

ARTIGO DE PESQUISA/RESEARCH PAPER

Transformer: O Poder da Atenção no Auxílio ao Diagnóstico de Múltiplas Doenças Neurodegenerativas

Transformer: The Power of Attention to Support the Diagnosis of Multiple Neurodegenerative Diseases

Giordana de Farias F. B. Bucci  [Universidade Federal de Goiás | giordanabucci@discente.ufg.br]

Ana Luísa de Bastos Chagas  [Universidade Federal de Goiás | analuisa23@discente.ufg.br]

Juliana Paula Félix  [Pontifícia Universidade Católica de Goiás e Universidade Federal de Goiás | julianafelix@ufg.br]

Hugo A. D. do Nascimento  [Universidade Federal de Goiás | hadn@ufg.br]

Fabrizio Soares  [Universidade Federal de Goiás | fabrizio@ufg.br]

 Instituto de Informática (INF) - UFG, Alameda Palmeiras, Quadra D, Câmpus Samambaia, 74690-900, Goiânia, Goiás, Brasil.

Resumo. As doenças neurodegenerativas (DNDs) causam, entre outros sintomas, instabilidade da marcha, possuem natureza incurável e apresentam um longo e difícil processo de diagnóstico. Por isso, diversos estudos investigam a marcha por meio de modelos de inteligência artificial como uma alternativa para auxiliar no diagnóstico dessas doenças. Este trabalho apresenta os principais resultados obtidos durante a iniciação científica com um método inovador de detecção de DNDs utilizando um *Encoder-Only Transformer* associado a análise de marcha em uma tarefa de multi-classificação. Os resultados alcançam valores altos de acurácia e indicam uma alternativa promissora de identificação de DNDs.

Abstract. Neurodegenerative diseases (NDDs) cause, among other symptoms, gait instability, have an incurable nature, and present a long and challenging diagnostic process. For this reason, several studies have investigated gait using artificial intelligence models as an alternative to assist in the diagnosis of these diseases. This study presents the main results obtained during an undergraduate research project using an innovative method for detecting NDDs, using an Encoder-Only Transformer combined with gait analysis in a multi-class classification task. The results achieve high accuracy values and indicate a promising alternative for NDD identification.

Palavras-chave: Doenças neurodegenerativas, diagnóstico, transformer, marcha

Keywords: Neurodegenerative diseases, diagnosis, transformer, gait

Recebido/Received: 22 May 2025 • Aceito/Accepted: 24 June 2025 • Publicado/Published: 17 July 2025

1 Introdução

As doenças neurodegenerativas (DNDs), tais como Esclerose Lateral Amiotrófica (ELA), Doença de Parkinson (DP) e Doença de Huntington (DH), afetam diretamente o sistema nervoso e causam, inclusive, movimentos involuntários, fraqueza dos músculos e marcha instável [Maragakis and Rothstein, 2006]. Não há cura para as DNDs [Heemels, 2016] e, por falta de testes definitivos da doença [Erkkinen *et al.*, 2018], o diagnóstico delas é feito por um acompanhamento da progressão dos sintomas ao longo do tempo [Mayeux, 2003]. Isso resulta em reconhecimento tardio das doenças, retardando no início do tratamento e agravamento dos sintomas.

A análise de marcha é um atrativo método alternativo para auxiliar no diagnóstico precoce de DNDs, já que alterações na caminhada ocorrem dentre os primeiros sintomas e se manifesta de maneiras diferentes em cada DND [Hausdorff *et al.*, 2000]. Na literatura, o estudo da marcha é comumente encontrado em conjunto com modelos de aprendizado de máquina clássicos [Rao *et al.*, 2025] ou profundo, como *Gated Recurrent Unit* (GRU) [Zhao *et al.*, 2018], Redes Neurais Convolucionais (CNN) [Erdaş *et al.*, 2023; Erdaş and Sümer, 2024] e combinações de *Long Short-Term Memory* com CNN [Erdaş *et al.*, 2021; Amooei *et al.*, 2023].

Enquanto esses trabalhos se baseiam em arquiteturas tra-

dicionais de aprendizado de máquina clássico ou profundo, estudos recentes demonstram o potencial dos modelos baseados em *transformers* para análise de dados sequenciais [Agarwal *et al.*, 2020]. O notável avanço refere-se à capacidade do Transformer [Vaswani *et al.*, 2017] em processar e capturar relações contextuais entre os elementos sequenciais, além de “aprender” as dependências à longo prazo utilizando principalmente mecanismos de atenção. Isso é computacionalmente vantajoso, já que o modelo não percorre ou itera sobre os dados, como fazem as redes neurais convolucionais (CNNs), as redes neurais recorrentes (RNNs), as redes *Gated Recurrent Unit* (GRUs) e as redes *Long Short-Term Memory* (LSTMs).

