



UFOP

Universidade Federal
de Ouro Preto

**Universidade Federal de Ouro Preto
Instituto de Ciências Exatas e Aplicadas
Departamento de Computação e Sistemas**

Classificação de Sinais de EEG para diagnóstico de epilepsia usando Machine Learning

Felipe Victor Marques de Sousa

TRABALHO DE CONCLUSÃO DE CURSO

ORIENTAÇÃO:

Dr. Glauco Ferreira Gazel Yared

Março, 2026

João Monlevade—MG

Felipe Victor Marques de Sousa

Classificação de Sinais de EEG para diagnóstico de epilepsia usando Machine Learning

Orientador: Dr. Glauco Ferreira Gazel Yared

Monografia apresentada ao curso de Engenharia de Computação do Instituto de Ciências Exatas e Aplicadas, da Universidade Federal de Ouro Preto, como requisito parcial para aprovação na Disciplina “Trabalho de Conclusão de Curso II”.

Universidade Federal de Ouro Preto

João Monlevade

Março de 2026

SISBIN - SISTEMA DE BIBLIOTECAS E INFORMAÇÃO

S725c Sousa, Felipe Victor Marques de.
Classificação de sinais de EEG para diagnóstico de epilepsia usando
machine learning. [manuscrito] / Felipe Victor Marques de Sousa. - 2026.
59 f.: il.: color., gráf., tab..

Orientador: Prof. Dr. Glauco Ferreira Gazel Yared.
Monografia (Bacharelado). Universidade Federal de Ouro Preto.
Instituto de Ciências Exatas e Aplicadas. Graduação em Engenharia de
Computação .

1. Aprendizado do computador. 2. Diagnóstico por computador. 3.
Epilepsia. 4. Eletroencefalografia. 5. Processamento de sinais. 6. Redes
neurais (Computação). I. Yared, Glauco Ferreira Gazel. II. Universidade
Federal de Ouro Preto. III. Título.

CDU 616.853:004.85

Bibliotecário(a) Responsável: Flavia Reis - CRB6/2431



FOLHA DE APROVAÇÃO

Felipe Victor Marques de Sousa

Classificação de sinais de EEG para diagnóstico de epilepsia usando Machine Learning

Monografia apresentada ao Curso de Engenharia de Computação da Universidade Federal de Ouro Preto como requisito parcial para obtenção do título de Bacharel em Engenharia de Computação

Aprovada em 4 de março de 2026.

Membros da banca

Doutor - Glauco Ferreira Gazel Yared - Orientador - Universidade Federal de Ouro Preto
Doutora - Gilda Aparecida de Assis - Universidade Federal de Ouro Preto
Doutor - Ronan Silva Ferreira - Universidade Federal de Ouro Preto

Glauco Ferreira Gazel Yared, orientador do trabalho, aprovou a versão final e autorizou seu depósito na Biblioteca Digital de Trabalhos de Conclusão de Curso da UFOP em 25/03/2026.



Documento assinado eletronicamente por **Glauco Ferreira Gazel Yared, PROFESSOR DE MAGISTERIO SUPERIOR**, em 25/03/2026, às 19:15, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site http://sei.ufop.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **1081318** e o código CRC **644DBC46**.

Este trabalho é dedicado à minha família, que nesta e em muitas gerações passadas, lutou por se manter de pé, por cuidar bem e ensinar aos seus e, diante de tudo isso, verá que um filho teu não foge à luta (como diz nosso saudosos hino nacional), e formará o primeiro Engenheiro da família Marques e Sousa. Essa vitória é de Margarida (vó), é de Eliane (mãe), é de Maria (vó que orgulhosa se aqui estivesse), é de Nilson (pai) e também é de Luiza (irmã), bem como de Patrick (irmão).

Agradecimentos

Agradeço todo o apoio da família e amigos, mas quero agradecer, sobretudo, a mim mesmo, que diante das incertezas deste caminho, dos riscos de se jogar no mundo e das dificuldades, largou um emprego estável, ficou distante de momentos importantes em família, para buscar e viver essa graduação durante todos esses anos.

Reconheço a necessidade de agradecer às políticas de permanência da PRACE e ao programa de extensão do Projeto de Inclusão Digital (PID), que contribuíram não só para a minha formação, mas também para a dedicação exclusiva aos estudos.

“O inventor, como a natureza de Linneu, não faz saltos; progride de manso, evolui.”

DUMONT, S. O que eu vi, o que nós veremos. São Paulo, 1918.

Resumo

A epilepsia é uma desordem neurológica crônica que impõe restrições à qualidade de vida dos pacientes, acarretando riscos físicos, como traumas decorrentes de quedas, e impactos psicossociais associados à imprevisibilidade das crises. O monitoramento contínuo apresenta-se como uma solução vital para a mitigação desses riscos, porém a análise manual prolongada de eletroencefalogramas (EEG) é inviável em larga escala. Este trabalho propõe e avalia comparativamente arquiteturas de *Deep Learning* (CNN 1D, CNN 2D com Transformada Wavelet Contínua e C-RNN com mecanismo de Atenção) frente a modelos de *Machine Learning* clássico (KNN e XGBoost). O objetivo central foi validar algoritmos capazes de operar com segurança clínica, priorizando a minimização da taxa de Falsos Negativos por Hora (FN/h). Utilizando o banco de dados CHB-MIT, em que os pacientes são majoritariamente pediátricos, os modelos foram submetidos a um protocolo de validação por janelas deslizantes em 10,6 horas de registros contínuos. Os resultados demonstraram a superioridade das abordagens baseadas em engenharia de características avançada: o modelo XGBoost, operando com limiar de decisão de 0,60 e alimentado por parâmetros de Hjorth e potência espectral, atingiu o melhor desempenho de segurança, registrando a menor taxa de omissões (3,12 FN/h) e a maior acurácia global (99,30%). Em comparação, arquiteturas profundas como a C-RNN com atenção, embora específicas, apresentaram taxas superiores (17,42 FN/h), comprometendo a detecção de eventos críticos. Conclui-se que o modelo XGBoost oferece o equilíbrio ideal entre segurança (alta sensibilidade) e usabilidade operacional, validando sua aplicação em dispositivos embarcados de alerta.

Palavras-chave: Epilepsia. Eletroencefalograma. Machine Learning. Redes neurais. EEG.

Abstract

Epilepsy is a chronic neurological disorder that imposes constraints on patients' quality of life, leading to physical risks, such as trauma resulting from falls, and psychosocial impacts associated with the unpredictability of seizures. Continuous monitoring emerges as a vital solution for mitigating these risks; however, prolonged manual analysis of electroencephalograms (EEG) is infeasible on a large scale. This work proposes and comparatively evaluates Deep Learning architectures (1D CNN, 2D CNN with Continuous Wavelet Transform, and C-RNN with Attention mechanism) against classical Machine Learning models (KNN and XGBoost). The central objective was to validate algorithms capable of operating with clinical safety, prioritizing the minimization of the False Negatives per Hour (FN/h) rate. Using the CHB-MIT database, which is primarily pediatric, the models were subjected to a sliding window validation protocol over 10.6 hours of continuous recordings. Results demonstrated the superiority of approaches based on advanced feature engineering: the XGBoost model, operating with a decision threshold of 0.60 and fed with Hjorth parameters and spectral power, achieved the best safety performance, recording the lowest omission rate (3.12 FN/h) and the highest global Accuracy (99.30%). In comparison, deep architectures such as the C-RNN with Attention, despite their specificity, presented higher rates (17.42 FN/h), compromising the detection of critical events. It is concluded that the XGBoost model offers the ideal balance between safety (high sensitivity) and operational usability, validating its application in embedded alert devices.

Key-words: Epilepsy. Electroencephalogram. Machine Learning. Neural networks. EEG.

Lista de ilustrações

Figura 1 – Representação esquemática de um neurônio e suas subdivisões funcionais.	18
Figura 2 – Sistema de posicionamento internacional de eletrodos 10-20.	21
Figura 3 – Comparação de Morfologia do Sinal de EEG amostrado da base CHB-MIT.	22
Figura 4 – Sinal Spike característico.	23
Figura 5 – Comparação de potência do sinal antes e após remoção de ruídos. . . .	25
Figura 6 – Infográfico do funcionamento de uma rede CNN com camada Bi-LSTM.	30
Figura 7 – Perfil demográfico dos pacientes do conjunto CHB-MIT.	32
Figura 8 – Curadoria dos sinais para remoção de artefatos descartáveis	34
Figura 9 – Janela deslizante em um sinal amostrado da base CHB-MIT	34
Figura 10 – Gráfico que mostra a separação dos Folds.	36
Figura 11 – Gráfico que mostra a separação dos Folds para 100 janelas	36
Figura 12 – Valores apropriados para o decaimento e a largura de banda da wavelet.	39
Figura 13 – Escalograma de uma amostra do sinal de EEG.	40
Figura 14 – Validação amostrada em Barcode para o modelo CNN Clássico.	47
Figura 15 – Matriz de Confusão para o modelo CNN Clássico.	47
Figura 16 – Validação amostrada em Barcode para o modelo CNN + Wav	48
Figura 17 – Matriz de Confusão para o modelo CNN + Wav	48
Figura 18 – Validação amostrada em Barcode para o modelo CRNN-Att	49
Figura 19 – Matriz de Confusão para o modelo CRNN-Att	49
Figura 20 – Melhor K com base na métrica de Manhattan	50
Figura 21 – Matriz de Confusão para o modelo KNN	50
Figura 22 – Validação amostrada em Barcode para o modelo XGB	51
Figura 23 – Comportamento do modelo durante crise (Probabilidade vs Tempo) . .	51
Figura 24 – Matriz de Confusão para o modelo XGB	52
Figura 25 – Importância dos Atributos (Feature Importance) no XGBoost	52

Lista de tabelas

Tabela 1 – Comparação final: Acurácia Global e métricas de desempenho clínico ordenadas por arquitetura.	46
Tabela 2 – Quadro síntese das conclusões do estudo.	54

Lista de abreviaturas e siglas

AUC	<i>Area Under the Curve</i> (Área Sob a Curva)
CHB-MIT	<i>Children's Hospital Boston - Massachusetts Institute of Technology</i>
CNN	<i>Convolutional Neural Network</i> (Rede Neural Convolucional)
C-RNN	<i>Convolutional Recurrent Neural Network</i>
CWT	<i>Continuous Wavelet Transform</i> (Transformada Wavelet Contínua)
DECSI	Departamento de Computação e Sistemas
DL	<i>Deep Learning</i> (Aprendizado Profundo)
EC	Engenharia de Computação
EEG	Eletroencefalograma
FHWA	<i>First Half-Wave Amplitude</i> (Amplitude da primeira meia-onda)
FN	Falsos Negativos
FP	Falsos Positivos
ICEA	Instituto de Ciências Exatas e Aplicadas
JM	João Monlevade
KNN	<i>K-Nearest Neighbors</i> (K-Vizinhos Mais Próximos)
LSTM	<i>Long Short-Term Memory</i>
ML	<i>Machine Learning</i> (Aprendizado de Máquina)
PSD	<i>Power Spectral Density</i> (Densidade Espectral de Potência)
RMS	<i>Root Mean Square</i> (Raiz Quadrada da Média)
ROC	<i>Receiver Operating Characteristic</i>
RNN	<i>Recurrent Neural Network</i> (Rede Neural Recorrente)
SHWA	<i>Second Half-Wave Amplitude</i> (Amplitude da segunda meia-onda)
STFT	<i>Short-Time Fourier Transform</i> (Transformada de Fourier de Curto Prazo)

TN	Verdadeiros Negativos
TP	Verdadeiros Positivos
UFOP	Universidade Federal de Ouro Preto
XGBoost	<i>Extreme Gradient Boosting</i>

Lista de símbolos

α	Banda de frequência Alpha (8–13 Hz) ou peso do mecanismo de atenção
β	Banda de frequência Beta (13–30 Hz)
δ	Banda de frequência Delta (0,5–4 Hz)
γ	Banda de frequência Gamma (30–45 Hz)
θ	Banda de frequência Theta (4–8 Hz)
$\psi(t)$	Função Wavelet Mãe
ω_0	Frequência central da Wavelet de Morlet
σ	Parâmetro de largura de banda (Wavelet) ou função de ativação Sigmoide
τ	Limiar de decisão (<i>Threshold</i>) ou parâmetro de translação temporal
f_s	Frequência de amostragem (128 Hz)
k	Número de vizinhos mais próximos (Hiperparâmetro do KNN)
s	Parâmetro de escala da Transformada Wavelet
$x(t)$	Sinal de EEG no domínio do tempo
X	Matriz ou tensor de entrada (Features)
y	Rótulo real da classe (<i>Ground Truth</i>)
$\hat{y}$	Rótulo predito pelo modelo
μ	Média aritmética
R	Conjunto dos números reais
$\in$	Pertence
Σ	Somatório
$\int$	Integral
$*$	Operação de convolução

