Journal of Health Informatics

J Health Inform. 2026;18:e1860
DOI: https://doi.org/10.59681/2175-4411.v18.iEspecial.2026.1860

ARTIGO COMPLETO
CBIS 2026

Identificação da doença de Parkinson utilizando dados de sensores capturados durante o desenho de espirais

Identification of Parkinson's disease using sensor data captured during spiral drawing

Geovanna Cunha Andrade Silva[1], Giordana de Farias Franco Bueno Bucci[1], Fabrizzio Soares[1], Juliana Paula Felix[1] ✉

ORCID: 0009-0007-9628-7569 · 0009-0009-2511-610X · 0000-0003-1598-1377 · 0000-0003-4095-1639

[1] Universidade Federal de Goiás, Goiânia (GO), Brasil.

FICHA DO ARTIGO

TIPO
Artigo completo · CBIS 2026

IDENTIFICADOR
e1860

DOI
10.59681/2175-4411.v18.iEspecial.2026.1860

RECEBIDO
17/05/2026

APROVADO
20/07/2026

PROCESSO DE AVALIAÇÃO
Double Blind Review

EDITOR RESPONSÁVEL
Juliano de Souza Gaspar

ACESSO
Aberto, CC BY 4.0

AUTOR CORRESPONDENTE
Juliana Paula Felix
julianafelix@ufg.br

LICENÇA

Este é um artigo publicado em acesso aberto sob a licença Creative Commons CC BY 4.0, que permite uso, distribuição e reprodução em qualquer meio, desde que o trabalho original seja corretamente citado.

CC BY 4.0

RESUMO

Objetivo: Propor e avaliar uma abordagem de aprendizado de máquina para a identificação da Doença de Parkinson a partir de um conjunto de características extraídas de desenhos espirais realizados em mesa digitalizadora. **Método:** Utilizou-se o conjunto de dados público Parkinson Disease Spiral Drawings Using Digitized Graphics Tablet como entrada para cinco classificadores clássicos, avaliados por validação cruzada estratificada com sobreamostragem SMOTE (Synthetic Minority Over-sampling Technique) aplicada somente ao conjunto de treino. **Resultados:** O Random Forest obteve o melhor desempenho, com acurácia de 91,71%, sensibilidade de 94,85% e *F1-Score* de 94,69%. **Conclusão:** Os resultados sugerem o potencial da abordagem como ferramenta de apoio digital ao diagnóstico da Doença de Parkinson.

Descritores: doença de Parkinson; aprendizado de máquina; diagnóstico por computador.

ABSTRACT

Objective: To propose and evaluate a machine learning approach for the identification of Parkinson's Disease, based on a set of features extracted from spiral drawings performed on a digitized tablet. **Method:** The publicly available Parkinson Disease Spiral Drawings Using Digitized Graphics Tablet dataset was used as input to five classic classifiers, evaluated by stratified cross-validation with SMOTE (Synthetic Minority Over-sampling Technique) oversampling applied only to the training set. **Results:** Random Forest achieved the best performance, with an accuracy of 91.71%, a sensitivity of 94.85% and an *F1-Score* of 94.69%. **Conclusion:** The results suggest the potential of the approach as a digital diagnostic support tool for Parkinson's Disease.

Keywords: Parkinson disease; machine learning; diagnosis; computer-assisted.

CBIS 2026 XXI CONGRESSO BRASILEIRO DE INFORMÁTICA EM SAÚDE

Artigo apresentado no XXI Congresso Brasileiro de Informática em Saúde (CBIS 2026), realizado de 23 a 25 de setembro de 2026, em Brasília (DF), e publicado na seção CBIS 2026 do Journal of Health Informatics.

COMO CITAR ESTE ARTIGO: Silva GCA, Bucci GFFB, Soares F, Felix JP. Identificação da doença de Parkinson utilizando dados de sensores capturados durante o desenho de espirais. J Health Inform 2026;18:e1860. doi: https://doi.org/10.59681/2175-4411.v18.iEspecial.2026.1860

1 INTRODUÇÃO

A Doença de Parkinson (DP) é a segunda patologia neurodegenerativa mais prevalente no mundo, afetando aproximadamente 1% da população acima de 60 anos [1]. Os sintomas cardinais incluem tremores em repouso, rigidez muscular, bradicinesia e instabilidade postural, manifestações que comprometem progressivamente a capacidade motora dos pacientes e afetam de forma expressiva a qualidade de vida [2]. A DP compromete, ainda, funções relacionadas à escrita manual e ao controle motor fino, tornando tarefas cotidianas cada vez mais difíceis à medida que a doença avança [2].

Apesar de ser uma doença amplamente estudada, a DP não tem cura conhecida [3] e seu diagnóstico clínico convencional baseia-se em critérios subjetivos, como a Unified Parkinson's Disease Rating Scale (UPDRS), que depende da experiência do avaliador e pode apresentar variabilidade entre observadores [4]. Não existe um exame laboratorial ou de imagem definitivo para confirmação do diagnóstico, o que frequentemente resulta em identificação tardia da doença [5]. Esse atraso limita o acesso a intervenções que poderiam retardar a progressão dos sintomas e melhorar o prognóstico.