Apesar dessa inovação em análises de séries temporais com *transformer*, apenas alguns trabalhos implementam esse modelo no domínio de doenças neurodegenerativas e focam somente na Doença de Parkinson. Nguyen *et al.* [2022] abordaram o diagnóstico de Parkinson como uma tarefa de classificação binária, enquanto Sun and Zhang [2022] e Naimi *et al.* [2023] abordaram a detecção da severidade da doença de Parkinson como uma tarefa de classificação multi-classe.

Nesse sentido, o principal objetivo deste trabalho de iniciação científica foi investigar métodos e técnicas de aprendizado profundo para auxiliar no diagnóstico de DNDs. Neste documento, apresentam-se os principais resultados, que con-

sistem na proposta de um novo método para identificação de DH, DP, ELA e pacientes de controle (CO) em uma tarefa de multi-classificação. O método utiliza análise de marcha associada a um *Encoder-Only Transformer*, arquitetura *transformer* baseada somente em codificador.

O restante deste artigo está organizado da seguinte forma: a Seção 2 descreve sobre os trabalhos relacionados; a Seção 3 detalha sobre a base matemática que fundamenta o modelo *transformer*; a Seção 4 apresenta a metodologia proposta, detalhando a base de dados, o pré-processamento dos dados, o modelo Encoder-Only Transformer implementado e a forma de avaliação dos resultados; a Seção 5 discute sobre os resultados obtidos; por fim, a Seção 6 discorre sobre as conclusões.

2 Trabalhos Relacionados

Em relação à tarefa de classificação multi-classe para identificar Esclerose Lateral Amiotrófica (ELA), Doença de Parkinson (DP), Doença de Huntington (DH) e grupo de controle (CO), diversos estudos investigaram métodos de aprendizado profundo.

Zhao *et al.* [2018] desenvolveram um modelo GRU de duas camadas para analisar os atributos da dinâmica da marcha, que foram divididos em janelas de comprimento fixo de 20 ciclos completos da marcha, e relataram uma acurácia média de 95,35%. Ning *et al.* [2018] propuseram uma combinação de uma CNN unidimensional e camadas LSTM para analisar os sinais obtidos da marcha segmentados em janelas de 2 segundos com sobreposição de 50%, o que resultou em 99,50% de acurácia.

Erdaş *et al.* [2021] apresentaram uma abordagem que converte os atributos da dinâmica de marcha em *QR-code*, que são então analisados por um modelo que integra as abordagens de CNN e LSTM, alcançando 89,44% de acurácia. Fraiwan and Hassain [2021] abordaram a classificação multi-classe com o modelo AdaBoost para analisar características estatísticas temporais simples – raiz quadrada média (RMS), variância, curtose e assimetria — extraídas manualmente dos sinais brutos de marcha, resultando em 99,17% de acurácia. Amooei *et al.* [2023] desenvolveram um método de classificação utilizando uma combinação de CNN, transformada *wavelet* e LSTM para analisar os sinais de marcha divididos em janelas deslizantes com 50% de sobreposição e transformados em espectrogramas RGB, apontando 95,37% de acurácia.

Erdaş *et al.* [2023] compararam métodos de aprendizado de máquina e aprendizado profundo para a tarefa multi-classe, sendo o melhor resultado de 68,11% utilizando CNN unidimensional sobre os atributos da dinâmica de marcha. Faisal *et al.* [2023] propuseram uma arquitetura de rede de aprendizado profundo chamada *NDDNet*, para processar simultaneamente o sinal bruto obtido a partir da marcha, as fases da marcha e as características extraídas desses sinais, obtendo um resultado geral de 83,00% de acurácia. Por fim, Erdaş and Sümer [2024] simplificaram sua abordagem anterior baseada em *QR-code*, utilizando apenas uma arquitetura de redes neurais convolucionais (CNN) como classificador, o que resultou em uma acurácia de 84,65%.

3 Fundamentação Teórica

3.1 Mecanismo de Atenção

Os mecanismos de atenção são técnicas utilizadas para aprimorar a capacidade de um modelo em focar nas partes mais relevantes dos dados de entrada, dado um contexto específico. Esses mecanismos baseiam-se em três componentes principais: *query*, que representa a informação que o modelo está buscando; *key*, que corresponde à a informação armazenada em cada elemento de entrada; e *value*, que contém os dados reais associados a cada elemento de entrada. Com esses componentes, calcula-se a similaridade entre todos os elementos de entrada, independentemente de sua posição na sequência e, como resultado, atribui-se dinamicamente pesos às diferentes partes da entrada, o que representa sua importância para a saída atual.