Sumário

1	INTRODUÇÃO	15
1.1	O problema de pesquisa	15
1.2	Objetivos	16
1.3	Metodologia	16
1.4	Organização do trabalho	17
2	REVISÃO BIBLIOGRÁFICA	18
2.1	Fisiologia Cerebral e Fundamentos do EEG	18
2.1.1	A Base Eletroquímica: Potencial de Membrana e Condução	18
2.1.2	O Eletroencefalograma (EEG)	19
2.1.3	Sistemas de Aquisição e Padronização 10-20	20
2.1.3.1	O Sistema Internacional 10-20	20
2.2	Fisiopatologia da Epilepsia	21
2.3	Processamento e Caracterização Matemática do Sinal	22
2.3.1	Dinâmica do Sinal: Do Caos à Sincronia	22
2.3.2	Análise Mimética e Morfologia no Domínio do Tempo	23
2.3.3	Energia e Estatísticas de Ordem Superior (HOS)	24
2.3.3.1	Energia do Sinal (E)	24
2.3.3.2	Curtose e Assimetria	24
2.3.4	Análise Tempo-Frequência	24
2.3.5	Pré-processamento e Filtragem	24
2.3.6	Extração de Características	26
2.3.7	Transformada Wavelet Contínua	27
2.4	Algoritmos de Aprendizado de Máquina e Deep Learning	28
2.4.1	Redes Neurais Convolucionais Unidimensionais (1D-CNN)	28
2.4.2	Arquiteturas Híbridas: CRNN e Bi-LSTM	29
2.4.3	K-Nearest Neighbors (KNN)	30
2.4.4	XGBoost e Métodos de Ensemble	31
3	DESENVOLVIMENTO	32
3.0.1	Base de Dados CHB-MIT	32
3.0.2	Curadoria dos canais e pré-processamento	33
3.0.3	Janelas deslizantes	34
3.1	Processamento e Organização dos Dados	35
3.1.1	Arquitetura e Treinamento da Rede Neural Convolucional (CNN)	37
3.1.2	Arquitetura CNN-2D e Pré-processamento via CWT	38

3.1.2.1	Geração de Escalogramas e Paralelismo de CPU	40
3.1.2.2	Rede neural da CNN-2D e Otimização em Hardware	40
3.1.3	Arquitetura Híbrida C-RNN com Mecanismo de Atenção	41
3.1.3.1	Refinamento por Mineração de Exemplos Difíceis (<i>Hard Negative Mining</i>)	42
3.1.4	Arquitetura K-Nearest Neighbors (KNN) e Engenharia de Características	42
3.1.4.1	Extração de Características Manuais (<i>Feature Engineering</i>)	42
3.1.4.2	Balanceamento de Classes e Normalização	43
3.1.4.3	Otimização de Hiperparâmetros via <i>Grid Search</i>	43
3.1.5	Classificação via Extreme Gradient Boosting (XGBoost)	43
3.1.5.1	Extração de Parâmetros de Hjorth e Potência Espectral	43
3.1.5.2	Interpretabilidade e Importância dos Atributos	44
3.1.5.3	Suavização Temporal e Eficácia Clínica	44
3.1.5.4	Validação na Base Real (Janela Deslizante)	44
4	RESULTADOS	46
4.0.0.1	CNN Clássica (1D)	46
4.0.0.2	CNN + Transformada Wavelet (CWT)	47
4.0.0.3	C-RNN com Mecanismo de Atenção	48
4.0.0.4	K-Nearest Neighbors (Baseline)	49
4.0.0.5	XGBoost (Modelo Otimizado)	51
5	CONCLUSÃO	53
5.0.1	Trabalhos Futuros	54
	REFERÊNCIAS	56

1 Introdução

A epilepsia é um dos distúrbios neurológicos crônicos mais comuns, afetando aproximadamente 50 milhões de pessoas em todo o mundo (LIU et al., 2021; WANI; KARKI, 2024; VETRIKANI; BOBBY, 2017). Clinicamente, é definida como uma desordem resultante da ocorrência transitória de atividade neuronal anormal e excessiva no córtex cerebral (VETRIKANI; BOBBY, 2017; WANI; KARKI, 2024; SANEI; CHAMBERS, 2007). O seu sintoma principal é a crise epiléptica, caracterizada por mudanças súbitas nos movimentos, comportamento ou percepção sensorial, podendo durar de alguns segundos a vários minutos (VETRIKANI; BOBBY, 2017; SANEI; CHAMBERS, 2007).

No cenário brasileiro, a prevalência da doença apresenta uma natureza heterogênea devido à vasta diversidade socioeconômica e geográfica do país. No Rio de Janeiro, por exemplo, pesquisas realizadas em comunidades urbanas de baixa renda revelaram uma prevalência acumulada de 16,3 casos por 1.000 habitantes, com uma taxa de epilepsia ativa estimada em 5,1 por 1.000 (GOMES et al., 2002). Embora esses números possam ser inferiores aos encontrados em outros países da América Latina, a condição exige atenção contínua das autoridades de saúde devido ao seu impacto clínico e social.

Os impactos da patologia são multidimensionais. As crises de ausência (*petit mal*) podem comprometer o desempenho escolar, enquanto as crises tônico-clônicas generalizadas envolvem perda de consciência e riscos de injúrias físicas graves (SANEI; CHAMBERS, 2007; VETRIKANI; BOBBY, 2017). A recorrência desses episódios sem controle adequado pode levar a déficits cognitivos progressivos, depressão e isolamento social (LIU et al., 2021; FATMA; SINGH; SIDDIQUI, 2023).

1.1 O problema de pesquisa

Apesar da importância do diagnóstico precoce, a detecção e o monitoramento da epilepsia dependem fundamentalmente do Eletroencefalograma (EEG) (SANEI; CHAMBERS, 2007; VETRIKANI; BOBBY, 2017; WANI; KARKI, 2024). O problema central reside no fato de que a análise visual de registros de longa duração, que podem se estender por dias, é uma tarefa tediosa, demorada e dispendiosa. Ela exige especialistas altamente treinados para identificar padrões sutis em meio a ruídos e artefatos (LIU et al., 2021; FATMA; SINGH; SIDDIQUI, 2023; SANEI; CHAMBERS, 2007).

A fadiga humana e a subjetividade na interpretação podem levar a erros de diagnóstico ou à não detecção de eventos críticos. Nesse contexto, o desenvolvimento de sistemas automatizados baseados em *Machine Learning* (ML) e *Deep Learning* tornou-se

uma prioridade científica (LIU et al., 2021; WANI; KARKI, 2024). A lacuna que este trabalho busca preencher é a necessidade de ferramentas computacionais robustas que auxiliem a tomada de decisão médica, oferecendo maior precisão e rapidez na análise dos sinais.

1.2 Objetivos

O **objetivo geral** deste trabalho é desenvolver e avaliar um sistema automatizado para a detecção de crises epiléticas em sinais de EEG, visando auxiliar o diagnóstico clínico e reduzir a carga de trabalho manual dos especialistas (LIU et al., 2021; FATMA; SINGH; SIDDIQUI, 2023).

Para alcançar este propósito, definem-se os seguintes objetivos específicos:

- **Implementar** o processamento de sinais utilizando a base de dados *CHB-MIT Scalp EEG Database*, focando em registros de sujeitos pediátricos com crises rotuladas;
- **Investigar** técnicas de pré-processamento para remoção de artefatos de rede elétrica, ruídos musculares e padronização dos dados, garantindo a qualidade do sinal de entrada;
- **Avaliar** a eficácia de modelos de *Deep Learning*, especificamente arquiteturas híbridas baseadas em Redes Neurais Convolucionais (CNN) e Redes Neurais Recorrentes (LSTM/bi-LSTM), comparando seu desempenho com métodos tradicionais (WANI; KARKI, 2024).

1.3 Metodologia

A metodologia proposta para a execução deste trabalho está estruturada em quatro estágios, compondo o fluxo de reconhecimento de padrões:

1. **Estágio de Pré-processamento (*Preprocessing stage*)**: envolve a limpeza dos sinais brutos através de técnicas de sinais e sistemas, como a aplicação de filtros FIR passa-baixa (geralmente com corte em 40 Hz) e filtros *notch* para remoção de ruídos de rede elétrica e artefatos oculares ou musculares (WANI; KARKI, 2024; SANEI; CHAMBERS, 2007).
2. **Segmentação por Janela Deslizante (*Sliding window segmentation*)**: os sinais contínuos são divididos em segmentos menores de tempo para facilitar o treinamento dos modelos e capturar transições sutis na atividade elétrica (WANI; KARKI, 2024).

3. **Estágio de Extração de Características (*Feature extraction stage*):** identificação de atributos distintivos nos domínios do tempo e da frequência, como energia espectral e morfologia de ondas, que servem de entrada para os classificadores (VETRIKANI; BOBBY, 2017; FATMA; SINGH; SIDDIQUI, 2023).
4. **Estágio de Classificação (*Classification stage*):** uso de modelos de aprendizado de máquina e redes neurais profundas para rotular os sinais como normais ou epiléticos, visando maximizar não somente a acurácia e especificidade do sistema, mas sobretudo a sensibilidade, que é o indicador mais importante no contexto de dados médicos. (LIU et al., 2021; WANI; KARKI, 2024).

1.4 Organização do trabalho

O restante deste trabalho está organizado da seguinte forma: o capítulo 2 apresenta a fundamentação teórica, abordando os aspectos fisiológicos da epilepsia e os conceitos matemáticos do processamento digital de sinais. O Capítulo 3 descreve detalhadamente a metodologia, incluindo a caracterização da base de dados CHB-MIT e as arquiteturas de redes neurais propostas. O Capítulo 4 apresenta os resultados experimentais e as discussões sobre o desempenho dos classificadores. Por fim, o Capítulo 5 encerra o trabalho com as conclusões e sugestões para pesquisas futuras.

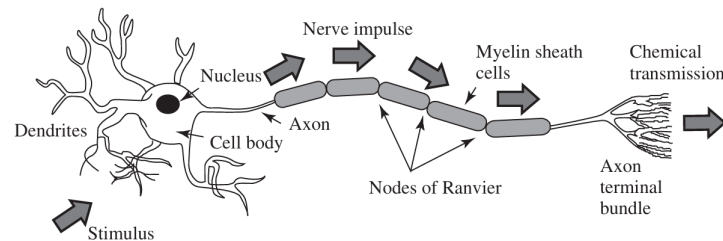
2 Revisão Bibliográfica

Este capítulo apresenta a fundamentação teórica necessária para a compreensão do trabalho, estabelecendo uma ponte entre a neurofisiologia e a engenharia biomédica. A discussão abrange desde a gênese eletroquímica dos sinais neurais até as técnicas avançadas de processamento digital de sinais (PDS) utilizadas para a detecção de padrões epiléticos.

2.1 Fisiologia Cerebral e Fundamentos do EEG

O sistema nervoso central é composto fundamentalmente por células nervosas, conhecidas como neurônios, e células da glia, que oferecem suporte metabólico e estrutural (SANEI; CHAMBERS, 2007). Estima-se que o cérebro humano possua bilhões de neurônios, cada um atuando como uma unidade de processamento de informação. A interação entre estas unidades ocorre através de junções especializadas chamadas sinapses, formando redes complexas responsáveis pela cognição e controle motor (WANI; KARKI, 2024).

Figura 1 – Representação esquemática de um neurônio e suas subdivisões funcionais.



Fonte: (SANEI; CHAMBERS, 2007)

Na Figura 1 podemos compreender a estrutura anatômica de um neurônio e a base do seu funcionamento. Os dendritos (*dendrites*) atuam como as principais vias receptoras, captando o estímulo (*stimulus*) inicial proveniente de outras células. Esse sinal é integrado no corpo celular (*cell body*), que abriga o núcleo (*nucleus*) e coordena as atividades metabólicas.

2.1.1 A Base Eletroquímica: Potencial de Membrana e Condução

Para que o estímulo captado trafegue pela estrutura neural, ocorre um processo eletroquímico mediado pela membrana celular. Esta membrana é uma bicamada lipídica que atua como um isolante elétrico, separando o meio intracelular do extracelular. O Potencial de Membrana é definido pela diferença de potencial elétrico (tensão) gerada pela

distribuição desigual de íons através dessa barreira, principalmente Sódio (Na^+), Potássio (K^+) e Cloro (Cl^-).

A dinâmica da propagação do sinal ocorre em duas fases elétricas principais:

- **Potencial de Repouso:** em estado basal, a célula é mais negativa internamente devido à alta permeabilidade ao K^+ e à ação da bomba de sódio-potássio. O neurônio mantém uma diferença de potencial de aproximadamente -70 mV.
- **Potencial de Ação (Impulso Nervoso):** se a soma dos estímulos nos dendritos despolarizar a célula até um limiar crítico, canais iônicos de Na^+ dependentes de tensão se abrem, permitindo um influxo massivo de cargas positivas. Isso causa uma inversão abrupta e transitória de polaridade (despolarização), seguida de um retorno ao estado basal (repolarização) (SANEI; CHAMBERS, 2007).

Uma vez gerado, este impulso nervoso (*nerve impulse*) propaga-se de forma unidirecional ao longo do axônio (*axon*), um prolongamento cilíndrico especializado na condução elétrica. A eficiência dessa condução é garantida pelas células da bainha de mielina (*myelin sheath cells*), que envolvem o axônio. Ao longo da bainha existem falhas chamadas nódulos de Ranvier (*nodes of Ranvier*), regiões altamente concentradas em canais iônicos dependentes de tensão. Nesses nódulos, a diferença de potencial é constantemente regenerada, permitindo que o impulso "salte" de um nódulo para outro (condução saltatória), acelerando drasticamente a velocidade do sinal.

Por fim, ao alcançar os terminais axonais (*axon terminal bundle*), o potencial de ação promove a entrada de íons cálcio (Ca^{2+}), desencadeando a transmissão química (*chemical transmission*). Vesículas liberam neurotransmissores na fenda sináptica, convertendo o sinal elétrico do axônio em uma mensagem química capaz de alterar o potencial de membrana da célula vizinha e reiniciar o ciclo.