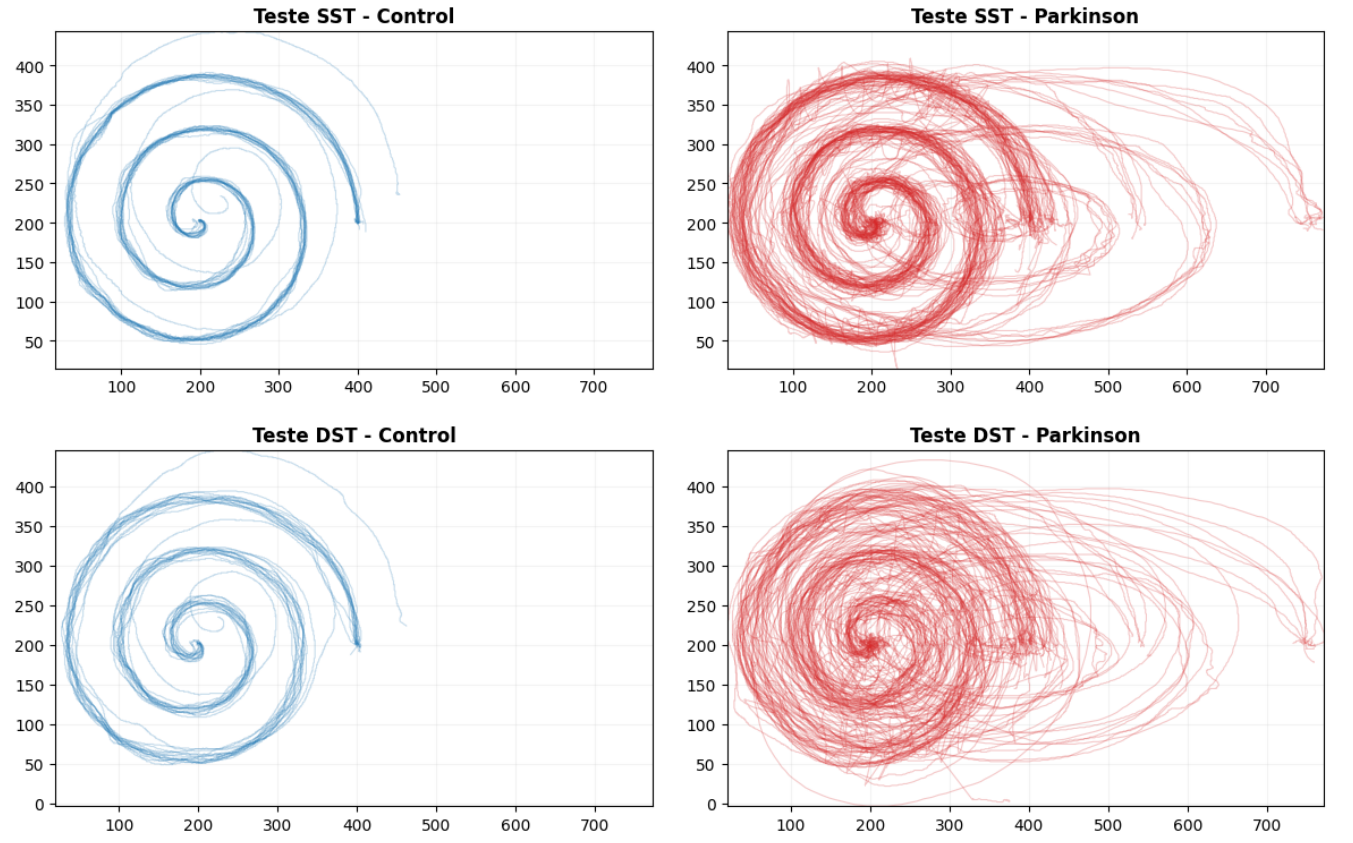
Nesse contexto, a Saúde Digital emerge como campo estratégico para o desenvolvimento de biomarcadores objetivos e acessíveis. A análise computacional de tarefas motoras simples, como o desenho de espirais em mesas digitalizadoras, tem demonstrado potencial para capturar alterações cinemáticas sutis associadas à DP, mesmo em estágios iniciais [6,7]. Diferentemente de exames de imagem ou testes genéticos, essa abordagem é não invasiva, de baixo custo e pode ser implementada em contextos ambulatoriais com recursos tecnológicos simples, o que a torna especialmente relevante para sistemas de saúde com recursos limitados, como o Sistema Único de Saúde (SUS) [8]. A mesa digitalizadora permite registrar, além das coordenadas espaciais do traço, variáveis como pressão da caneta, ângulo de empunhadura e marcações temporais, fornecendo um conjunto expressivo de dados para análise por algoritmos de aprendizado de máquina.

Na literatura, abordagens baseadas em imagens de espirais têm sido exploradas, como Alniemi e Mahmood [9], que aplicaram uma rede neural convolucional sobre imagens de espirais desenhadas à mão, alcançando 93,33% de acurácia; e Farhah [10], que empregou modelos de aprendizado profundo com *transfer learning*, obtendo 89% de acurácia com a arquitetura InceptionV3. Adicionalmente, é possível observar diversos trabalhos que usam o conjunto de dados público chamado *Parkinson Disease Spiral Drawings Using Digitized Graphics Tablet* [11] e exploram a classificação binária entre DP e Controle (CO) a partir de sinais de desenhos espirais digitalizados. Wang *et al.* [12] propuseram um modelo baseado em Support Vector Machines (SVM) com extração de características via Histogram of Oriented Gradients (HOG), utilizando validação cruzada de cinco *folds*, alcançando 93,30% de acurácia. Kamble *et al.* [13] realizaram uma análise detalhada de características cinemáticas derivadas dos testes *Static Spiral Test* (SST) e *Dynamic Spiral Test* (DST), incluindo velocidade, aceleração e *jerk*, obtendo 91,60% de acurácia nos modelos de Regressão Logística e Random Forest com validação cruzada estratificada de dez *folds*. Thakur *et al.* [14] empregaram uma Máquina de Boltzmann Restrita (RBM) combinada com Perceptron Multicamadas (MLP) para a classificação, adotando partição simples (*hold-out*) com 40% dos dados reservados para teste, atingindo 95,32% de acurácia. Vaish [15] aplicou Random Forest com partição simples com divisão de 80/20 entre treino e teste para identificação precoce da DP, obtendo 70,97% de acurácia. Embora esses trabalhos apresentem resultados expressivos, Wang *et al.* [12] utilizam imagens das espirais como entrada, descartando as informações cinemáticas ricas registradas pelos sensores, enquanto Kamble *et al.* [13] e Thakur *et al.* [14] analisam SST e DST de forma independente, sem explorar a complementaridade entre os testes. Além disso, é válido ressaltar que nenhum destes trabalhos aplica técnicas de balanceamento de classes para lidar com o desequilíbrio entre grupos.