Matematicamente, dada uma sequência de entrada de n elementos, representada por $X = (x_1, x_2, \dots, x_n)$, com $x_i \in \mathbb{R}^{d_k}$, onde d_k é a dimensão de cada elemento, os mecanismos de atenção podem ser expressos pela Equação (Eq.) 1. As matrizes Q , K e V representam *queries*, *keys* e *values*, com dimensões d_k , d_k e d_v , respectivamente. O produto interno QK^T calcula a similaridade entre os elementos de entrada, e o fator de escala $1/\sqrt{d_k}$ ajuda a estabilizar o cálculo para valores grandes de d_k . Os valores resultantes são normalizados usando a função softmax (Eq. 2) para calcular os pesos finais de atenção na combinação dos valores.

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V. \quad (1)$$

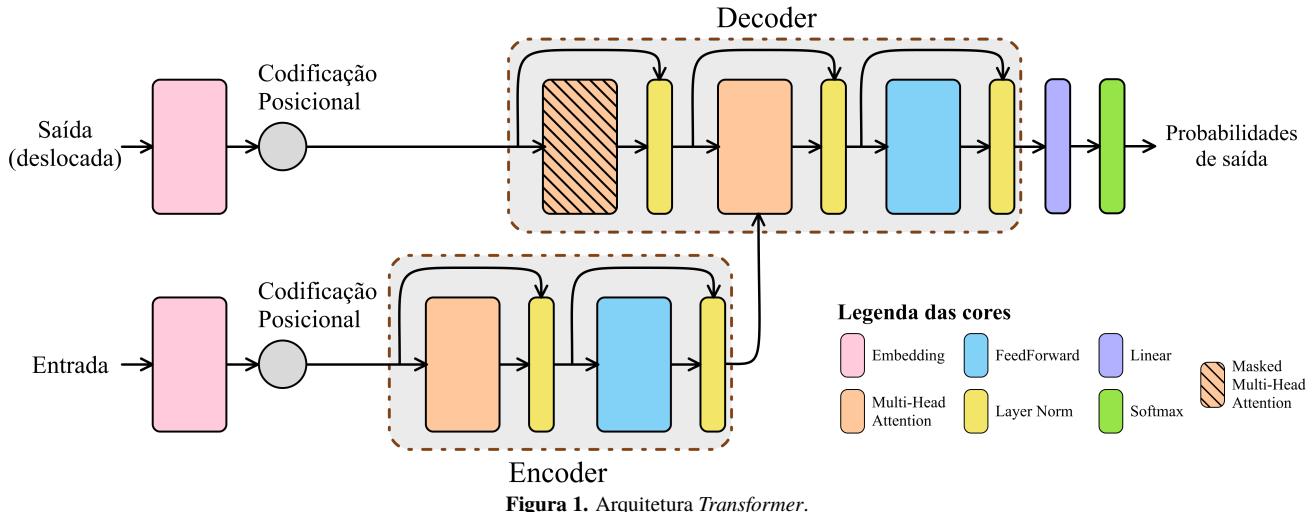
$$\text{softmax}(x_i) = \frac{e^{x_i}}{\sum_{j=1}^n e^{x_j}}. \quad (2)$$

3.2 Transformer

Introduzida por Vaswani *et al.* [2017], a arquitetura geral do *transformer*, como mostrada na Figura 1, é composta por pilhas de uma estrutura *encoder-decoder* e algumas funções adicionais, como *embeddings*, codificação posicional e *softmax*.

O *encoder* (codificador) consiste em uma pilha de n camadas idênticas, cada uma contendo duas subcamadas. A primeira subcamada é um mecanismo de atenção (*Multi-Head Attention*) e a segunda é uma rede totalmente conectada *FeedForward*. Conexões residuais seguidas de camadas de normalização (*Layer Norm*) são aplicadas entre esses componentes. Para suportar as conexões residuais, todos as camadas do modelo, incluindo as camadas de *embedding*, produzem vetores de saída com dimensão d_{model} .

O *decoder* (decodificador) consiste em uma pilha de n camadas idênticas, cada uma com três subcamadas. A primeira subcamada é um mecanismo de atenção com máscara (*Masked Multi-Head Attention*) que impede determinadas posições de verificarem posições subsequentes. A segunda subcamada consiste em uma atenção multi-cabeças (*Multi-Head Attention*) que considera a saída da pilha de *encoder*, enquanto a terceira subcamada é uma subcamada de alimentação direta (*FeedForward*). Essa máscara na primeira subcamada, combinada com o deslocamento dos *embeddings* de saída em uma posição a direita, garante que as previsões



para uma posição (i) dependam apenas das saídas conhecidas em posições anteriores ($< i$). Similar ao *encoder*, conexões residuais seguidas de camadas de normalização (*Layer Norm*) são aplicadas entre esses componentes.