2.1.2 O Eletroencefalograma (EEG)

O **Eletroencefalograma (EEG)** é a ferramenta clínica que registra essa atividade elétrica de forma não invasiva. O sinal captado pelo EEG de superfície representa a somatória espacial e temporal dos potenciais pós-sinápticos de milhões de neurônios piramidais no córtex. Devido às camadas de tecido (dura-máter, crânio e couro cabeludo), este sinal sofre atenuação severa e atua como um filtro passa-baixa natural (SANEI; CHAMBERS, 2007; VETRIKANI; BOBBY, 2017).

Clinicamente, este sinal é decomposto em bandas de frequência ou ritmos:

- **Delta (0,5–4 Hz):** Ondas de alta amplitude, predominantes em sono profundo e na infância.

- **Theta (4–8 Hz):** Associado à sonolência, acesso à memória e meditação.
- **Alpha (8–13 Hz):** Típico de estados de vigília relaxada, com os olhos fechados.
- **Beta (13–30 Hz):** Relacionado ao pensamento ativo, ansiedade e foco cognitivo.
- **Gamma (>30 Hz):** Associado ao processamento sensorial complexo e ligação de informações (binding problem) (SANEI; CHAMBERS, 2007).

2.1.3 Sistemas de Aquisição e Padronização 10-20

A aquisição do sinal de EEG pode ser categorizada em dois métodos principais, dependendo da invasividade e da localização dos eletrodos:

1. **EEG de Superfície (*Scalp EEG*):** método não invasivo em que os eletrodos são fixados sobre o couro cabeludo. Embora o sinal sofra atenuação devido à resistência do crânio e tecidos moles, é o método padrão para diagnóstico inicial e monitoramento de longo prazo, sendo a modalidade utilizada na base de dados deste trabalho (*CHB-MIT Scalp EEG Database*) (GUTTAG, 2010; SANEI; CHAMBERS, 2007).
2. **EEG Intracraniano (iEEG):** método invasivo que requer intervenção cirúrgica. Inclui o Eletrocorticograma (ECoG), em que eletrodos são dispostos diretamente sobre o córtex cerebral (sob a dura-máter), e eletrodos de profundidade, inseridos em estruturas cerebrais específicas (SANEI; CHAMBERS, 2007). Embora ofereça melhor resolução espacial e maior amplitude de sinal (devido à ausência da barreira óssea), seu uso é restrito a avaliações pré-cirúrgicas complexas.

2.1.3.1 O Sistema Internacional 10-20

Para garantir a reprodutibilidade dos exames e a consistência na análise espacial das atividades cerebrais entre diferentes laboratórios e pacientes, a Federação Internacional de Sociedades de Eletroencefalografia e Neurofisiologia Clínica recomendou o **Sistema Internacional 10-20** de posicionamento de eletrodos (SANEI; CHAMBERS, 2007).

A nomenclatura 10-20 refere-se ao fato de que os eletrodos são posicionados a distâncias de 10% ou 20% da distância total entre marcos anatômicos específicos do crânio: o **Násio** (a depressão na parte superior do nariz, entre os olhos) e o **Ínio** (a protuberância óssea na base do crânio, na parte posterior da cabeça) (SANEI; CHAMBERS, 2007).

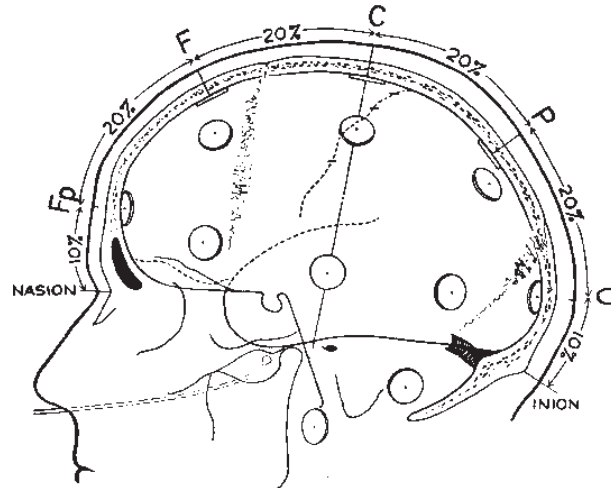
A identificação dos eletrodos combina letras e números para indicar a localização anatômica e o hemisfério (SANEI; CHAMBERS, 2007; VETRIKANI; BOBBY, 2017):

- **Letras (Lobos):** F (Frontal), T (Temporal), C (Central), P (Parietal) e O (Occipital). Note que a região Central (C) é usada como referência, embora não exista um lobo central anatomicamente definido.

- **Números (Lateralidade):** números ímpares referem-se a eletrodos no hemisfério esquerdo, enquanto números pares referem-se ao hemisfério direito. A letra z (de zero) indica eletrodos posicionados na linha média sagital do crânio (ex: Fz, Cz, Pz).

A Figura 2 ilustra a disposição espacial padrão deste sistema, essencial para a análise multicanal realizada nesta monografia.

Figura 2 – Sistema de posicionamento internacional de eletrodos 10-20.



Fonte: (KLEM et al., 1999)

2.2 Fisiopatologia da Epilepsia

A epilepsia é um distúrbio neurológico crônico caracterizado pela predisposição duradoura a gerar crises epiléticas (VETRIKANI; BOBBY, 2017). Uma crise é definida como a ocorrência transitória de sinais e/ou sintomas devido a uma atividade neuronal anormal, excessiva ou síncrona no cérebro (WANI; KARKI, 2024).

Do ponto de vista da dinâmica de sistemas, o cérebro saudável opera em um regime caótico (complexidade necessária para processar informações variadas). Durante uma crise, ocorre uma transição de fase: o cérebro transita do estado caótico para um estado ordenado/sincronizado, em que grandes populações de neurônios disparam simultaneamente tempestades elétricas, reduzindo a capacidade de processamento de informação.

Temporalmente, a dinâmica da epilepsia é dividida em quatro estados cruciais para sistemas de predição (FATMA; SINGH; SIDDIQUI, 2023):

1. **Interictal:** o intervalo entre as crises (estado basal do paciente).
2. **Pré-ictal:** o período de transição que antecede o evento. É nesta janela que biomarcadores sutis aparecem, permitindo a predição.

3. **Ictal:** a crise propriamente dita, caracterizada por alta energia e sincronia.
4. **Pós-ictal:** o período de recuperação imediata após a crise.

2.3 Processamento e Caracterização Matemática do Sinal

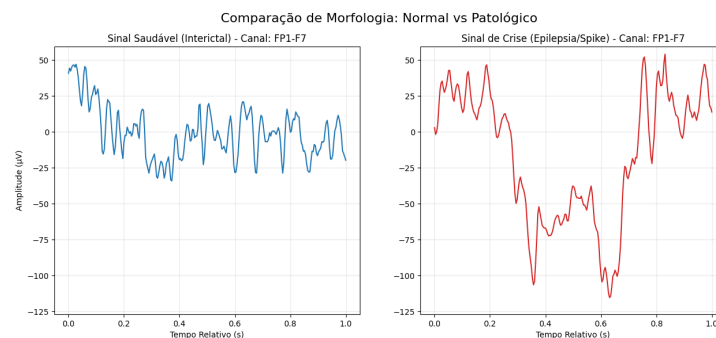
A detecção automatizada de epilepsia depende fundamentalmente da tradução de padrões visuais clínicos em descritores matemáticos quantitativos. O sinal de EEG é inerentemente estocástico, não-estacionário e, em estados patológicos, apresenta transições abruptas de comportamento dinâmico (SANEI; CHAMBERS, 2007).

2.3.1 Dinâmica do Sinal: Do Caos à Sincronia

Em um cérebro saudável, a atividade elétrica exibe um comportamento **caótico** de alta dimensão, refletindo a complexidade necessária para o processamento de informações variadas. Uma crise epiléptica, paradoxalmente, representa um colapso dessa complexidade: o sistema transita para um estado altamente **ordenado e hipsíncrono** (SANEI; CHAMBERS, 2007).

Matematicamente, essa transição é caracterizada pela redução da dimensão fractal e pelo aumento da periodicidade do sinal. O evento ictal é marcado por descargas paroxísticas que se manifestam como *spikes* (pontas), ondas agudas e complexos ponta-onda (SANEI; CHAMBERS, 2007; FATMA; SINGH; SIDDIQUI, 2023).

Figura 3 – Comparação de Morfologia do Sinal de EEG amostrado da base CHB-MIT.



Fonte: Autor

A figura 3 mostra como é a diferença entre um Sinal Saudável (gráfico à esquerda) e um Sinal de Crise (gráfico à direita) amostrados do canal FP1-F7. O sinal saudável se mostra mais comportado em relação à amplitude, variando entre aproximadamente $-25\mu V$ até pouco menos de $50\mu V$.

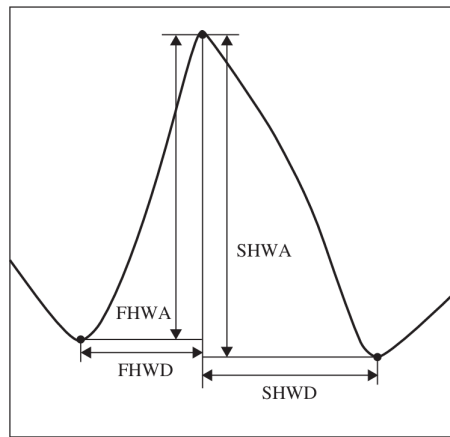
Por outro lado, o sinal de Crise mostra uma variação de amplitude mais abrupta, atingindo picos próximos de $50\mu V$ e caindo para perto de $-125\mu V$.

2.3.2 Análise Mimética e Morfologia no Domínio do Tempo

A análise mimética busca reproduzir computacionalmente o processo de inspeção visual realizado pelo neurofisiologista. O elemento central desta análise é a **ponta epiléptica** (*epileptic spike*), definida como um transiente com duração típica entre 20 e 70 ms (SANEI; CHAMBERS, 2007).

Para quantificar a geometria dessa onda, utilizam-se métricas baseadas na decomposição da onda em duas metades (subida e descida), conforme proposto por Anusha et al. (ANUSHA; MATHEWS; PUTHANKATTIL, 2012):

Figura 4 – Sinal Spike característico.



Fonte: (SANEI; CHAMBERS, 2007)

Conforme apresentado na figura 4, podemos compreender a morfologia do sinal característico com base na amplitude e na duração do sinal, em termos de FHWA, SHWA, FHWD e SHWD.

- **Amplitude (FHWA e SHWA):** a *First Half-Wave Amplitude* (FHWA) mede a diferença de potencial do início ao pico da onda, enquanto a *Second Half-Wave Amplitude* (SHWA) mede do pico ao fim da onda.
- **Duração (FHWD e SHWD):** correspondem aos intervalos de tempo de subida e descida, respectivamente.
- **Inclinação (Slope):** a agudeza da onda, crítica para diferenciar uma ponta epiléptica de um artefato ou ritmo de fundo, é calculada pela razão entre amplitude e duração. Uma inclinação elevada é um forte indicativo de atividade epileptiforme:

$$Slope_{media} \approx \frac{1}{N} \sum_{i=1}^N \frac{Amplitude_i}{Duração_i} \quad (2.1)$$

2.3.3 Energia e Estatísticas de Ordem Superior (HOS)

Métodos baseados apenas em média e variância (estatísticas de 1ª e 2ª ordem) assumem que o sinal tende a seguir uma distribuição Gaussiana. No entanto, o EEG epiléptico é intrinsecamente não-Gaussiano e não-linear em função de ondas agudas provenientes do sinal característicos Spike, o que torna o sinal epiléptico imperfeito e assimétrico (SANEI; CHAMBERS, 2007). Para capturar essa natureza, utilizam-se:

2.3.3.1 Energia do Sinal (E)

Reflete a intensidade da descarga neuronal síncrona. Em uma janela de tempo discreta de comprimento L e sinal n , a energia instantânea é a soma quadrática das amplitudes $x(p)$, o que amplifica os eventos de alta variação de potencial elétrico típicos da crise (SANEI; CHAMBERS, 2007):

$$E(n) = \frac{1}{L} \sum_{p=n-L/2}^{n-1+L/2} x^2(p) \quad (2.2)$$

2.3.3.2 Curtose e Assimetria

As Estatísticas de Ordem Superior (HOS) são fundamentais para detecção de transientes. A **Curtose** (*Kurtosis*), ou 4º momento estatístico, mede o achatamento da distribuição. Sinais com pontas esporádicas (distribuição leptocúrtica) apresentam alta curtose, diferenciando-se do ruído de fundo Gaussiano (SANEI; CHAMBERS, 2007; FATMA; SINGH; SIDDIQUI, 2023).

$$Kurtosis = \frac{E[(x - \mu)^4]}{\sigma^4} \quad (2.3)$$

em que μ é a média e σ é o desvio padrão.

2.3.4 Análise Tempo-Frequência

Devido à não-estacionariedade, a simples Transformada de Fourier (FFT) é insuficiente, pois perde a informação temporal dos eventos transientes. Abordagens modernas utilizam a Transformada Wavelet Contínua (CWT), que decompõe o sinal em coeficientes de escala e translação, permitindo visualizar a evolução da energia em diferentes bandas de frequência ao longo do tempo (escalogramas) (WANI; KARKI, 2024; SANEI; CHAMBERS, 2007). Isso permite detectar o início exato da atividade rítmica (teta ou alfa) que frequentemente precede ou acompanha a crise.