Diante desse cenário, este artigo propõe uma abordagem de aprendizado de máquina para a classificação da DP com base em características cinemáticas, espectrais, de pressão, de empunhadura, de fluidez e de hesitação, extraídas diretamente das séries temporais registradas por sensores durante o desenho de espirais. O diferencial metodológico central deste trabalho compreende a exploração de características derivadas de sensores em vez de imagens de espirais, a fusão das características do SST e do DST em um único vetor por participante para classificação conjunta, em vez da análise individual de cada teste, e a aplicação de sobreamostragem SMOTE (Synthetic Minority Over-sampling Technique) internamente a cada *fold* de validação cruzada para mitigar o desbalanceamento de classes. Dessa maneira, o objetivo deste trabalho é propor e avaliar uma abordagem de aprendizado de máquina para a classificação binária da Doença de Parkinson, a partir de características extraídas desses testes e fundidas em um único vetor por participante.

O restante deste artigo está organizado da seguinte forma: na Seção 2, descrevem-se os métodos, o que inclui o conjunto de dados, o pré-processamento dos dados, os classificadores, a avaliação e as métricas; na Seção 3 apresentam-se os resultados; na Seção 4 discutem-se as descobertas em relação à literatura; e na Seção 5 abordam-se as conclusões e perspectivas futuras.

Figura 1. Visualização dos desenhos espirais por grupo e modalidade de teste



Fonte: Elaborado pelos autores (2026).

2 MÉTODOS

2.1 Conjunto de Dados

O conjunto de dados utilizado neste trabalho é o Parkinson Disease Spiral Drawings Using Digitized Graphics Tablet, disponível publicamente por Isenkul *et al.* [11] em 2014. Esse conjunto foi selecionado por ser amplamente adotado na literatura de classificação da DP a partir de sinais motores [12-15], o que permite a comparação direta dos resultados obtidos com trabalhos relacionados.

Os dados foram coletados no Departamento de Neurologia do Hospital Cerrahpaşa da Universidade de Istambul, Turquia, sob supervisão médica especializada. Os participantes realizaram três tipos de testes motores utilizando uma mesa digitalizadora Wacom Cintiq 12WX, com intervalos entre amostras entre 7 e 10 ms nos registros analisados. Os testes foram: Static Spiral Test (SST), em que o participante desenha sobre um padrão de espiral exibido na tela, avaliando a precisão do controle motor fino; Dynamic Spiral Test (DST), em que o padrão de espiral pisca em intervalos regulares, exigindo que o participante continue o desenho de memória, avaliando a manutenção do padrão motor sob demanda cognitiva; e o Circular Motion Test (CMT), em que o participante desenha círculos ao redor de um ponto fixo, avaliando a fluidez do movimento circular contínuo.

A mesa digitalizadora registrou, a cada amostra temporal, as seguintes variáveis: coordenadas X e Y (pixels) da posição da caneta na tela, coordenada Z de afastamento da caneta em relação à superfície (valores próximos de zero durante o contato), nível de pressão (adimensional, escala de 0 a 1023, em que os valores mais altos indicam maior pressão exercida sobre a superfície), ângulo de empunhadura da caneta (décimos de grau, faixa observada de 0 a 3580), marcação temporal (milissegundos) e identificador do tipo de teste (0 para SST, 1 para DST e 2 para CMT). Desse modo, um total de 7 variáveis são registradas por amostra temporal.

Na Figura 1, são mostrados os traços registrados durante os dois testes motores (SST e DST) para os grupos DP e CO, evidenciando visualmente as diferenças cinemáticas entre indivíduos saudáveis e pacientes diagnosticados com Doença de Parkinson.

Tabela 1. Distribuição dos participantes e registros por modalidade de teste

Grupo	Participantes	SST	DST	CMT
Controle	15	15	15	7
Parkinson	62	61	57	47
Total	77	76	72	54

Fonte: Elaborado pelos autores (2026).

O conjunto de dados original disponibilizado por Isenkul *et al.* [11], referido como subconjunto HandWriting (HW), contém registros de 40 participantes: 15 indivíduos saudáveis (CO) e 25 indivíduos com DP diagnosticada clinicamente por neurologistas. Para ampliar o número de amostras da classe Parkinson, incorporou-se o subconjunto *NewDataset*, que contém 37 participantes adicionais com DP coletados sob protocolo similar. O conjunto total compreende 77 participantes (15 controles e 62 pacientes com DP), os quais realizaram até três modalidades de teste, como é mostrado na Tabela 1.