As subcamadas de *Multi-Head Attention*, conforme mostrado nas Eq. 3 e Eq. 4, concatenam os resultados da execução de h funções de atenção em paralelo, usando *queries*, *keys* e *values* projetados linearmente (com projeções aprendidas diferentes) para dimensões d_k , d_k e d_v , respectivamente, onde $d_k = d_v = (d_{model}/h)$.

$$\text{Attention}_{MH}(Q, K, V) = \text{concat}(\text{head}_1, \dots, \text{head}_h)W^O, \quad (3)$$

$$\text{head}_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V), \quad (4)$$

onde $W^O \in \mathbb{R}^{hd_v \times d_{model}}$, $W_i^Q \in \mathbb{R}^{d_{model} \times d_k}$, $W_i^K \in \mathbb{R}^{d_{model} \times d_k}$ e $W_i^V \in \mathbb{R}^{d_{model} \times d_v}$ são as matrizes de parâmetros para as projeções.

As subcamadas de *FeedForward*, representadas pela Eq. 5, são compostas por duas transformações lineares com a função de ativação ReLU (Eq. 6) entre elas, e são aplicadas idêntica e independentemente a cada posição da sequência. Ao introduzir não linearidade e aplicar expansão e compressão dimensional, a *FeedForward* permite que o *transformer* aprenda relações mais complexas, obtenha representações mais abstratas e foque em transformações locais.

$$\text{FeedForward}(x) = \text{ReLU}(xW_1 + b_1)W_2 + b_2. \quad (5)$$

$$\text{ReLU}(x) = \max(0, x). \quad (6)$$

As camadas de codificação posicional agregam (somam aos *embeddings*) informações sobre a posição relativa ou absoluta dos elementos na sequência, permitindo que o modelo utilize a informação de ordem da sequência, mesmo sem apresentar qualquer convolução ou recorrência. Por fim, as camadas *embedding* são usadas para converter os dados de entrada e saída em vetores de dimensão d_{model} , e a função *softmax* é utilizada para converter a saída do *decoder* em probabilidades de saída.

4 Metodologia

A Figura 2 apresenta de modo geral a metodologia discutida neste trabalho. As seções seguintes detalham cada uma das etapas analisadas.

4.1 Base de Dados

A base de dados pública *Gait in Neurodegenerative Disease Database* (GaitNDD), fornecida por Hausdorff *et al.* [1997, 2000] e disponível no sítio PhysioNet¹ foi utilizada no presente trabalho. A base contém dados de até 5 minutos da marcha de 20 indivíduos diagnosticados com Doença de Huntington, 15 com Doença de Parkinson, 13 com Esclerose Lateral Amiotrófica e 16 indivíduos saudáveis de controle, totalizando 64 indivíduos.

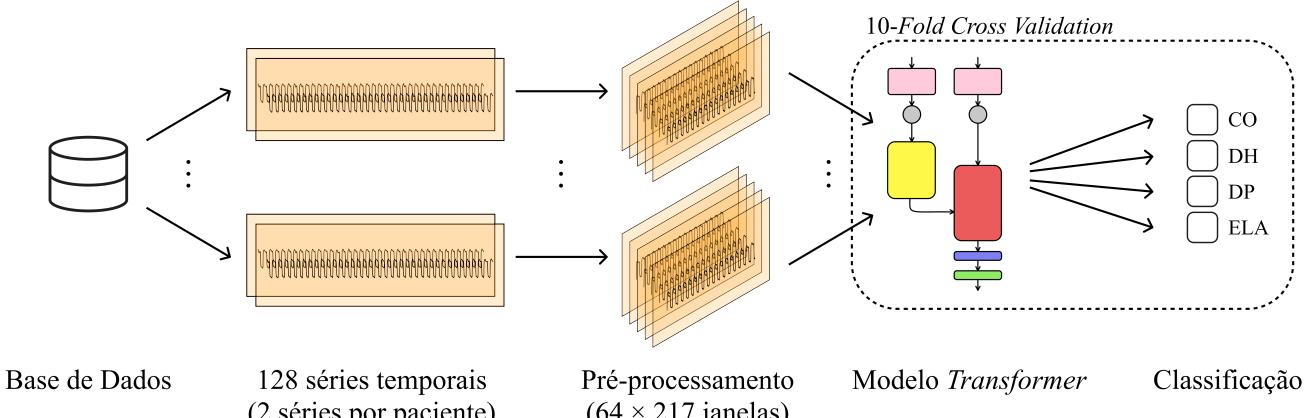
Esses sinais de marcha foram coletados com uma frequência de 300 Hz de maneira não invasiva, por meio de sensores sensitivos de força [Hausdorff *et al.*, 1995] acoplados a ambos os pés (esquerdo e direito) dos participantes. Com isso, foi solicitado a cada voluntário que caminhasse durante 5 minutos, sem apoio e em ritmo autodeterminado, em um corredor de 77 m de comprimento. Isso resultou em duas séries temporais (uma para cada pé) de 90 mil pontos de dados (300 Hz \times 5 \times 60 segundos).