2.3.5 Pré-processamento e Filtragem

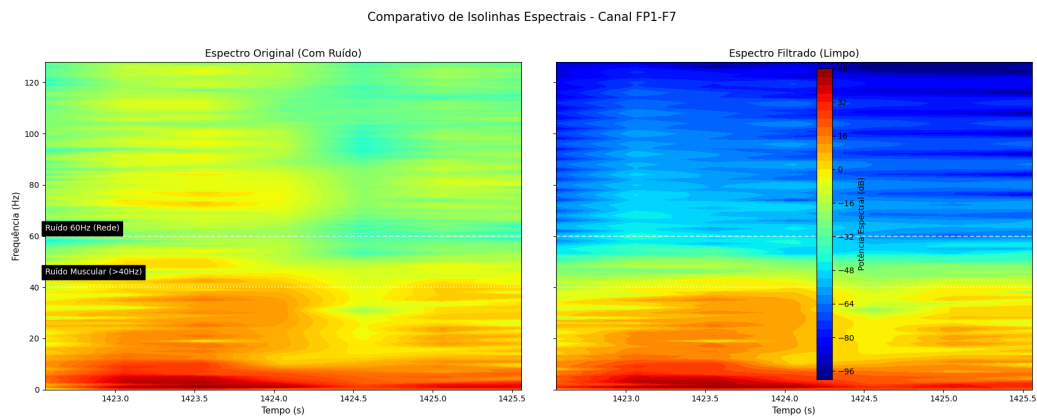
O EEG é um sinal de baixíssima amplitude (na ordem de microvolts), sendo altamente suscetível a ruídos. Antes da análise, é necessário remover artefatos que não

têm origem cerebral:

- **Ruído de Rede (Powerline):** interferência eletromagnética de 50 Hz ou 60 Hz proveniente da rede elétrica.
- **Artefatos Fisiológicos:** potenciais gerados por outros órgãos, como batimentos cardíacos (ECG), movimento ocular (EOG) e contração muscular (EMG).

Técnicas comuns incluem filtros *notch* para a frequência da rede e filtros passa-banda (ex: 0,5–40 Hz) para isolar as frequências cerebrais de interesse.

Figura 5 – Comparação de potência do sinal antes e após remoção de ruídos.



Fonte: elaboração do autor com base nos dados coletados

Na figura 5, uma amostra de uma parte do sinal mostra o resultado do processo de limpeza de ruído provocado por rede elétrica e também por atividade muscular. No gráfico à esquerda, o espectro original relaciona a frequência em termos de potência (em dB). No gráfico à direita, após o filtro que limpa os ruídos da rede elétrica, nota-se que a potência espectral vai de aproximadamente 16dB para faixas acima de 48db. Se tratando de potência de sinal, isso significa que o ruído não é mais significativo, o que é bom para a saúde do sinal como um todo.

Entretanto, para ruídos musculares, devido à proximidade com faixas provenientes da atividade cerebral, não é possível notar uma diferença significativa no espectro de potência em comparação com a amostra ruidosa. Isso é bastante comum para sinais característicos. Nem sempre é possível limpar todo tipo de ruído. É importante ressaltar que esta é uma pequena amostra do sinal trabalhado. O espectro em outras posições no tempo, podem ter tido um sucesso maior na limpeza de ruídos musculares por que a difusão com os sinais de importância não é tão significativa.

2.3.6 Extração de Características

A extração de características (*feature extraction*) é uma etapa fundamental no processamento de sinais digitais e na construção de modelos de aprendizado de máquina. Este processo consiste na transformação de dados brutos (*raw data*), frequentemente ruidosos e de alta dimensionalidade, em um conjunto reduzido de variáveis descritivas, denominado vetor de características (SANEI; CHAMBERS, 2007). O objetivo primário é mitigar a redundância da informação e isolar os padrões matemáticos mais discriminativos que representam a fisiopatologia subjacente, como a transição para o estado ictal (FATMA; SINGH; SIDDIQUI, 2023).

No contexto de sinais biológicos, especialmente no tratamento de dados de EEG que são inerentemente não-estacionários, o volume de informações gerado por múltiplas amostras temporais e diversos canais impõe desafios computacionais significativos. Algoritmos de classificação sofrem com a maldição da dimensionalidade caso sejam alimentados diretamente com o sinal bruto. Portanto, a engenharia de características atua como um tradutor essencial entre a dinâmica física do sinal e a matemática do algoritmo preditivo, permitindo a separação entre classes (ex: normal vs. crise) (LIU et al., 2021).

Na literatura de processamento de sinais de EEG, as características são tipicamente projetadas e agrupadas conforme o domínio de análise:

- **Domínio do Tempo:** envolve a extração de métricas diretas da série temporal. Características comuns incluem a energia do sinal, a amplitude de pico e a duração das meias-ondas (FHWA, SHWA, FHWD, SHWD), que tentam mimetizar a análise visual de pontas epiléticas (ANUSHA; MATHEWS; PUTHANKATTIL, 2012). Além disso, devido à natureza não-linear e não-gaussiana do EEG durante as crises, momentos estatísticos de ordem superior, como assimetria (*skewness*) e curtose, são amplamente empregados para capturar transientes abruptos (SANEI; CHAMBERS, 2007).
- **Domínio da Frequência e Tempo-Frequência:** dada a natureza rítmica da atividade cerebral (ritmos Alpha, Beta, etc.), a representação espectral fornece informações estruturais críticas. Técnicas clássicas utilizam a estimativa da Densidade Espectral de Potência (PSD) para quantificar a energia em bandas de frequência específicas (SANEI; CHAMBERS, 2007). Para lidar com a não-estacionariedade do sinal epilético, abordagens modernas utilizam a Transformada Wavelet Contínua (CWT), que decompõem o sinal em coeficientes de detalhe e aproximação, preservando a resolução temporal de eventos transientes (FATMA; SINGH; SIDDIQUI, 2023; WANI; KARKI, 2024).

Ao converter janelas temporais de amostras brutas em vetores de atributos estatís-

ticos e espectrais, o *pipeline* reduz o custo computacional e projeta os dados em um novo espaço de características (*feature space*). Este procedimento busca maximizar a separabilidade linear ou não-linear entre as classes, facilitando a convergência de classificadores como o XGBoost, usado neste trabalho (VETRIKANI; BOBBY, 2017).

2.3.7 Transformada Wavelet Contínua

O estudo de sinais não-estacionários, como é o caso de um EEG impõe desafios às técnicas clássicas de processamento. Embora a Transformada de Fourier seja uma ferramenta consagrada para identificar quais frequências compõem um sinal, ela falha em determinar *quando* essas frequências ocorrem, uma vez que utiliza funções globais (senóides infinitas) que não possuem localização temporal precisa (DOMINGUES et al., 2016).

Para superar essa limitação e analisar fenômenos transientes, utiliza-se a teoria Wavelet. Segundo Domingues et al. (2016), a Transformada Wavelet funciona como um microscópio matemático: a capacidade de ampliação é dada pelo parâmetro de escala (a) e a capacidade óptica é definida pela escolha da função analisadora (ψ), denominada wavelet-mãe.

Matematicamente, a Transformada Wavelet Contínua (CWT) de um sinal $f(t)$ é definida pela integral linear que mede a similaridade entre o sinal e a wavelet-mãe deslocada e escalonada (DOMINGUES et al., 2016):

$$W_{\psi}f(a, \tau) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t) \psi^* \left(\frac{t - \tau}{a} \right) dt \quad (2.4)$$

sendo:

- a (escala): parâmetro de dilatação ($a > 1$) ou compressão ($a < 1$). Escalas grandes correspondem a baixas frequências (visão global), enquanto escalas pequenas analisam altas frequências (detalhes) (DOMINGUES et al., 2016);
- τ (translação): parâmetro que desliza a wavelet ao longo do tempo, permitindo a localização temporal dos eventos (DOMINGUES et al., 2016);
- ψ^* : representa o complexo conjugado da função wavelet escolhida.

Diferentemente da Transformada de Fourier, a CWT permite uma multiresolução: janelas longas para baixas frequências e janelas curtas para altas frequências. (DOMINGUES et al., 2016).

Entre as diversas famílias de wavelets, a Wavelet de Morlet destaca-se na análise de sinais oscilatórios biológicos. Ela é definida como uma onda plana modulada por uma envólucra Gaussiana (DOMINGUES et al., 2016):

$$\psi(t) = \exp\left(-\frac{t^2}{2\sigma^2}\right) \exp(i\omega_0 t) \quad (2.5)$$

em que σ controla o decaimento (resolução temporal) e ω_0 a frequência central. Por ser uma função complexa, a Morlet permite extrair tanto o módulo (energia) quanto a fase (sincronia) do sinal decomposto (DOMINGUES et al., 2016). O resultado dessa transformação é visualizado através do Escalograma, uma representação bidimensional da densidade de energia do sinal no plano tempo-escala (DOMINGUES et al., 2016).

2.4 Algoritmos de Aprendizado de Máquina e Deep Learning

A seleção dos algoritmos para detecção de crises epilépticas baseia-se na capacidade de cada modelo em lidar com as características intrínsecas do sinal de EEG: a não-estacionariedade, a alta dimensionalidade e a dependência temporal. A literatura classifica estas abordagens em aprendizado supervisionado, não supervisionado e, mais recentemente, aprendizado profundo (*Deep Learning*), capaz de analisar dados em larga escala de forma inteligente (SARKER, 2021). A seguir, fundamenta-se o funcionamento de cada técnica utilizada.

2.4.1 Redes Neurais Convolucionais Unidimensionais (1D-CNN)

Embora as Redes Neurais Convolucionais (CNNs) tenham raízes históricas no Neocognitron de 1979 e sejam tradicionalmente associadas ao processamento de imagens (dados 2D) (SCHMIDHUBER, 2015), sua variante unidimensional (1D-CNN) é projetada especificamente para processar sequências temporais, como o sinal bruto de EEG.

Funcionamento:

A 1D-CNN opera através de camadas de *convolução*, onde filtros (ou *kernels*) deslizam ao longo do eixo do tempo do sinal de entrada. Essa arquitetura permite a detecção automática de características importantes sem intervenção manual excessiva (SARKER, 2021). Matematicamente, para um sinal de entrada $x(t)$ e um filtro $w(\tau)$, a operação de convolução discreta gera um mapa de características $c(t)$:

$$c(t) = (x * w)(t) = \sum_{\tau} x(t - \tau) \cdot w(\tau) \quad (2.6)$$

Diferente das redes neurais tradicionais que tratam cada ponto de dados isoladamente, a convolução preserva a relação espacial local entre pontos vizinhos. Isso permite

que a rede aprenda a reconhecer padrões morfológicos locais característicos da epilepsia, como espículas (*spikes*) e ondas agudas, independentemente de onde eles ocorrem na janela de tempo. O uso de pesos compartilhados nestas camadas reduz significativamente a complexidade descritiva da rede (SCHMIDHUBER, 2015).

Após a convolução, camadas de *pooling* (subamostragem), como o *Max-Pooling*, são aplicadas para reduzir a dimensionalidade dos dados, mantendo apenas as ativações mais relevantes e tornando o modelo invariante a pequenas translações temporais (SCHMIDHUBER, 2015; SARKER, 2021). Estudos recentes, como o de (WANI; KARKI, 2024), demonstram que modelos CNN puros podem atingir acurácias de até 97% em dados transformados.

2.4.2 Arquiteturas Híbridas: CRNN e Bi-LSTM

Enquanto as CNNs são excelentes na extração de características espaciais, elas podem carecer de mecanismos para interpretar dependências temporais de longa duração. Para resolver isso, utilizam-se arquiteturas híbridas que acoplam a CNN a Redes Neurais Recorrentes (RNN), especificamente as *Long Short-Term Memory* (LSTM).

Funcionamento da Bi-LSTM:

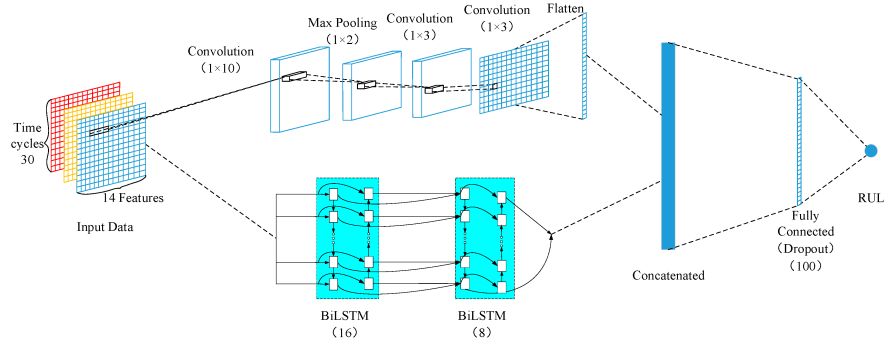
A LSTM tradicional foi introduzida para resolver o problema fundamental do Deep Learning: o desaparecimento ou explosão do gradiente (*vanishing/exploding gradient*) que ocorre em RNNs convencionais ao tentar aprender longos intervalos de tempo (SCHMIDHUBER, 2015). A LSTM soluciona isso através de uma unidade de memória chamada *Constant Error Carousel* (CEC), que mantém o erro estável, e é controlada por portões (*gates*) não-lineares: portão de entrada, portão de esquecimento e portão de saída (SCHMIDHUBER, 2015; SARKER, 2021).