Conforme é possível observar na Tabela 1, apenas 7 participantes do grupo controle realizaram o CMT. Além disso, a inspeção visual dos traços registrados nessa modalidade revelou padrões geométricos incompatíveis com o movimento de espiral, uma vez que todos os registros apresentavam traços em grade na mesma região restrita da mesa digitalizadora, sugerindo possível corrupção nos dados originais. Assim, em razão do reduzido número de amostras de controle e dos problemas de qualidade identificados, optou-se por não utilizar os dados do CMT. Neste trabalho, então, foram utilizados apenas os participantes que completaram ambas as modalidades SST e DST, totalizando 72 participantes: 15 de controle e 57 indivíduos com DP.

2.2 Pré-processamento de Dados

O pré-processamento dos sinais foi realizado em três etapas. Inicialmente, realizou-se a sincronização temporal (reamostragem), uma vez que os registros apresentam intervalos entre amostras distintos entre si. Para garantir um intervalo de tempo constante, definido conforme a equação (a), os sinais foram reamostrados para 133 Hz por interpolação linear, o que permite o cálculo preciso das derivadas cinemáticas [6].

$$\Delta t = \frac{1}{fs} \quad \text{(a)}$$

O sinal contém ruído de alta frequência oriundo da quantização espacial do sensor e da variação do intervalo entre amostras, situado acima da faixa de interesse do tremor. Para atenuar esse ruído, aplicou-se o filtro de Savitzky-Golay [16] às coordenadas X e Y, com janela de 7 amostras (52,6 ms a 133 Hz) e polinômio de 2ª ordem. O filtro apresenta corte de −3 dB em 21 Hz e atenuação inferior a 0,4 dB até 12 Hz, preservando a faixa do tremor de repouso parkinsoniano (4–7 Hz) [17].

Para a extração da *característica* espectral de tremor (Tabela 2), optou-se pelo espectro da velocidade, e não da trajetória, pois a derivação atenua a componente de baixa frequência da própria forma da espiral, que mascararia o tremor caso calculada sobre a trajetória bruta. A potência espectral relativa à potência total do sinal foi calculada na banda de 3–7 Hz (ligeiramente mais ampla do que a faixa de 4–7 Hz do tremor de repouso parkinsoniano [17], de modo a acomodar variabilidade individual na frequência do tremor entre pacientes), consistente com a banda adotada por Kamble *et al.* [13] e Impedovo *et al.* [6]. A frequência de pico, por sua vez, foi localizada em faixa ampla (0,5–25 Hz), sem restrição à banda de tremor.

Por fim, foi feito um processo de engenharia de características em que foram extraídas 32 características organizadas em seis grupos funcionais, descritas na Tabela 2. As características foram definidas com base nas propostas de Kamble *et al.* [13] para características cinemáticas, de fluidez e hesitação, e de Thakur *et al.* [14] e Impedovo *et al.* [6] para características espectrais. Quatro características adicionais foram propostas neste trabalho: taxa de variação da pressão da caneta, raio máximo em relação ao centroide geométrico do traço, comprimento por segundo e taxa de picos de velocidade por segundo. Esta última é uma adaptação do Número de Mudanças de Velocidade (ou *Number of Changes of Velocity*, NCV) proposto por Kamble *et al.* [13], refinada com critério de proeminência mínima de 5% da velocidade máxima para maior robustez a ruídos de hardware.

Com isso, é obtido um único vetor de 64 características por participante, resultante da concatenação das 32 características extraídas do SST com as 32 extraídas do DST. Essa estratégia permite capturar a complementaridade entre as duas modalidades, precisão do controle motor fino (SST) e manutenção do padrão motor sob demanda cognitiva (DST), em uma única representação por participante, diferentemente de abordagens que analisam cada teste de forma independente [13,14].

Tabela 2. Características extraídas por grupo funcional

Grupo	Número	Variáveis de entrada	Características
Pressão	5	Pressão	Média, desvio padrão, mínimo, máximo e taxa de variação da pressão da caneta
Grip	4	GripAngle	Média, desvio padrão, mínimo e máximo do ângulo de empunhadura
Cinemática	9	X, Y (suavizados)	Velocidade média, desvio padrão e máxima; velocidade radial média; aceleração média e máxima; *jerk* médio, máximo e RMS
Fluidez	5	X, Y (suavizados)	Total de picos de velocidade, taxa de picos por segundo, comprimento total do traço, comprimento por segundo e raio máximo
Hesitação	3	Timestamp	Tempo total de desenho, tempo no ar (*air time*) e razão entre *air time* e tempo de desenho
Espectral	6	Velocidade derivada de X, Y (suavizados); Pressão	Potência relativa e frequência de pico do tremor (velocidade e pressão, método de Welch, faixa 3–7 Hz) e entropia de amplitude de velocidade e pressão
Total	32		

Fonte: Elaborado pelos autores (2026).

2.3 Classificação

Para a tarefa de classificação binária entre pacientes com DP e controles saudáveis, cinco classificadores foram avaliados: SVM, Regressão Logística (RL), Análise Discriminante Linear (LDA), Naive Bayes (NB) e Random Forest (RF). O classificador SVM com *kernel* de função de base radial (RBF) busca o hiperplano de margem máxima que separa as classes no espaço de características, maximizando a distância entre os exemplos mais próximos de cada classe, denominados vetores de suporte [18]. O *kernel* RBF projeta implicitamente os dados em um espaço de alta dimensão, permitindo a separação de classes não linearmente separáveis no espaço original.