4.2 Pré-processamento

A fim de mitigar os efeitos de inicialização reportados pelos autores da base de dados [Hausdorff *et al.*, 1997, 2000], os dados coletados nos primeiros 20 segundos de cada série temporal foram descartados. Em seguida, os valores faltantes (*NaN values*) foram tratados, sendo substituídos pelo próximo valor válido na série e, ainda, qualquer valor *NaN* restante foi substituído pelo valor válido anterior.

Posteriormente, utilizou-se uma técnica de janelamento, que segmenta os sinais em janelas deslizantes de 10 segundos com passo (*step*) de 1 segundo para maximizar a quantidade de dados provenientes de um mesmo participante. Os valores das janelas foram normalizados para valores entre 0 e 3,5 [Hausdorff *et al.*, 1995]. Assim, cada um dos 64 pacientes passa a ser representado por dois conjuntos de 271 janelas, cada uma com 3.000 pontos de dados. Portanto, os dados da

¹<https://physionet.org/content/gaitndd/1.0.0>

**Figura 2.** Fluxograma da metodologia proposta.

rede *transformer* são representados por um vetor tridimensional ($64 \times 271, 3000, 2$).

4.3 Classificação

Para realizar a tarefa de multi-classificação, foi utilizada uma arquitetura *transformer* baseada apenas em codificador, nomeado *Encoder-Only Transformer*. O modelo projeta linearmente a dimensão de entrada 2 (representando os sinais de marcha obtidos dos dois pés) em um espaço de *embedding* de $d_{model} = 64$.

Essa projeção é aplicada para capturar o significado dos elementos em um espaço de maior dimensão e para supor tar todas as conexões residuais do *transformer*. Em seguida, aplica-se aos dados uma codificação posicional (*positional encoding*), que alterna entre funções seno e cosseno de diferentes frequências para fornecer informações posicionais exclusivas para cada elemento na sequência de entrada, detalhada pelas Equações (Eq.) 7 e 8. Esse valores posicionais são incorporados aos *embeddings* de entrada para garantir que o modelo possa utilizá-la.

$$PE(pos, i) = \begin{cases} \text{sen}(pos \times div_{term}(i)), & \text{se } i \text{ é par.} \\ \text{cos}(pos \times div_{term}(i)), & \text{se } i \text{ é ímpar.} \end{cases} \quad (7)$$

$$div_{term}(i) = \exp\left(\frac{-i \times \ln 10.000}{d_{model}}\right). \quad (8)$$

Os dados incorporados e codificados são subsequentemente processados por uma pilha de $n = 2$ camadas do codificador. Dentro de cada uma dessas camadas, a subcamada de *multi-head self-attention* opera com $h = 4$ cabeças de atenção, e a subcamada *feedforward* possui dimensionalidade de 2.048. O *dropout*, configurado para 0,1, é aplicado após o mecanismo de atenção e em cada etapa da *feedforward* para mitigar o *overfitting*.

Após o codificador, a representação da sequência é reduzida a um único vetor por meio de uma operação de média (*mean pooling*) ao longo da dimensão temporal, condensando as informações temporais. Por fim, a representação reduzida passa por uma camada de normalização seguida por uma camada linear, resultando nos *logits* para a tarefa de classificação de quatro classes.

O treinamento foi realizado com 50 épocas, com entropia cruzada para a função de perda e Adam para a função do otimizador, com uma taxa de aprendizado de 0,001.

4.4 Avaliação

A validação cruzada com 10 folds (*10-Fold Cross-Validation*) foi utilizada para avaliar o desempenho do modelo. Para uma avaliação numérica, utilizou-se a acurácia (Eq. 9), que indica a taxa de classificações corretas considerando todas as amostras e *folds*, enquanto a matriz de confusão final, que sumariza as matrizes de confusão de cada *fold*, foi utilizada para uma avaliação visual. O código para classificação e avaliação foi desenvolvido em Python 3.12.3, utilizando as bibliotecas PyTorch e Scikit-Learn.

$$\text{Acurácia (Acc.)} = \frac{\text{predições corretas}}{\text{todas as predições}}. \quad (9)$$

5 Resultados e Discussão

O modelo *Encoder-Only Transformer* proposto alcançou uma acurácia média final de 99,47% na tarefa de classificação multi-classe para identificar os indivíduos como pertencentes a uma das quatro classes – Esclerose Lateral Amiotrófica (ELA), Doença de Parkinson (DP), Doença de Huntington (DH) ou controle (CO). Essa alta acurácia destaca a eficácia da abordagem baseada em *transformers* na distinção entre diferentes doenças neurodegenerativas com base em dados de marcha.