A evolução para **Bi-LSTM (Bidirectional LSTM)** processa a sequência de entrada em duas direções simultaneamente, uma técnica essencial para tarefas de reconhecimento de sequência (SCHMIDHUBER, 2015):

1. **Camada Forward:** processa o sinal do passado para o futuro ($t \rightarrow t + n$).
2. **Camada Backward:** processa o sinal do futuro para o passado ($t + n \rightarrow t$).

Isso é crucial na análise de epilepsia, pois o estado de um ponto no sinal é melhor compreendido se o modelo tiver acesso tanto ao contexto pré-ictal quanto ao pós-ictal. Essa capacidade de processar dados sequenciais torna a LSTM (e suas variantes) distinta das redes *feed-forward* convencionais (SARKER, 2021). (WANI; KARKI, 2024) demonstraram que a fusão de CNN com Bi-LSTM eleva a sensibilidade de detecção para níveis acima de 99,8%.

Figura 6 – Infográfico do funcionamento de uma rede CNN com camada Bi-LSTM.



Fonte: (ZHAO et al., 2020)

A figura 6 mostra uma representação de uma rede CNN com camada Bi-LSTM. Na entrada, o sinal puro, que depois passa por camadas de Convolução e extrai as features mais significativas. Depois, segue para uma camada de Max Pooling, para a redução de dimensionalidade do sinal, conforme as características extraídas na convolução.

Na parte inferior da figura, temos a rede Bi-LSTM, que olha para o sinal no passado e no futuro, processando-o de forma convencional e também de trás para frente.

Por fim, temos a rede neural que vai receber o processamento resultante das camadas e finalizar o processo de classificação.

2.4.3 K-Nearest Neighbors (KNN)

O algoritmo **K-Nearest Neighbors (KNN)** é um método de aprendizado supervisionado baseado em instâncias (*instance-based learning*), também conhecido como "aprendizado preguiçoso" (*lazy learning*), pois não constrói um modelo interno geral durante o treinamento, apenas armazena os dados (SARKER, 2021).

A classificação de uma nova amostra de EEG (um vetor de características x_q) ocorre em tempo de execução comparando-a com os vetores armazenados. O algoritmo calcula a distância entre a nova amostra e os pontos vizinhos. Embora diversas métricas possam ser usadas, a Distância Euclidiana é a mais comum (SARKER, 2021):

$$d(x_q, x_i) = \sqrt{\sum_{j=1}^n (x_{qj} - x_{ij})^2} \quad (2.7)$$

Onde x_{qj} e x_{ij} representam a j -ésima característica da amostra de consulta e da amostra de treino, respectivamente. O processo segue os passos:

1. Calcular a distância entre a nova amostra e todos os pontos armazenados.
2. Identificar os k vizinhos mais próximos.

3. Definir a classificação por votação majoritária desses vizinhos (SARKER, 2021).

Apesar de sua simplicidade, o KNN é robusto em dados ruidosos, servindo como um excelente comparativo de base (*baseline*) (SARKER, 2021).

2.4.4 XGBoost e Métodos de Ensemble

O *Extreme Gradient Boosting* (XGBoost) é um sistema de *Tree Boosting* escalável que tem alcançado resultados de estado da arte em diversos desafios de aprendizado de máquina (CHEN; GUESTRIN, 2016). Ele pertence à família de algoritmos de *Ensemble Learning*, combinando múltiplos modelos fracos (árvores de decisão) para criar um preditor forte.

Funcionamento e Demonstração Matemática:

O XGBoost otimiza um objetivo regularizado para equilibrar a precisão preditiva e a complexidade do modelo, evitando *overfitting*. Para um conjunto de dados com n exemplos e m características, o modelo utiliza K funções aditivas para prever a saída $\hat{y}_i$ (CHEN; GUESTRIN, 2016):

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i), \quad f_k \in \mathcal{F} \quad (2.8)$$

Em que $\mathcal{F}$ é o espaço das árvores de regressão. O objetivo a ser minimizado ($\mathcal{L}$) inclui uma função de perda diferenciável l e um termo de regularização Ω :

$$\mathcal{L}(\phi) = \sum_i l(\hat{y}_i, y_i) + \sum_k \Omega(f_k) \quad (2.9)$$

O termo de regularização $\Omega(f)$ é definido por (CHEN; GUESTRIN, 2016):

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \|w\|^2 \quad (2.10)$$

Por fim, temos que T é o número de folhas na árvore e w são os pesos das folhas. Esta regularização é uma vantagem crítica do XGBoost sobre métodos tradicionais de *Gradient Boosting*, penalizando modelos excessivamente complexos (CHEN; GUESTRIN, 2016). Além disso, o algoritmo utiliza aproximações de segunda ordem (séries de Taylor) para otimizar rapidamente a função objetivo e implementa um algoritmo de busca de divisão consciente de esparsidade (*sparsity-aware*), o que é particularmente útil para lidar com valores ausentes ou dados esparsos comuns em cenários reais (CHEN; GUESTRIN, 2016).

3 Desenvolvimento

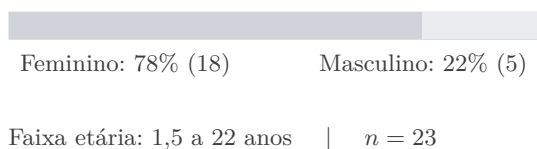
Este capítulo tem como objetivo apresentar todo o processo de desenvolvimento do trabalho, começando no tratamento dos dados da base CHB-MIT até processo de classificação utilizando algoritmos de CNN, KNN e XGBoost.

3.0.1 Base de Dados CHB-MIT

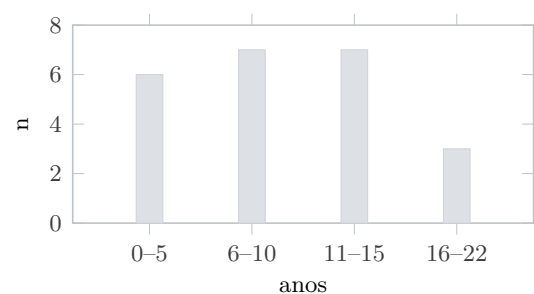
Foi utilizada a base pública *CHB-MIT Scalp EEG Database*, coletada no *Children's Hospital Boston*. Conforme a figura 7 apresenta em gráficos de distribuição por gênero, por idade e por faixa etária, a base é composta por registros eletroencefalográficos de sujeitos pediátricos com crises intratáveis, monitorados para avaliação de intervenção cirúrgica após a retirada de medicação anticonvulsivante. O conjunto de dados original agrupa 23 casos provenientes de 22 sujeitos, estratificados demograficamente em 5 indivíduos do sexo masculino (com idades entre 3 e 22 anos) e 17 do sexo feminino (com idades entre 1,5 e 19 anos) (GUTTAG, 2010). Cada registro contém sinais multicanais adquiridos a uma taxa de amostragem de 256 Hz com resolução de 16 bits, acompanhados de anotações temporais precisas sobre o início e o fim dos eventos de crise epilética.

Figura 7 – Perfil demográfico dos pacientes do conjunto CHB-MIT.

Distribuição por gênero



Distribuição por idade (faixas)



Resumo: média $\approx 9,6$ anos; mediana = 9 anos; mais de 85% dos pacientes possuem menos de 15 anos.

Distribuição por faixa etária

Faixa etária (anos)	n
0-5	6
6-10	7
11-15	7
16-22	3

Fonte: CHB-MIT Scalp EEG Database.

A rotulagem dos eventos epiléticos na base é realizada de forma temporal e precisa, em nível de arquivo. O conjunto disponibiliza o arquivo **RECORDS**, que lista todos os 664

registros disponíveis, e o arquivo `RECORDS-WITH-SEIZURES`, que identifica os 129 arquivos que contêm pelo menos um evento de crise.

Para cada arquivo listado em `RECORDS-WITH-SEIZURES`, existe um arquivo auxiliar de anotação com extensão `.seizure`, no qual o início e o término de cada crise são explicitamente marcados por colchetes, indicando os intervalos temporais exatos em que ocorre atividade ictal. Adicionalmente, os arquivos `chbXX-summary.txt` fornecem informações complementares sobre a montagem dos canais e registram, em segundos, o tempo decorrido desde o início do arquivo `.edf` até o início e o fim de cada crise anotada.

Esse esquema de anotação permite que a rotulagem seja realizada de forma amostral ou por janelas temporais, possibilitando a segmentação dos sinais em trechos rotulados como ictais (crise) ou não ictais (normal). Assim, embora a base não forneça rótulos explícitos por amostra, as anotações temporais permitem a geração controlada de rótulos binários a partir da interseção entre as janelas de EEG e os intervalos de crise descritos nos arquivos de anotação.

3.0.2 Curadoria dos canais e pré-processamento

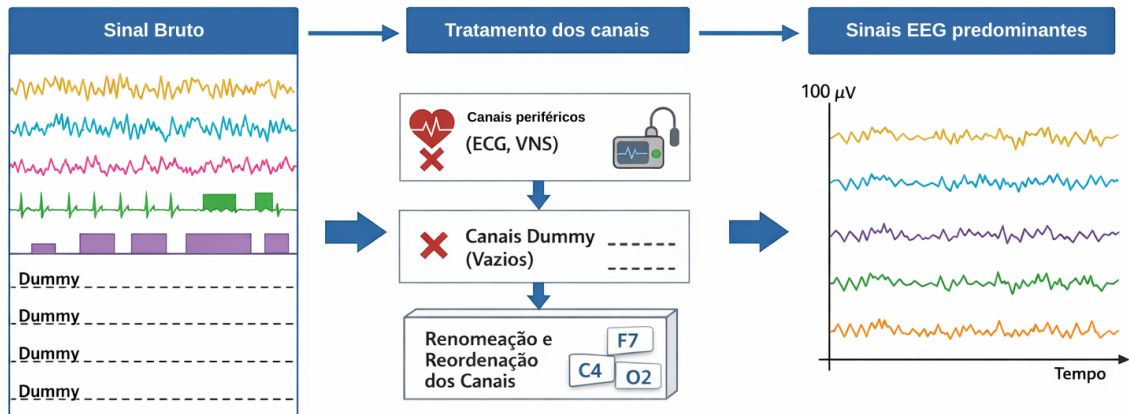
Na etapa inicial do pipeline, foi realizada a curadoria dos sinais eletroencefalográficos, com o objetivo de garantir consistência estrutural e relevância fisiológica dos dados utilizados no treinamento dos modelos de classificação.

Embora a base forneça registros multicanais adquiridos a uma frequência de amostragem de 256 Hz, nem todos os canais presentes nos arquivos são pertinentes à análise de crises epiléticas. Além dos canais de EEG, diversos registros incluem sinais fisiológicos periféricos, como ECG (eletrocardiograma), sinais associados a dispositivos de estimulação do nervo vago (VNS), bem como canais *dummy*, inseridos apenas para fins de formatação e alinhamento visual dos traçados. Esses canais não carregam informação relevante para a tarefa de detecção de crises e, portanto, foram descartados.

A curadoria dos canais foi realizada por meio de um processo de *Data Curation*, no qual apenas um subconjunto previamente definido de canais de EEG foi mantido. Para assegurar a padronização entre diferentes registros e pacientes, os canais selecionados foram identificados utilizando correspondência por expressões regulares, permitindo a inclusão de variações nominais existentes nos arquivos originais. Em seguida, os canais foram renomeados de forma consistente, garantindo uma ordem e nomenclatura uniformes ao longo de toda a base de dados. Registros que não continham todos os canais requeridos foram automaticamente descartados do pipeline.

Essas etapas podem ser visualizadas na figura 8, que mostra o fluxo partindo do Sinal Bruto, seguindo para o Tratamento dos canais, e por fim, para os sinais EEG predominantes.

Figura 8 – Curadoria dos sinais para remoção de artefatos descartáveis



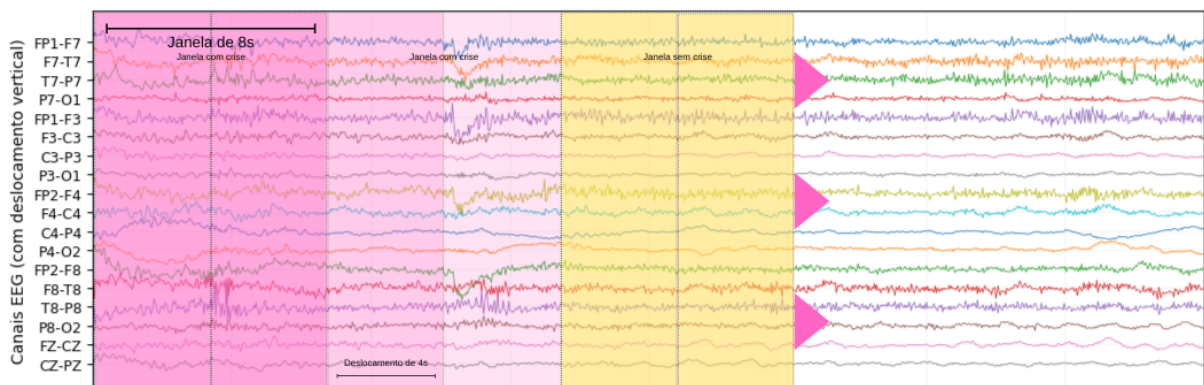
Fonte: elaboração do autor com base nos dados coletados

3.0.3 Janelas deslizantes

Após a curadoria dos canais, os sinais contínuos são segmentados em janelas temporais deslizantes de 8 segundos, com um deslocamento de 4 segundos entre janelas consecutivas, resultando em uma sobreposição de 50%. Cada janela é rotulada de acordo com a presença ou ausência de atividade epiléptica, utilizando as anotações temporais fornecidas pela base. Uma janela é considerada positiva (crise) caso contenha qualquer fração de amostras associadas a um evento epiléptico.

