinson's disease: Current trends and challenges," *Frontiers in Neurology*, vol. 13, 874203, 2022.

[10] M. A. Thenganatt and E. D. Louis, "Distinguishing essential tremor from Parkinson's disease: Bedside tests and laboratory evaluations," *Expert Review of Neurotherapeutics*, vol. 12, pp. 687–696, 2012.

[11] O. Alniemi and H. F. Mahmood, "Convolutional neural network for the detection of Parkinson disease based on hand-draw spiral images," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 30, no. 1, pp. 267–275, 2023.

[12] V. Rajasekar, K. Sathya, and P. Suresh, "Parkinson's disease detection using deep learning," *Grenze International Journal of Engineering & Technology (GIJET)*, vol. 10, no. 2, Part 4, pp. 3281–3287, 2024.

[13] M. Bange, G. Gonzalez-Escamilla, T. Marquardt *et al.*, "Deficient interhemispheric connectivity underlies movement irregularities in Parkinson's disease," *Journal of Parkinson's Disease*, vol. 12, pp. 381–395, 2022.

[14] M. S. Luciano, C. Wang, R. A. Ortega *et al.*, "Digitized spiral drawing: A possible biomarker for early Parkinson's disease," *PLoS ONE*, vol. 11, e0162799, 2016.

[15] P. Varalakshmi, K. Lakshmi, and G. V. S. S. Raju, "Parkinson's disease detection using deep learning and hybrid models on spiral drawings," *Grenze International Journal of Engineering and Technology*, vol. 10, no. 2, part 4, pp. 3281–3287, Jun. 2024

[16] Y. L. Cun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, 2015.

[17] M. Díaz, D. Contreras, M. Parra, J. Villalobos, and F. Martínez, "Use of deep learning for Parkinson's disease detection through drawing tasks," *Electronics*, vol. 10, no. 8, p. 907, 2021.

[18] K. Mader. (2019). Parkinson's drawings. *Kaggle* [Online]. Available: https://www.kaggle.com/datasets/kmader/parkinsons-drawings

[19] P. Zham, D. K. Kumar, P. Dabnichki, S. P. Arjunan, and S. Raghav, "Distinguishing different stages of Parkinson's disease using composite index of speed and pen-pressure of sketching a spiral," *Frontiers in Neurology*, vol. 8, p. 435, 2017. doi: 10.3389/fneur.2017.00435

[20] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, pp. 273–297, 1995.

[21] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[22] S. Haykin, *Neural Networks and Learning Machines*, 3rd ed. Pearson, 2009.

[23] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Trans. Inf. Theory*, vol. 13, no. 1, pp. 21–27, 1967.

[24] M. R. Al-Kasasbeh, A. R. Al-Ali, A. A. Al-Ali, and H. Alashaikh, "Handwriting-based Parkinson's disease detection using machine learning approaches: A review," *IEEE Access*, vol. 9, pp. 74195–74209, 2021.

[25] S. Shanmugam and C. Arumugam, "Hybrid ladybug Hawk optimization-enabled deep learning for multimodal Parkinson's disease classification using voice signals and hand-drawn images," *Network: Computation in Neural Systems*, pp. 1–43, 2025. https://doi.org/10.1080/0954898X.2025.2457955