Para uma análise mais detalhada dos resultados da multi-classificação, uma matriz de confusão foi gerada para cada *fold*. A Tabela 1 apresenta a matriz de confusão final para os 10 *folds*, obtida pela soma das matrizes individuais de cada *fold* e normalizada para cada classe. Essa matriz agregada exibe uma forte diagonal principal, indicando que o modelo classifica corretamente a maioria dos indivíduos em suas respectivas categorias, com mínimas classificações incorretas.

A Tabela 2 apresenta uma comparação com outros estudos na literatura que classificam doenças neurodegenerativas em tarefas de multi-classificação usando a base de dados GaitNDD. Os resultados mostram que o modelo *Encoder-Only Transformer* apresenta ótimo desempenho para classificação multi-classe de doenças neurodegenerativas. Comparado aos métodos existentes na literatura, o modelo apresentado neste artigo obteve um desempenho similar ao modelo de CNN + LSTM de Ning et al. [2018] e superior aos métodos de *deep learning* tradicionais. Além disso, é importante enfatizar que o presente estudo também apresenta resultados mais robustos, considerando o uso da validação cruzada de

Tabela 1. Matriz de confusão que representa a acurácia de cada classe.

	Classe Predita				
	Esclerose Lateral Amiotrófica (ELA)	Doença de Huntington (DH)	Doença de Parkinson (DP)	Controle (CO)	
Classe Real	ELA	99,35%	0,37%	0,28%	0,00%
	DH	0,09%	99,48%	0,24%	0,18%
	DP	0,34%	0,15%	99,34%	0,17%
	CO	0,05%	0,09%	0,18%	99,68%

10 *folds*, o que proporciona uma avaliação mais confiável do desempenho do modelo quando comparado com o método mais comum, mas menos rigoroso, de *hold-out*, adotado pela maioria dos trabalhos relacionados.

Esses achados ressaltam a importância da inovação na computação aplicada à saúde referente ao uso de *transformers*. A alta eficácia e competitividade dessa abordagem na distinção entre diferentes doenças neurodegenerativas por meio da análise de marcha evidenciam seu potencial como uma alternativa superior aos modelos tradicionais de aprendizado profundo. Além disso, essa inovação pode permitir uma identificação de padrões sutis da doença, possibilitando diagnóstico precoce, intervenções e tratamentos antecipados, retardando na progressão da doença e, consequentemente, melhora na qualidade de vida dos pacientes.

Tabela 2. Comparação das acurárias dos métodos apresentados na literatura para uma multi-classificação das doenças neurodegenerativas.

Trabalho	Modelo	Acurácia
Zhao <i>et al.</i> [2018]	GRU	95,35%
Ning <i>et al.</i> [2018]	CNN + LSTM	99,50%
Erdaş <i>et al.</i> [2021]	CNN + LSTM	89,44%
Fraiwan and Hassanin [2021]	Adaboost	99,17%
Amooei <i>et al.</i> [2023]	CNN + LSTM e Transformada Wavelet	95,37%
Erdaş <i>et al.</i> [2023]	1d CNN	68,11%
Faisal <i>et al.</i> [2023]	NDDNet	83,00%
Erdaş and Sümer [2024]	CNN	84,65%
Este trabalho	Encoder-Only Transformer	99,47%

Gated Recurrent Unit (GRU), Rede Neural Convolucionial (CNN), Long Short-Term Memory (LSTM)

6 Conclusão

Este trabalho propôs uma nova abordagem para a detecção de doenças neurodegenerativas utilizando o *Encoder-Only Transformer*, um modelo de arquitetura baseada em *transformer*. Para atingir esse objetivo, foram analisadas séries temporais obtidas dos pés esquerdo e direito de indivíduos diagnosticados com Esclerose Lateral Amiotrófica, Doença de Parkinson, Doença de Huntington e de indivíduos de controle, durante uma caminhada de 5 minutos sem apoio. Os sinais obtidos

de cada pé foram segmentados em janelas menores e processados pelo modelo *transformer* para realizar uma tarefa de multi-classificação em quatro categorias (DH, DP, ELA, CO).

O método proposto obteve uma acurácia de 99,47%, evidenciando desempenho superior em relação a modelos tradicionais de aprendizado profundo presentes na literatura. Esses resultados indicam que o uso da arquitetura *transformer*, em conjunto com a análise dos dados de marcha, pode favorecer o diagnóstico automático de doenças neurodegenerativas. Dessa forma, a abordagem tem potencial para viabilizar a identificação precoce da condição, permitindo o início antecipado do tratamento, o que pode contribuir para desacelerar a progressão da doença e promover uma melhor qualidade de vida ao paciente.