O algoritmo RL modela a probabilidade de pertencimento a cada classe por meio de uma função sigmoide aplicada a uma combinação linear das características, sendo um classificador interpretável e eficiente para problemas binários [19]. O modelo LDA busca a projeção linear que maximiza a separabilidade entre classes, assumindo distribuição gaussiana dos dados e covariâncias iguais entre os grupos [20]. O classificador NB gaussiano aplica o teorema de Bayes assumindo independência condicional entre as características, estimando a probabilidade de cada classe a partir de distribuições gaussianas ajustadas aos dados de treino [21]. O modelo RF constrói um conjunto de árvores de decisão treinadas sobre subamostras aleatórias dos dados e características, combinando suas predições por votação majoritária para obter um classificador robusto e resistente a *overfitting* [22].

2.4 Avaliação e Métricas

A avaliação dos modelos foi realizada por validação cruzada com K = 5 *folds* usando *StratifiedKFold*, que é uma estratégia para garantir a proporção entre as classes em cada *fold*. Como cada participante contribui com um único vetor de características (resultante da fusão das características do SST e do DST), a separação por participante é garantida naturalmente pela estrutura dos dados. Para mitigar o desequilíbrio entre classes (57 DP e 15 CO), aplicou-se a técnica de sobreamostragem SMOTE [23], que gera amostras sintéticas da classe minoritária por interpolação entre exemplos reais vizinhos. O SMOTE foi aplicado exclusivamente ao conjunto de treino de cada *fold*, garantindo que nenhuma amostra sintética fosse incluída no conjunto de teste e preservando a integridade da avaliação.

As métricas de avaliação empregadas foram: acurácia, definida pela equação (b), que mede a proporção de classificações corretas sobre o total de amostras; sensibilidade (*recall*), definida pela equação (c), que mede a proporção de pacientes com DP corretamente identificados (verdadeiros positivos); especificidade, definida pela equação (d), que mede a proporção de controles saudáveis corretamente identificados; e *F1-score*, definido pela equação (e), que representa a média harmônica entre precisão e sensibilidade. A AUC (Area Under the ROC Curve) também foi calculada como medida complementar da capacidade discriminativa do modelo. É importante mencionar que os valores reportados correspondem à média das métricas obtidas nos cinco *folds*.

$$acurácia = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad \text{(b)}$$

$$sensibilidade = \frac{TP}{(TP + FN)} \quad \text{(c)}$$

$$especificidade = \frac{TN}{(TN + FP)} \quad \text{(d)}$$

$$f1score = 2 \times \frac{(Precisão \times Sensibilidade)}{(Precisão + Sensibilidade)} \quad \text{(e)}$$

3 RESULTADOS

Os resultados dos cinco classificadores avaliados com StratifiedKFold (K = 5) são apresentados na Tabela 3. O Random Forest obteve o melhor desempenho geral, com acurácia de 91,71%, sensibilidade de 94,85%, *F1-Score* de 94,69% e AUC de 98,79%, sendo selecionado como modelo de referência.

Tabela 3. Desempenho dos classificadores avaliados

Classificadores	Acurácia	Sensibilidade	Especificidade	*F1-Score*	AUC
SVM	88,95%	93,03%	73,33%	92,95%	94,85%
Regressão Logística	79,24%	80,61%	73,33%	85,57%	91,82%
LDA	52,86%	54,55%	46,67%	64,16%	58,64%
Naive Bayes	76,19%	75,15%	80,00%	82,16%	84,19%
Random Forest	91,71%	94,85%	80,00%	94,69%	98,79%

Fonte: Elaborado pelos autores (2026).

Tabela 4. Desempenho do modelo Random Forest

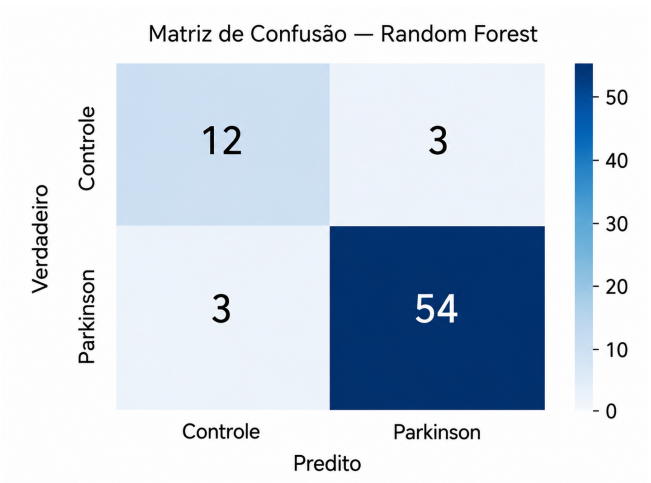
Nº do *Fold*	Acurácia	Sensibilidade	Especificidade	*F1-Score*	AUC
Fold 1	86,67%	83,33%	100,00%	90,91%	100,00%
Fold 2	93,33%	100,00%	66,67%	96,00%	100,00%
Fold 3	92,86%	90,91%	100,00%	95,24%	100,00%
Fold 4	92,86%	100,00%	66,67%	95,65%	96,97%
Fold 5	92,86%	100,00%	66,67%	95,65%	96,97%
Média dos Folds	91,71%	94,85%	80,00%	94,69%	98,79%

Fonte: Elaborado pelos autores (2026).