É possível ver como isso funciona na figura 9, em que temos os canais (representados por linhas) e cada etapa de janela (representado por colunas). A janela de 8 segundos (em rosa) pode ou não possuir eventos de crise, com base nas informações do rótulo. Se houver, podemos classificar como crise (1). Uma janela sem crise foi representada em amarelo, sendo tratada como (0). Desta forma, podemos deslizar sobre essas janelas a cada 4 segundos e classificar janela por janela, até que alcancemos o final da janela.

Figura 9 – Janela deslizante em um sinal amostrado da base CHB-MIT



Fonte: elaboração do autor com base nos dados coletados

3.1 Processamento e Organização dos Dados

Devido ao acentuado desbalanceamento natural da base de dados, aplicou-se uma estratégia de subamostragem das janelas negativas. Inicialmente, todas as janelas contendo crises foram preservadas, enquanto selecionou-se aleatoriamente uma fração das janelas sem crise para compor o *pool* inicial. Este conjunto consolidou-se em uma estrutura tensorial de dimensões (9505, 18, 1024), servindo como base para a geração das partições de treinamento e teste.

Posteriormente, os dados foram submetidos a um algoritmo de partição para a implementação de uma validação cruzada baseada na geração de 20 *folds* aleatórios. A estruturação de cada partição seguiu uma lógica de reamostragem independente através de um gerador de números aleatórios global (*default_rng*), garantindo que, para cada *fold*, selecionou-se a totalidade das janelas de crise ($n = 2581$) e uma subamostragem distinta das janelas normais disponíveis no *pool* ($N_{pool} = 6924$).

Este processo assegura um cenário de decisão equilibrado para cada experimento, conforme definido pela Equação 3.1:

$$\mathcal{T}_f \in R^{N_f \times C \times W} \quad (3.1)$$

onde $N_f = 5162$, composto por 50% de amostras da classe ictal e 50% da classe não ictal.

A robustez da metodologia é validada pela análise de similaridade entre *folds* consecutivos, que apresentou uma média estável de aproximadamente 37,3%. Este índice de interseção não é aleatório, sendo o resultado estatístico direto da probabilidade de colisão entre dois sorteios aleatórios sem reposição de tamanho n em um universo N_{pool} . A interseção esperada (E) é definida pela Equação 3.2 (ROSS, 2018):

$$E = \frac{n^2}{N_{pool}} \quad (3.2)$$

Aplicando os valores reais do diagnóstico do *dataset*, onde $n = 2581$ e $N_{pool} = 6924$, obtém-se:

$$E = \frac{2581^2}{6924} = \frac{6.661.561}{6924} \approx 962,1 \quad (3.3)$$

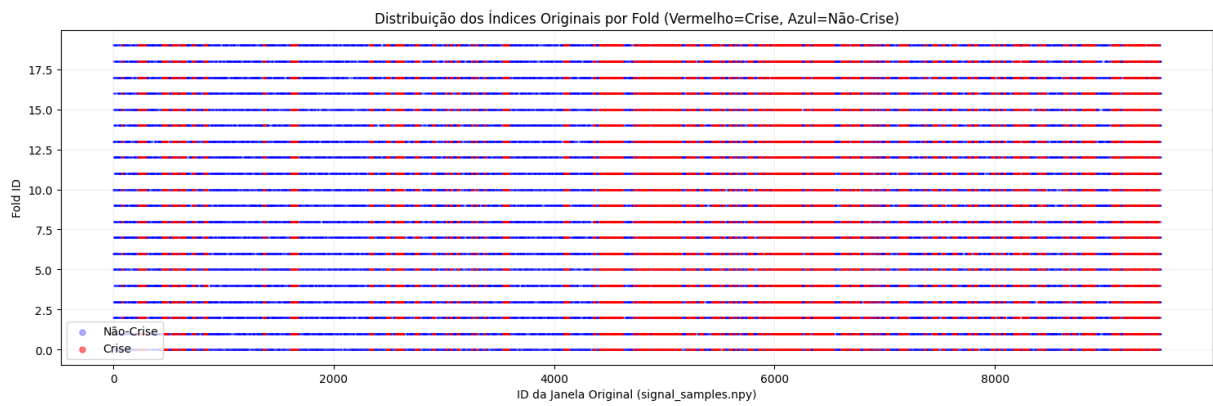
Consequentemente, a similaridade esperada (S) entre a classe não ictal de dois *folds* sucessivos é dada pela razão entre a interseção calculada e o número total de amostras sorteadas por classe:

$$S = \frac{E}{n} = \frac{962,1}{2581} \approx 37,27\% \quad (3.4)$$

A convergência dos resultados experimentais para este patamar teórico confirma que o algoritmo de geração operou em sua capacidade máxima de variabilidade estocástica. Tal rigor metodológico assegura que, enquanto a sensibilidade é preservada pela manutenção integral dos eventos ictais, a especificidade do modelo é exaustivamente testada contra diferentes combinações de atividades interictais, mitigando o risco de sobreajuste (*overfitting*) e garantindo a robustez estatística dos classificadores.

Note que a figura 10 mostra a separação de uma amostra de folds. Em azul, temos as janelas sem crise. Em vermelho, são representadas as janelas de crise.

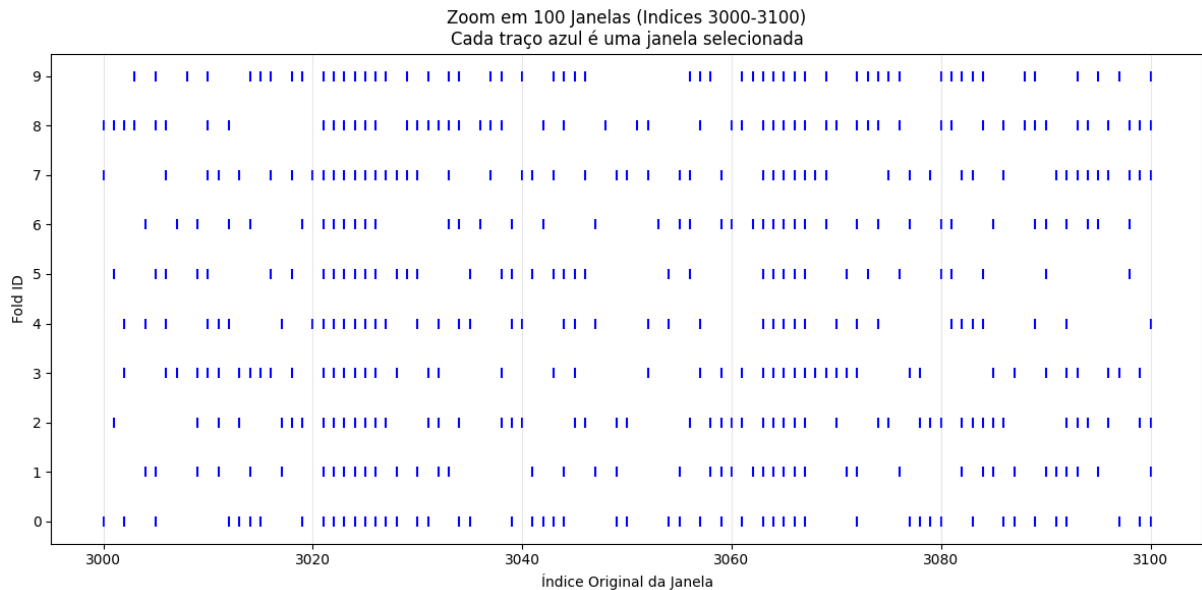
Figura 10 – Gráfico que mostra a separação dos Folds.



Fonte: elaboração do autor com base nos dados coletados

Podemos obter essa separação aproximando em 100 janelas para melhor visualização, conforme apresentado na figura 11.

Figura 11 – Gráfico que mostra a separação dos Folds para 100 janelas



Fonte: elaboração do autor com base nos dados coletados

Para cada janela de cada Fold aleatório gerado, existe um arquivo separado apontando o índice exato da mesma janela na base original e desbalanceada. Esse cuidado foi necessário para que, ao validar a eficiência do modelo com a base original, seja necessário comparar janela por janela, o que se faz com o mesmo índice esperado.

3.1.1 Arquitetura e Treinamento da Rede Neural Convolutacional (CNN)

A primeira abordagem de aprendizado profundo implementada baseou-se em uma arquitetura de Rede Neural Convolutacional unidimensional (CNN-1D), projetada para a extração automática de características morfológicas dos sinais de EEG. A escolha por convoluções 1D fundamenta-se na natureza temporal dos sinais, permitindo que o modelo aprenda padrões locais de descargas epileptiformes ao longo do eixo temporal de cada canal.

A estrutura da rede foi composta por três blocos convolucionais progressivos, conforme detalhado a seguir:

- **Bloco 01:** camada de convolução com 32 filtros e *kernel* de tamanho 64, visando a captura de oscilações de baixa frequência e padrões amplos no sinal.
- **Bloco 02:** camada de convolução com 64 filtros e *kernel* de tamanho 32, refinando a extração para componentes de frequência intermediária.
- **Bloco 03:** camada de convolução final com 128 filtros e *kernel* de tamanho 16, focada em transições rápidas e espículas (*spikes*).

Cada bloco foi acompanhado por camadas de *Batch Normalization* para acelerar a convergência, *MaxPooling1D* com fator 4 para redução da dimensionalidade espacial e *Dropout* de 0,3 para mitigação de *overfitting*. Após as convoluções, aplicou-se uma camada de *Global Average Pooling* (GAP) para reduzir o mapa de características a um vetor de dimensão fixa, seguida por uma camada densa de 64 neurônios e uma camada de saída com ativação *Sigmoid*.

Para a viabilização computacional do treinamento com os 20 *folds* gerados, utilizou-se a infraestrutura de aceleradores de hardware NVIDIA A100. Foram implementadas três estratégias de otimização de alto desempenho:

1. **Mixed Precision Training:** configuração da política de precisão mista (*float16*), utilizando os *Tensor Cores* da GPU para reduzir o tempo de processamento e o consumo de memória, mantendo a estabilidade numérica na camada de saída em *float32*.

2. **XLA (Accelerated Linear Algebra):** compilação do grafo computacional via *JIT Compile*, otimizando a execução das operações algébricas em nível de *kernel* de hardware.
3. **Pipeline de Dados Otimizado:** utilização da API *tf.data* com *Prefetch* e paralelismo de CPU, garantindo que o carregamento das amostras ocorra simultaneamente ao processamento dos lotes (*batches*) de tamanho 1024 pela GPU.

O treinamento foi conduzido com o otimizador Adam ($lr = 0,001$) e a função de perda *Binary Cross-Entropy*. Implementou-se um mecanismo de *Early Stopping* com paciência de 8 épocas, monitorando a perda de validação para restaurar os melhores pesos encontrados, garantindo a generalização do modelo para dados não vistos durante o ajuste dos parâmetros.

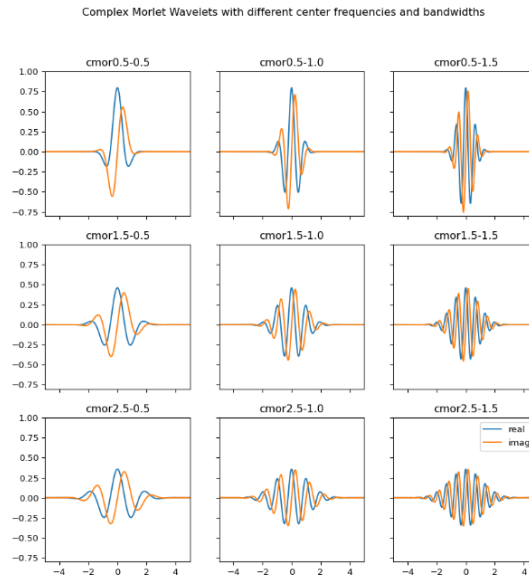
3.1.2 Arquitetura CNN-2D e Pré-processamento via CWT

Para alimentar a arquitetura CNN-2D, implementou-se uma etapa de pré-processamento baseada na Transformada Wavelet Contínua (CWT), visando a extração de características tempo-frequência dos sinais de EEG. Esta abordagem permite capturar a assinatura visual de eventos transientes, como complexos ponta-onda, essenciais para a classificação de crises (SANEI; CHAMBERS, 2007).

A decomposição foi realizada utilizando a função *Wavelet de Morlet Complexa*, selecionada por sua capacidade de preservar a informação de fase e por sua similaridade morfológica com potenciais oscilatórios cerebrais. Conforme definido na fundamentação teórica, a morfologia desta wavelet é controlada pelos parâmetros de decaimento (σ) e frequência central (ω_0). Neste estudo, adotaram-se os valores padrões sugeridos por Domingues et al. (2016) para garantir a condição de admissibilidade e o equilíbrio no plano tempo-frequência:

$$\sigma = 1.5 \quad e \quad \omega_0 = 1.0 \quad (3.5)$$

Figura 12 – Valores apropriados para o decaimento e a largura de banda da wavelet.