Como sugestão de trabalhos futuros, propõe-se o refinamento das estratégias de segmentação dos dados para evitar possíveis viéses. Tal medida poderá viabilizar o desenvolvimento de modelos mais precisos e eficientes para o apoio ao diagnóstico de doenças neurodegenerativas. Além dos resultados apresentados neste documento, um artigo relacionado ao uso de redes *transformer* para classificação de doenças neurodegenerativas foi aceito para publicação em uma conferência internacional [Bucci *et al.*, 2025]. Resultados preliminares e colaborações em trabalhos relacionados ao diagnóstico de DNDs foram publicados em uma revista [Silva *et al.*, 2024], conferência nacional [Chagas *et al.*, 2024a] e conferências locais [Bucci *et al.*, 2024; Chagas *et al.*, 2024b].

Declarações complementares

Agradecimentos

O presente trabalho foi realizado com apoio do Conselho Nacional de Desenvolvimento Científico (CNPq) e da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) – Código de Financiamento 001, e Edital Nº 30/2022 – Programa Emergencial de Solidariedade Acadêmica (PDPG).

Contribuições dos autores

GFFBB – metodologia, investigação, experimentos e escrita; ALBC – experimentos e escrita; JPJ – concepção, supervisão e validação; HADN – supervisão; FS – administração do projeto e obtenção de financiamento. Giordana Bucci é a principal contribuidora deste manuscrito. Todos os autores leram e aprovaram o manuscrito final.

Conflitos de interesse

Os autores declaram que não têm nenhum conflito de interesses

Disponibilidade de dados e materiais

O conjunto de dados analisado durante o estudo atual estão disponíveis

veis em physionet.org.