O Random Forest alcançou acurácia média de 91,71% e sensibilidade de 94,85%, indicando que o modelo identifica corretamente a grande maioria dos pacientes com DP, o que é clinicamente desejável em um contexto de triagem diagnóstica, em que falsos negativos (pacientes com DP classificados como saudáveis) são especialmente prejudiciais. A especificidade de 80,00%, menor do que a sensibilidade, reflete o desequilíbrio original entre as classes presentes na base de dados (57 pacientes com DP contra apenas 15 controles saudáveis), parcialmente mitigado pela aplicação do SMOTE no conjunto de treino. A variabilidade residual da especificidade entre *folds* decorre do número reduzido de controles no conjunto de teste, em que o SMOTE não é aplicado, de modo que o modelo permanece mais conservador ao classificar indivíduos como saudáveis, priorizando a detecção da doença. O detalhamento do desempenho por *fold* é apresentado na Tabela 4.

A análise individual dos *folds* (Tabela 4) evidencia desempenho consistente: a acurácia variou entre 86,67% (*Fold* 1) e 93,33% (*Fold* 2), com valores próximos da média de 91,71% nos cinco *folds*, o que sustenta o argumento de robustez da abordagem. A sensibilidade manteve-se elevada e estável em quatro dos cinco *folds* (entre 90,91% e 100,00%), confirmando que a detecção de pacientes com DP é estável independentemente da partição dos dados. A especificidade apresentou a maior variabilidade entre *folds*, oscilando de 66,67% (*Fold* 2) a 100,00% (*Fold* 3), instabilidade explicada pelo número reduzido de controles nos dados originais mitigado pelo uso do SMOTE. A matriz de confusão acumulada dos cinco *folds* (Figura 2) consolida esse padrão, evidenciando que os erros do modelo se concentram proporcionalmente em falsos positivos (controles classificados como DP) e não em falsos negativos, o que é um comportamento adequado para uma ferramenta de triagem, em que deixar de identificar um paciente é o desfecho mais crítico.

Figura 2. Matriz de Confusão do Random Forest



Fonte: Elaborado pelos autores (2026).

A AUC de 98,79% confirma que o modelo possui excelente capacidade de discriminação entre as classes em diferentes limiares de decisão. O *F1-Score* de 94,69% reflete o equilíbrio entre precisão e sensibilidade, indicando que, quando o modelo classifica um indivíduo como paciente com DP, essa classificação é majoritariamente correta.

Tabela 5. Comparação com a literatura

Trabalho	Entrada	Testes	Classificador	Acurácia
Wang et al. [12]	Imagens (HOG)	SST e DST	SVM	93,30%
Kamble et al. [13]	Sinais de sensores	SST e DST	RL, RF	91,60%
Thakur et al. [14]	Sinais de sensores	SST e DST	RBM + MLP	95,32%
Vaish [15]	Sinais de sensores	SST, DST e CMT	Random Forest	70,97%
Alniemi e Mahmood [9]	Imagens	Espiral	CNN	93,33%
Farhah [10]	Imagens	Espiral	InceptionV3	89,00%
Este trabalho	Sinais de sensores	SST + DST	Random Forest + SMOTE	91,71%

Fonte: Elaborado pelos autores (2026).

4 DISCUSSÃO

Diante dos resultados apresentados, na Tabela 5 é apresentada a comparação direta dos resultados obtidos neste trabalho com os principais estudos da literatura. Os quatro primeiros trabalhos [12-15] utilizam o mesmo conjunto de dados deste trabalho, enquanto Alniemi e Mahmood [9] e Farhah [10] empregam dados distintos de imagens de espirais, incluídos para ilustrar abordagens baseadas em imagens. Todos os trabalhos comparados realizam classificação binária (DP vs. CO), assim como o presente estudo.

Ao comparar com a literatura, é possível observar que o trabalho de Thakur *et al.* [14] alcança a maior acurácia (95,32%) utilizando RBM + MLP, um modelo de aprendizado profundo com maior complexidade computacional. Wang *et al.* [12], com SVM, obtêm 93,30%, enquanto Kamble *et al.* [13], com modelos mais simples como Regressão Logística e Random Forest, chegam a 91,60%. Vaish [15], com Random Forest, apresenta a menor acurácia do grupo comparado (70,97%).

O modelo proposto neste trabalho alcançou 91,71% de acurácia, posicionando-se de forma equiparável aos trabalhos comparados. Wang *et al.* [12] utilizam imagens das espirais como entrada, extraindo características visuais via HOG, descartando as informações cinemáticas registradas pelos sensores da mesa digitalizadora. Kamble *et al.* [13] e Thakur *et al.* [14] exploram sinais temporais, porém analisam SST e DST de forma independente, gerando um vetor de características por teste. Nenhum dos trabalhos comparados aplica técnicas de balanceamento de classes. Este trabalho se diferencia ao propor a fusão dos dois testes em um único vetor por participante, capturando a complementaridade entre as modalidades, e ao aplicar o SMOTE internamente no conjunto de treino a cada *fold* para mitigar o efeito do desequilíbrio de classes. Esses aspectos metodológicos contribuem para estimativas de desempenho mais robustas e para uma representação mais completa do perfil motor do paciente, o que reforça o potencial da abordagem proposta como ferramenta de apoio ao diagnóstico digital da DP.