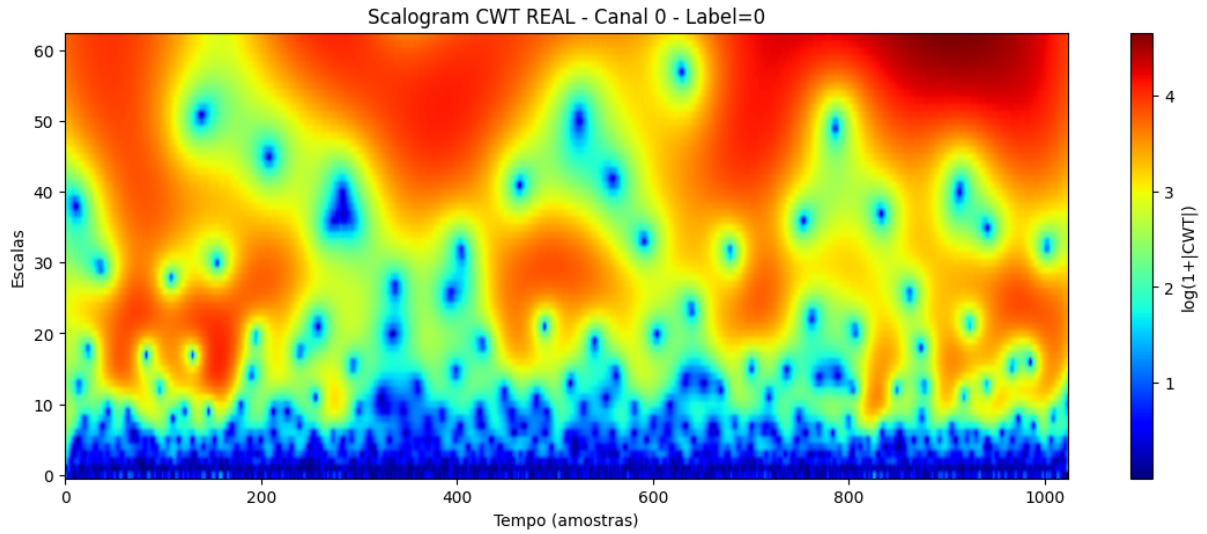
Fonte: PyWavelets <https://pywavelets.readthedocs.io/en/latest/ref/cwt.html>

A figura 12 mostra as diferenças de sinal resultante conforme configuração de decaimento e largura de banda.

A escolha do parâmetro de decaimento ($\sigma = 1, 5$), também referido como largura de banda (*bandwidth*), justifica-se pela necessidade de suavização do envelope gaussiano da wavelet, permitindo que a função capture a natureza não-estacionária das descargas epileptiformes sem introduzir artefatos de borda excessivos. Simultaneamente, a frequência central ($\omega_0 = 1, 0$) estabelece a oscilação base da wavelet, garantindo que o compromisso entre a resolução temporal e a resolução de frequência seja otimizado para as bandas de interesse clínico (Delta, Theta, Alpha e Beta).

O produto desta operação gerou o escalograma para cada segmento de EEG. Utilizou-se o módulo quadrático dos coeficientes wavelet ($|W_\psi f|^2$) para representar a distribuição de energia, resultando em imagens 2D que servem como tensores de entrada para a rede convolucional (WANI; KARKI, 2024; DOMINGUES et al., 2016).

Figura 13 – Escalograma de uma amostra do sinal de EEG.



Fonte: elaboração do autor com base nos dados coletados

A figura 13 mostra uma amostra de escalograma para o sinal no canal 0. Podemos expressar o sinal em termos de potência, o que facilita muito a identificar características que podem passar despercebidas por classificadores.

3.1.2.1 Geração de Escalogramas e Paralelismo de CPU

Dada a alta complexidade computacional da CWT, o processamento dos 20 *folds* foi otimizado através de uma arquitetura de paralelismo em nível de processo (*Process-PoolExecutor*). O sinal bruto de cada janela (1024×18) foi transformado em um tensor tridimensional de dimensões (63, 1024, 18), onde as magnitudes dos coeficientes foram submetidas a uma compressão logarítmica ($\log(1 + x)$) para estabilização da variância e preservação de detalhes em escalas de baixa energia. Os dados resultantes foram persistidos em precisão de meia-largura (*float16*), reduzindo em 50% o consumo de memória RAM e acelerando a transferência para a GPU.

3.1.2.2 Rede neural da CNN-2D e Otimização em Hardware

A rede neural proposta foi projetada para processar os escalogramas como imagens multicanal, tratando cada eletrodo do sistema internacional como uma dimensão de profundidade. A topologia segue uma estrutura convolucional profunda:

- **Camadas Convolucionais:** três blocos com filtros progressivos (32, 64 e 128) de tamanho 3×3 , intercalados com normalização por lote (*Batch Normalization*) e *MaxPooling2D* para redução da redundância espacial;

- **Global Average Pooling:** substituição das camadas densas excessivas pelo GAP 2D, reduzindo drasticamente o número de parâmetros e prevenindo o *overfitting* espectral;
- **Aceleração A100:** o treinamento utilizou a política de *Mixed Precision* (precisão mista), permitindo que os *Tensor Cores* da GPU NVIDIA A100 operassem em lotes de 128 amostras, enquanto a compilação JIT (*Just-In-Time*) via XLA otimizou o grafo computacional.

Para a validação do sistema em cenário de uso real, o modelo campeão foi submetido a uma varredura clínica completa na base original desbalanceada (10,6 horas de registro). Este processo integrou a geração dinâmica de CWT em *batches* com o cálculo de métricas globais, incluindo a análise da Curva ROC e a taxa de Falsos Positivos por Hora (FP/h), permitindo visualizar a convergência entre as crises reais e as previsões temporais através de representações do tipo *Barcode*.

3.1.3 Arquitetura Híbrida C-RNN com Mecanismo de Atenção

Para superar as limitações das redes convolucionais convencionais na modelagem de dependências temporais de longa duração, foi desenvolvida uma arquitetura híbrida denominada C-RNN (*Convolutional Recurrent Neural Network*), integrada a um mecanismo de auto-atenção (*Self-Attention*). Esta abordagem permite que o sistema não apenas identifique padrões morfológicos locais, mas também compreenda a progressão temporal dos eventos críticos.

A arquitetura foi estruturada em quatro estágios fundamentais:

1. **Fatiamento Temporal:** a janela de entrada de 8 segundos ($W = 1024$) é segmentada em 8 sub-blocos independentes de 1 segundo cada. Esta estratégia permite que a rede analise a evolução da crise segundo a segundo.
2. **Extração de Características Espaciais (CNN):** utilizou-se uma camada *TimeDistributed* para aplicar a mesma rede convolucional a cada sub-bloco. Foram empregadas camadas de convolução 1D com ativação ELU (*Exponential Linear Unit*), focadas em extrair vetores de características (*feature maps*) de cada segundo do sinal.
3. **Modelagem Temporal (Bi-LSTM):** os vetores extraídos são processados por uma camada LSTM Bidirecional de 64 unidades. A bidirecionalidade é crucial para que o modelo capture o contexto tanto do início quanto do desenvolvimento da descarga epileptiforme.

4. **Mecanismo de Atenção:** implementou-se uma camada de atenção baseada em *Softmax* para calcular pesos de importância para cada um dos 8 segundos. Isso permite que o modelo priorize segmentos onde as espículas (*spikes*) são mais evidentes, ignorando artefatos de movimento ou ruídos musculares.

3.1.3.1 Refinamento por Mineração de Exemplos Difíceis (*Hard Negative Mining*)

Após o treinamento inicial, observou-se que o modelo apresentava uma taxa bruta de falsos positivos considerável devido a artefatos biológicos similares a crises. Para mitigar este efeito, implementou-se um processo de *Hard Negative Mining* (Mineração de Exemplos Difíceis). O algoritmo realizou uma varredura completa na base original para identificar janelas normais classificadas erroneamente como crises (falsos positivos).

Estes erros foram extraídos e utilizados para compor um novo conjunto de treinamento de ajuste fino (*fine-tuning*), utilizando uma taxa de aprendizado reduzida ($lr = 10^{-5}$). Este procedimento forçou o modelo a aprender as sutis diferenças entre ruídos de alta amplitude e crises reais, elevando a especificidade do sistema.

3.1.4 Arquitetura K-Nearest Neighbors (KNN) e Engenharia de Características

Como base de comparação para os modelos de aprendizagem profunda, foi implementado o algoritmo *K-Nearest Neighbors* (KNN). Por ser um classificador baseado em instâncias, o KNN não possui uma fase de treinamento explícita para ajuste de pesos, realizando a classificação por meio do cálculo de distância entre o vetor de entrada e os k vizinhos mais próximos no espaço de características.

3.1.4.1 Extração de Características Manuais (*Feature Engineering*)

Devido à maldição da dimensionalidade e à sensibilidade do KNN a ruídos ruidosos em sinais brutos, aplicou-se uma extração de características (*handcrafted features*) em cada uma das janelas de 8 segundos. Foram extraídos 8 atributos por canal, totalizando um vetor de 144 dimensões por amostra (considerando a montagem de 18 canais), segmentados em dois domínios:

1. **Domínio do Tempo:** foram calculadas medidas estatísticas de primeira e segunda ordem, incluindo a média aritmética, o desvio padrão e a energia RMS (*Root Mean Square*).
2. **Domínio da Frequência:** utilizou-se o método de Welch para a estimativa da Densidade Espectral de Potência (PSD). A potência absoluta foi integrada nas bandas de frequência clássicas do eletroencefalograma: Delta (0,5–4 Hz), Theta (4–8 Hz), Alpha (8–13 Hz), Beta (13–30 Hz) e Gamma (30–45 Hz).

3.1.4.2 Balanceamento de Classes e Normalização

Dada a natureza desbalanceada do *dataset* original, aplicou-se a técnica de subamostragem aleatória (*undersampling*) na classe majoritária (interictal), equalizando-a à quantidade total de janelas de crise detectadas.

Após o balanceamento, os dados foram submetidos à **Padronização (*Z-score Normalization*)**. Como o KNN utiliza métricas de distância (como a Euclidiana, definida na Equação 3.6), a normalização é mandatória para garantir que atributos com ordens de grandeza distintas (como energia e média) possuam o mesmo peso no cálculo da similaridade.

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (3.6)$$

3.1.4.3 Otimização de Hiperparâmetros via *Grid Search*

A performance do baseline foi otimizada por meio de uma busca em grade (*Grid Search*) com validação cruzada de 5 dobras (*5-fold cross-validation*). O espaço de busca incluiu valores ímpares para o hiperparâmetro k (entre 1 e 31), visando evitar empates na classificação, e o teste das métricas de distância Euclidiana e Manhattan.

O objetivo desta etapa foi maximizar o *F1-Score*, servindo como um marco de desempenho (*benchmark*) para validar se a complexidade adicional das redes neurais recorrentes e convolucionais se justifica frente a um modelo estatístico clássico. A análise da curva de aprendizado (*elbow curve*) permitiu identificar o ponto de saturação onde o aumento de vizinhos deixava de prover ganhos significativos em sensibilidade.

3.1.5 Classificação via Extreme Gradient Boosting (XGBoost)

A arquitetura XGBoost foi implementada como uma alternativa robusta aos modelos lineares, utilizando um algoritmo de *Gradient Boosted Decision Trees* (GBDT). A principal motivação para o uso do XGBoost reside na sua capacidade de realizar a seleção implícita de atributos e fornecer interpretabilidade clínica através do ganho de informação (*feature importance*).

3.1.5.1 Extração de Parâmetros de Hjorth e Potência Espectral

Para representar a morfologia do sinal, foram calculados os descritores estatísticos de Hjorth, que permitem descrever o sinal no domínio do tempo com propriedades interpretáveis no domínio da frequência. A Mobilidade (Equação 3.7) representa a frequência média quadrática, enquanto a Complexidade (Equação 3.8) estima a largura de banda do sinal.

$$Mobilidade(x) = \sqrt{\frac{Var(\frac{dx}{dt})}{Var(x)}} \quad (3.7)$$

$$Complexidade(x) = \frac{Mobilidade(\frac{dx}{dt})}{Mobilidade(x)} \quad (3.8)$$

Os vetores de características foram normalizados logaritmicamente para estabilizar a variância das potências espectrais, resultando em uma matriz de atributos altamente discriminativa para a detecção de crises epiléticas.

3.1.5.2 Interpretabilidade e Importância dos Atributos

Uma vantagem crítica do XGBoost é a extração da importância das *features* baseada no ganho de informação. A análise revelou que os atributos relacionados às bandas **Gamma** e **Beta** figuraram consistentemente no topo da hierarquia de decisão. Tal observação corrobora a literatura neurológica, que associa o início de crises focais a descargas rítmicas de alta frequência.

3.1.5.3 Suavização Temporal e Eficácia Clínica

Para mitigar disparos isolados causados por artefatos de movimento, aplicou-se um filtro de suavização temporal (média móvel) sobre as probabilidades preditas. A otimização do limiar de decisão (*threshold*) para 0,7, combinada com uma janela de consenso de 3 segmentos (12 segundos), permitiu reduzir drasticamente a taxa de falsos positivos, elevando a especificidade para patamares superiores a 95% na base original desbalanceada, validando o modelo para monitoramento contínuo.