Referências

- Agarwal, K., Dheekollu, L., Dhama, G., Arora, A., Asthana, S., and Bhowmik, T. (2020). Deep learning based time series forecasting. In *2020 19th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pages 859–864. DOI: 10.1109/ICMLA51294.2020.00140.
- Amooei, E., Sharifi, A., and Manthouri, M. (2023). Early diagnosis of neurodegenerative diseases using cnn-lstm and wavelet transform. *Journal of Healthcare Informatics Research*, 7(1):104–124. DOI: 10.1007/s41666-023-00130-9.
- Bucci, G. d. F. F. B., Chagas, A. L., Felix, J., and Nascimento, H. (2024). Aprendizado de máquina profundo aplicado ao diagnóstico de doenças neurodegenerativas a partir da marcha. In *21º Congresso de pesquisa, ensino e extensão (CONPEEX), Goiânia, GO, Brasil, Novembro 06–08, 2024*, page 385. UFG. Disponível em: <https://anaisconpeex.ciar.ufg.br/edicoes/21/index.html>.
- Bucci, G. d. F. F. B., Felix, J., Salvini, R., Nascimento, H., and Soares, F. (2025). Encoder-only transformer for detecting multiple neurodegenerative diseases from gait analysis. In *2025 IEEE 49th Annual Computers, Software, and Applications Conference (COMPSAC), Toronto, Canada*, pages 1–5. DOI: 10.1109/COMPSAC65507.2025.00119.
- Chagas, A., Bucci, G., Félix, J., Fonseca, A., Nascimento, H., and Soares, F. (2024a). Avaliando a sobreamostragem de dados temporais de marcha no diagnóstico automático de doenças neurodegenerativas. In *Anais do XXIV Simpósio Brasileiro de Computação Aplicada à Saúde (SBCAS)*, pages 567–578, Porto Alegre, RS, Brasil. SBC. DOI: 10.5753/sbcas.2024.2776.
- Chagas, A. L. d. B., Felix, J., Bucci, G., and Nascimento, H. (2024b). Diagnóstico de doenças neurodegenerativas a partir da marcha: Uma aplicação de métodos clássicos de aprendizado de máquina. In *21º Congresso de pesquisa, ensino e extensão (CONPEEX), Goiânia, GO, Brazil, Novembro 06–08, 2024*, page 1023. UFG. Disponível em: <https://anaisconpeex.ciar.ufg.br/edicoes/21/index.html>.
- Erdaş, Ç. B. and Sümer, E. (2024). Cnn-based neurodegenerative disease classification using qr-represented gait data. *Brain and Behavior*, 14(10). DOI: 10.1002/brb3.70100.
- Erdaş, Ç. B., Sümer, E., and Kibaroğlu, S. (2021). Neurodegenerative disease detection and severity prediction using deep learning approaches. *Biomedical Signal Processing and Control*, 70:103069. DOI: 10.1016/j.bspc.2021.103069.
- Erdaş, Ç. B., Sümer, E., and Kibaroğlu, S. (2023). Neurodegenerative diseases detection and grading using gait dynamics. *Multimedia Tools and Applications*, 82(15):22925 – 22942. DOI: 10.1007/s11042-023-14461-7.
- Erkkinen, M. G., Kim, M.-O., and Geschwind, M. D. (2018). Clinical neurology and epidemiology of the major neurodegenerative diseases. *Cold Spring Harbor Perspectives in Biology*, 10(4):a033118. DOI: 10.1101/cshperspect.a033118.
- Faisal, M. A. A., Chowdhury, M. E. H., Mahbub, Z. B., Pedersen, S., Ahmed, M. U., Khandakar, A., Alhatou, M., Nabil, M., Ara, I., Bhuiyan, E. H., Mahmud, S., and AbdulMoniem, M. (2023). Nddnet: a deep learning model for predicting neurodegenerative diseases from gait pattern. *Applied Intelligence*, 53(17):20034 – 20046. DOI: 10.1007/s10489-023-04557-w.
- Fraiwan, L. and Hassanin, O. (2021). Computer-aided identification of degenerative neuromuscular diseases based on gait dynamics and ensemble decision tree classifiers. *Plos one*, 16(6):e0252380. DOI: 10.1371/journal.pone.0252380.
- Hausdorff, J. M., Ladin, Z., and Wei, J. Y. (1995). Footswitch system for measurement of the temporal parameters of gait. *Journal of Biomechanics*, 28(3):347–351. DOI: 10.1016/0021-9290(94)00074-e.
- Hausdorff, J. M., Lertratanakul, A., Cudkowicz, M. E., Peterson, A. L., Kaliton, D., and Goldberger, A. L. (2000). Dynamic Markers of Altered Gait Rhythm in Amyotrophic Lateral Sclerosis. *Journal of Applied Physiology*, 88(6):2045–2053. DOI: 10.1152/jappl.2000.88.6.2045.
- Hausdorff, J. M., Mitchell, S. L., Firtion, R., Peng, C.-K., Cudkowicz, M. E., Wei, J. Y., and Goldberger, A. L. (1997). Altered fractal dynamics of gait: reduced stride-interval correlations with aging and Huntington's disease. *Journal of Applied Physiology*, 82(1):262–269. DOI: 10.1152/jappl.1997.82.1.262.
- Heemels, M.-T. (2016). Neurodegenerative diseases. *Nature*, 539(7628):179–180. DOI: 10.1038/539179a.
- Maragakis, N. J. and Rothstein, J. D. (2006). Mechanisms of disease: astrocytes in neurodegenerative disease. *Nature Clinical Practice Neurology*, 2(12):679–689. DOI: 10.1038/ncpneuro0355.
- Mayeux, R. (2003). Epidemiology of neurodegeneration. *Annual Review of Neuroscience*, 26(1):81–104. DOI: 10.1146/annurev.neuro.26.043002.094919.
- Naimi, S., Bouachir, W., and Bilodeau, G.-A. (2023). 1d-convolutional transformer for parkinson disease diagnosis from gait. *Neural Computing and Applications*, 36(4):1947–1957. DOI: 10.1007/s00521-023-09193-6.
- Nguyen, D. M. D., Miah, M., Bilodeau, G.-A., and Bouachir, W. (2022). Transformers for 1d signals in parkinson's disease detection from gait. In *2022 26th International Conference on Pattern Recognition (ICPR)*, pages 5089–5095. DOI: 10.1109/ICPR56361.2022.9956330.
- Ning, Z., Li, L., and Jin, X. (2018). Classification of neurodegenerative diseases based on CNN and LSTM. In *2018 9th International Conference on Information Technology in Medicine and Education (ITME)*, pages 82–85. IEEE. DOI: 10.1109/ITME.2018.00029.
- Rao, H., Zeng, M., Zhao, X., and Miao, C. (2025). A survey of artificial intelligence in gait-based neurodegenerative disease diagnosis. *Neurocomputing*, 626:129533. DOI: 10.1016/j.neucom.2025.129533.
- Silva, M. I. d., Felix, J. P., Prado, T. d. S., Chagas, A. L. d. B., Bucci, G. d. F. F. B., da Fonseca, A. U., and Soares, F. (2024). Sobre a análise de sinais de voz para o diagnóstico da doença de parkinson. *Journal of Health Informatics*, 16(Especial). DOI: 10.59681/2175-4411.v16.iEspecial.2024.1305.
- Sun, H.-J. and Zhang, Z.-G. (2022). Transformer-based

- severity detection of parkinson's symptoms from gait. In *2022 15th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, pages 1–5. DOI: 10.1109/CISP-BMEI56279.2022.9980289.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., and Polosukhin, I. (2017). Attention is all you need. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, NIPS'17, page 6000–6010, Red Hook, NY, USA. Curran Associates Inc. Disponível em: <https://dl.acm.org/doi/10.5555/3295222.3295349>.
- Zhao, A., Qi, L., Li, J., Dong, J., and Yu, H. (2018). LSTM for diagnosis of neurodegenerative diseases using gait data. In Yu, H. and Dong, J., editors, *Ninth International Conference on Graphic and Image Processing (ICGIP 2017)*, volume 10615, page 106155B. International Society for Optics and Photonics, SPIE. DOI: 10.1117/12.2305277.