Cabe destacar que abordagens baseadas em imagens de espirais, embora promissoras, dependem da conversão dos sinais em representações visuais, descartando a informação temporal e cinemática registrada pelos sensores. Alniemi e Mahmood [9], com uma rede neural convolucional, e Farhah [10], com modelos de *transfer learning*, alcançaram, respectivamente, 93,33% e 89% de acurácia em conjuntos de imagens de espirais. A abordagem proposta neste trabalho, ao explorar diretamente as séries temporais dos sensores, preserva informações dinâmicas do movimento que são perdidas na conversão para imagem, mantendo desempenho competitivo com menor complexidade computacional do que modelos de aprendizado profundo.

Além da questão dos dados de entrada, também é relevante o custo computacional das diferentes abordagens. O modelo proposto utiliza um classificador clássico sobre um vetor de características tabulares, o que tende a exigir menor poder de processamento e tempo de treinamento do que arquiteturas de aprendizado profundo como as empregadas por Alniemi e Mahmood [9] e Farhah [10], geralmente dependentes de unidade de processamento gráfico (GPU). Entretanto, nenhum dos trabalhos comparados (incluindo este) reporta métricas de tempo de treino, tempo de inferência ou tamanho de modelo, o que impede uma comparação rigorosa e quantitativa de custo computacional entre as abordagens.

Por ser não invasiva e de baixo custo tecnológico, a abordagem tem potencial para triagem e monitoramento remoto em sistemas de saúde com recursos limitados, como o SUS [8]. Essa aplicabilidade, porém, ainda depende de validação em amostras maiores e mais diversas, dado o tamanho reduzido e o desbalanceamento dos dados utilizados.

5 CONCLUSÕES

Este trabalho propôs e avaliou uma abordagem de aprendizado de máquina para a classificação binária da Doença de Parkinson a partir da fusão, em um único

vetor por participante, de características dos testes SST e DST. O modelo Random Forest combinado à sobreamostragem SMOTE atingiu acurácia de 91,71%, sensibilidade de 94,85% e *F1-Score* de 94,69%, com desempenho consistente entre os cinco *folds*, respondendo ao objetivo proposto e sugerindo o potencial da abordagem como ferramenta de apoio ao diagnóstico digital da DP.

As principais limitações deste trabalho incluem a pouca quantidade de dados e o desequilíbrio entre classes no conjunto de dados, que impacta a especificidade do modelo e constitui uma fonte de viés algorítmico, uma vez que a sub-representação dos controles tende a enviesar o classificador em favor da classe majoritária. Adicionalmente, este trabalho não avaliou o custo computacional do modelo proposto, tampouco o comparou de forma quantitativa com os das abordagens da literatura, o que constitui uma limitação e uma oportunidade para trabalhos futuros. Assim, como trabalhos futuros, propõe-se a extensão para bases de dados com maior número de controles saudáveis, a exploração de técnicas adicionais de balanceamento de dados, a investigação de modelos de aprendizado profundo para classificação de séries temporais e a avaliação da capacidade de quantificação da severidade da DP a partir dos mesmos sinais.

Contribuição dos autores. Geovanna Cunha Andrade Silva contribuiu para a concepção e o delineamento do estudo, desenvolvimento de software, análise e interpretação dos dados, elaboração de figuras e redação da versão inicial do manuscrito. Giordana de Farias Franco Bueno Bucci contribuiu para a concepção e o delineamento do estudo, desenvolvimento de software, validação, análise e interpretação dos dados, e redação da versão inicial do manuscrito. Fabrizzio Soares contribuiu para a análise formal, supervisão da pesquisa e revisão crítica do manuscrito. Juliana Paula Felix contribuiu para a concepção e o delineamento do estudo, definição da metodologia, análise formal, supervisão da pesquisa e revisão crítica do manuscrito. Todos os autores leram e aprovaram a versão final do manuscrito.

Agradecimentos. O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES), Código de Financiamento 001, e Edital Nº 30/2022 – PDPG – Solidariedade Acadêmica.

Financiamento. Esta pesquisa não recebeu financiamento externo.

Disponibilidade de dados e materiais. Este estudo utilizou um conjunto de dados público, denominado Parkinson Disease Spiral Drawings Using Digitized Graphics Tablet, disponibilizado no repositório UCI Machine Learning Repository. O conjunto de dados está disponível em: https://archive.ics.uci.edu/dataset/395/parkinson+disease+spiral+drawings+using+digitized+graphics+tablet.

Conflitos de interesse. Os autores declaram não possuir conflitos de interesse.

Declaração de uso de IA. O sistema de Inteligência Artificial Claude (Anthropic) foi empregado neste trabalho como ferramenta de apoio técnico para redação e revisão de texto, compreendendo análise de clareza, gramática, estilo e organização estrutural do manuscrito. Não foi utilizado para geração de dados, resultados, tabelas ou conclusões, nem para substituição da autoria intelectual dos autores.

REFERÊNCIAS

1. Tysnes OB, Storstein A. Epidemiology of Parkinson's disease. J Neural Transm (Vienna). 2017;124(8):901-5. doi: https://doi.org/10.1007/s00702-017-1686-y
2. Jankovic J. Parkinson's disease: clinical features and diagnosis. J Neurol Neurosurg Psychiatry. 2008;79(4):368-76. doi: https://doi.org/10.1136/jnnp.2007.131045
3. Armstrong MJ, Okun MS. Diagnosis and Treatment of Parkinson Disease: A Review. JAMA. 2020;323(6):548-60. doi: https://doi.org/10.1001/jama.2019.22360
4. Goetz CG, Tilley BC, Shaftman SR, Stebbins GT, Fahn S, Martinez-Martin P, et al. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results. Mov Disord. 2008;23(15):2129-70. doi: https://doi.org/10.1002/mds.22340
5. Kalia LV, Lang AE. Parkinson's disease. Lancet. 2015;386(9996):896-912. doi: https://doi.org/10.1016/S0140-6736(14)61393-3
6. Impedovo D, Pirlo G, Vessio G. Dynamic Handwriting Analysis for Supporting Earlier Parkinson's Disease Diagnosis. Information. 2018;9(10):247. doi: https://doi.org/10.3390/info9100247
7. Aouraghe I, Khaissidi G, Mrabti M. A literature review of online handwriting analysis to detect Parkinson's disease at an early stage. Multimed Tools Appl. 2023;82(8):11923-48. doi: https://doi.org/10.1007/s11042-022-13759-2
8. Brasil. Ministério da Saúde. Sistema Único de Saúde (SUS) [Internet]. Brasília: Ministério da Saúde [citado 2026 maio 16]. Disponível em: https://www.gov.br/saude/pt-br/sus
9. Alniemi O, F. Mahmood H. Convolutional neural network for the detection of Parkinson disease based on hand-draw spiral images. IJEECS. 2023;30(1):267. doi: https://doi.org/10.11591/ijeecs.v30.i1.pp267-275
10. Farhah N. Utilizing deep learning models in an intelligent spiral drawing classification system for Parkinson's disease classification. Front Med (Lausanne). 2024;11:1453743. doi: https://doi.org/10.3389/fmed.2024.1453743
11. Isenkul M, Sakar B, Kursun O. Improved Spiral Test Using Digitized Graphics Tablet for Monitoring Parkinson's Disease. In: Proceedings of the 2nd International Conference on e-Health and Telemedicine (ICEHTM 2014). 2014. p. 171-5. Conjunto de dados disponível em: https://doi.org/10.24432/C5Q01S

12. Wang L, Liu H, Chen X, Li Y. A Lightweight Parkinson's Disease Diagnostic Method Based on Hand-Drawn Spiral Lines. In: Proceedings of the 2024 4th International Conference on Bioinformatics and Intelligent Computing. New York, NY, USA: ACM; 2024. p. 217-20. doi: https://doi.org/10.1145/3665689.3665726

13. Kamble M, Shrivastava P, Jain M. Digitized spiral drawing classification for Parkinson's disease diagnosis. Measurement: Sensors. 2021;16:100047. doi: https://doi.org/10.1016/j.measen.2021.100047

14. Thakur M, Dhanalakshmi S, Kuresan H, Senthil R, Narayanamoorthi R, Lai KW. Automated restricted Boltzmann machine classifier for early diagnosis of Parkinson's disease using digitized spiral drawings. J Ambient Intell Human Comput. 2023;14(1):175-89. doi: https://doi.org/10.1007/s12652-022-04361-3

15. Vaish A. A Machine Learning Approach for Early Identification of Prodromal Parkinson's Disease. Cureus. 2024;16(6):e63240. doi: https://doi.org/10.7759/cureus.63240

16. Savitzky A, Golay MJ. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. Anal. Chem. 1964;36(8):1627-39. doi: https://doi.org/10.1021/ac60214a047

17. Bhatia KP, Bain P, Bajaj N, Elble RJ, Hallett M, Louis ED, et al. Consensus Statement on the classification of tremors. from the task force on tremor of the International Parkinson and Movement Disorder Society. Mov Disord. 2018;33(1):75-87. doi: https://doi.org/10.1002/mds.27121

18. Cortes C, Vapnik V. Support-Vector Networks. Machine Learning. 1995;20(3):273-97. doi: https://doi.org/10.1023/a:1022627411411

19. Cox DR. The Regression Analysis of Binary Sequences. Journal of the Royal Statistical Society Series B: Statistical Methodology. 1958;20(2):215-32. doi: https://doi.org/10.1111/j.2517-6161.1958.tb00292.x

20. Fisher RA. The use of multiple measurements in taxonomic problems. Annals of Eugenics. 1936;7(2):179-88. doi: https://doi.org/10.1111/j.1469-1809.1936.tb02137.x

21. Domingos P, Pazzani M. On the Optimality of the Simple Bayesian Classifier under Zero-One Loss. Machine Learning. 1997;29(2-3):103-30. doi: https://doi.org/10.1023/a:1007413511361

22. Breiman L. Random Forests. Machine Learning. 2001;45(1):5-32. doi: https://doi.org/10.1023/a:1010933404324

23. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: Synthetic Minority Over-sampling Technique. jair. 2002;16:321-57. doi: https://doi.org/10.1613/jair.953

NOTAS EDITORIAIS

O Journal of Health Informatics é uma publicação científica mantida pela Sociedade Brasileira de Informática em Saúde, associação científica sem fins lucrativos dedicada ao desenvolvimento da Informática em Saúde e da Saúde Digital no Brasil. O processo editorial é conduzido com independência científica, sem interferência institucional, comercial ou de patrocinadores, e segue as boas práticas internacionais de ética em publicação científica.

Este artigo foi avaliado por pares, em sistema duplo-cego, por especialistas da área. A equipe editorial usou ferramentas de inteligência artificial, com supervisão humana, na triagem inicial, na conferência do texto, das referências e dos números e, quando necessário, no redesenho de figuras, quadros e tabelas, aprovado pelos autores na prova. As ferramentas e as salvaguardas estão na Política de uso de inteligência artificial da revista. As decisões editoriais foram tomadas por pessoas.

SBIS Sociedade Brasileira de Informática em Saúde