3.1.5.4 Validação na Base Real (Janela Deslizante)

Para compreender a eficácia do modelo em um cenário clínico prático, a validação foi executada simulando a passagem do tempo, utilizando a técnica de janela deslizante (*sliding window*). De forma simplificada, o processo consiste em percorrer as 10,6 horas de gravação do eletroencefalograma de forma cronológica, passo a passo, avaliando cada trecho do sinal individualmente, sem misturar ou balancear os dados.

A cada avanço, o sistema realiza uma verificação direta, confrontando a predição do algoritmo com o laudo médico real (gabarito). O fluxo de validação obedece à seguinte lógica interrogativa para cada janela:

1. O modelo analisa o trecho atual do sinal e determina sua predição (*Há uma crise acontecendo aqui?*).

2. O sistema consulta a base real para o exato mesmo instante de tempo (*O especialista também anotou uma crise aqui?*).
3. As respostas são então comparadas para gerar o diagnóstico de acerto ou erro:
 - Se o modelo prediz crise (1) e a base confirma crise (1), contabiliza-se um acerto crítico (**Verdadeiro Positivo**).
 - Se o modelo prediz crise (1), mas a base indica normalidade (0), contabiliza-se um alarme falso (**Falso Positivo**).
 - Se o modelo prediz normalidade (0), mas a base indica uma crise (1), o modelo falhou em detectar o evento (**Falso Negativo**).
 - Se o modelo prediz normalidade (0) e a base confirma normalidade (0), contabiliza-se um acerto de rotina (**Verdadeiro Negativo**).

Após registrar o resultado desse confronto, o algoritmo "desliza" para a próxima janela no tempo e repete exatamente as mesmas perguntas.

Ao final do processamento sequencial das 9.505 janelas, os erros e acertos são somados. É essa contabilidade simples, direta e cronológica que permite calcular métricas operacionais vitais, como a taxa de alarmes falsos acionados a cada hora de monitoramento (FP/h). Esse método garante que os resultados do estudo reflitam fielmente o comportamento que um dispositivo de *hardware* embarcado apresentaria ao lado do leito do paciente.

4 Resultados

A avaliação dos modelos propostos foi conduzida em um cenário de simulação clínica, utilizando uma base de dados desbalanceada composta por aproximadamente 10,6 horas de registros contínuos de EEG (9.505 janelas). O objetivo principal desta etapa foi validar não apenas a capacidade estatística de classificação (Acurácia e AUC), mas, sobretudo, a viabilidade operacional dos algoritmos em dispositivos de monitoramento, onde a taxa de falsos negativos por hora (FN/h) é um fator crítico para a usabilidade.

A Tabela 1 resume o desempenho global de todas as arquiteturas. Observa-se uma clara distinção entre os métodos baseados em aprendizagem profunda (CNNs e CRNN), métodos de conjunto (XGBoost) e o *baseline* clássico (KNN).

Tabela 1 – Comparação final: Acurácia Global e métricas de desempenho clínico ordenadas por arquitetura.

Modelo	Acc. (%)	Sens. (%)	Spec. (%)	AUC	FN (%) ↓	FP (%) ↓
CNN Clássica	95.74	96.78	95.35	0.9912	0.87	3.39
CNN+Wav	97.00	98.53	96.43	0.9958	0.40	2.60
CRNN-Att	97.73	92.87	99.54	0.9967	1.94	0.34
KNN (K=3)	86.93	84.50	89.36	0.9215	7.74	5.32
XGB (Thr=0.60)	99.30	98.72	99.51	0.9995	0.35	0.36

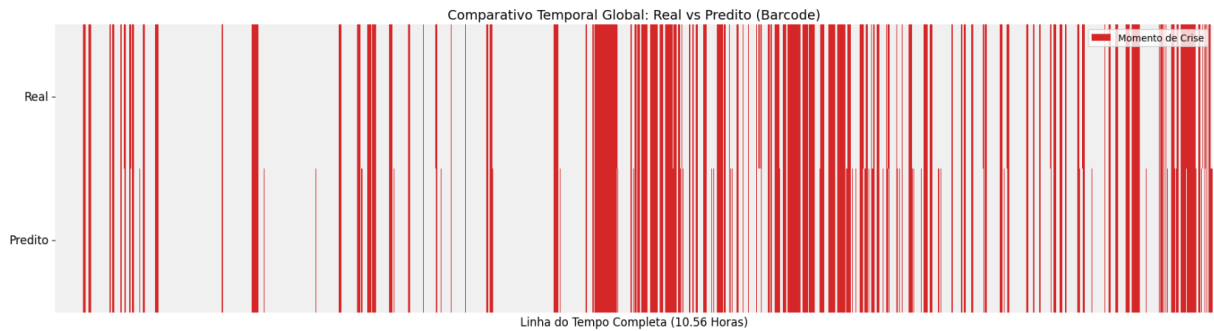
Legenda: Acc. = Acurácia Global; AUC = Área sob a curva ROC; FN = Crises perdidas; FP = Alarmes falsos; CNN Clássica = Modelo convolucional 1D; CNN+Wav = Modelo com Wavelet; CRNN-Att = C-RNN com Atenção; KNN = K-Nearest Neighbors; XGB = XGBoost.

4.0.0.1 CNN Clássica (1D)

A arquitetura CNN Clássica, operando diretamente sobre o sinal bruto no domínio do tempo, demonstrou uma alta capacidade de aprendizado de características, atingindo uma sensibilidade de 96,78%. A Figura 14 apresenta o gráfico de "código de barras" (*barcode plot*), onde a linha superior representa o gabarito (crises reais) e a inferior as predições.

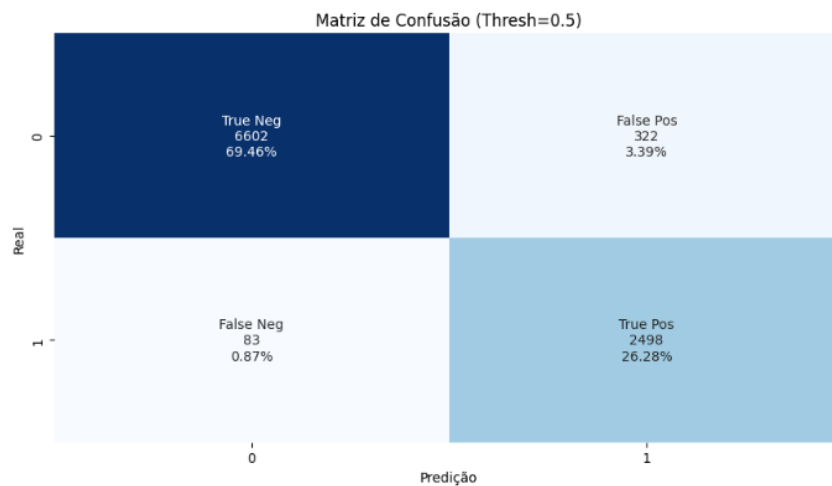
Embora o modelo detecte a grande maioria dos eventos ictais (baixa taxa de falsos negativos), observa-se uma fragmentação nas predições, resultando em uma taxa elevada de 30,49 falsos positivos por hora (FP/h). A matriz de confusão (Figura 15) corrobora este comportamento, evidenciando que, apesar da alta sensibilidade, a especificidade (95,35%) ainda permite a passagem de artefatos que mimetizam crises, como movimentos oculares ou ruídos musculares, comuns na análise de sinal bruto sem pré-processamento espectral.

Figura 14 – Validação amostrada em Barcode para o modelo CNN Clássico.



Fonte: elaboração do autor com base nos dados coletados

Figura 15 – Matriz de Confusão para o modelo CNN Clássico.



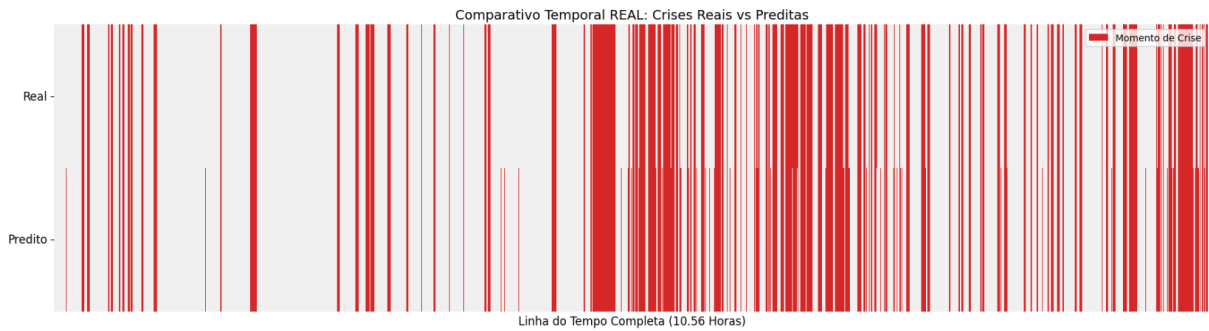
Fonte: elaboração do autor com base nos dados coletados

4.0.0.2 CNN + Transformada Wavelet (CWT)

A incorporação da Transformada Wavelet Contínua (CWT) proporcionou ao sistema a segunda melhor sensibilidade entre todos os modelos testados (98,53%), também com a segunda menor taxa de crises perdidas por hora (3,60 FN/h). A representação tempo-frequência via escalogramas permitiu que a rede identificasse padrões sutis de alta frequência que passaram despercebidos pela CNN 1D.

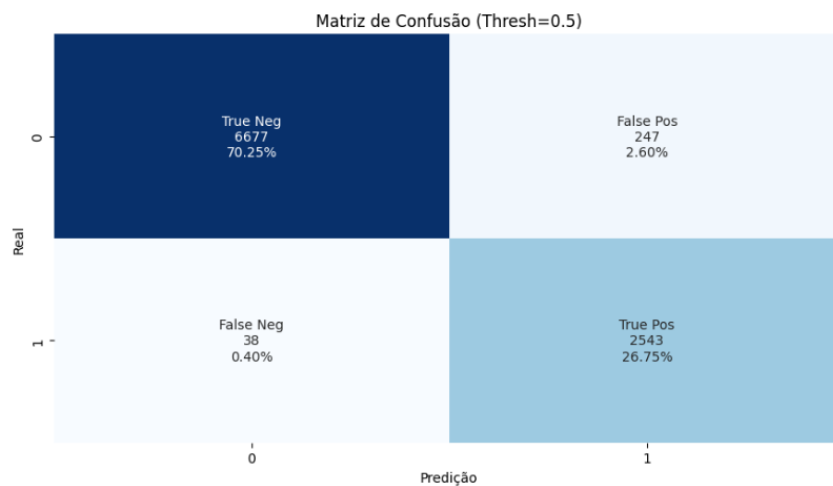
Conforme ilustrado na Figura 16, a detecção dos eventos de crise quando comparado com a base real é robusta e contínua. A CWT também dispõe de alta sensibilidade, decaindo na taxa de alarmes falsos, em relação ao modelo de CNN Clássica, mas mantendo a taxa de alarmes falsos em um patamar elevado (23,39 FP/h). A matriz de confusão (Figura 17) destaca a eficácia do modelo em minimizar os falsos negativos, tornando-o uma escolha ideal para cenários onde a segurança do paciente (não perder crises) é prioridade absoluta sobre o conforto (alarmes falsos).

Figura 16 – Validação amostrada em Barcode para o modelo CNN + Wav



Fonte: elaboração do autor com base nos dados coletados

Figura 17 – Matriz de Confusão para o modelo CNN + Wav



Fonte: elaboração do autor com base nos dados coletados

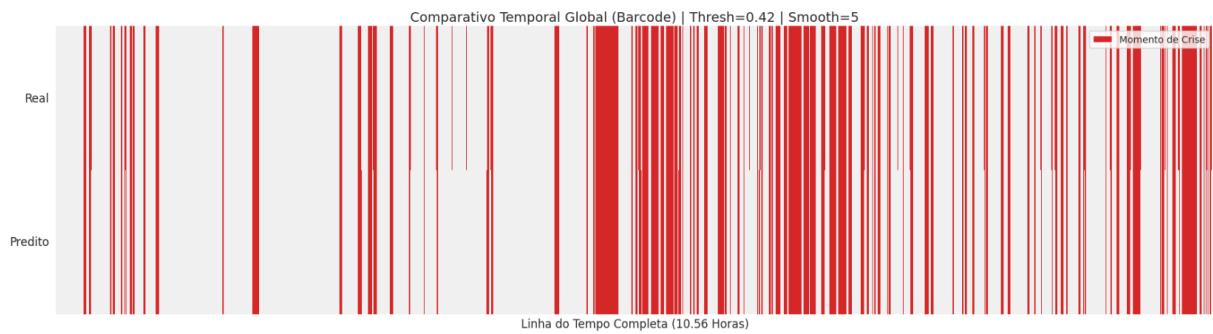
4.0.0.3 C-RNN com Mecanismo de Atenção

A arquitetura híbrida CRNN-Attention marcou uma transição significativa no compromisso entre sensibilidade e especificidade. Ao introduzir a memória de longo prazo (LSTM) e o mecanismo de atenção, o modelo aprendeu a ignorar transientes isolados que não possuem a evolução temporal típica de uma crise.

Isso resultou em uma redução drástica na taxa de falsos positivos, caindo para 3,03 FP/h, um valor dez vezes menor que a CNN Clássica. Entretanto, essa seletividade teve um custo: a sensibilidade recuou para 92,87%. Com a derrubada deste índice, temos um elevado índice de FN/h, o que ascende o alerta em relação aos modelos de CNN Clássica e CNN + Wav.

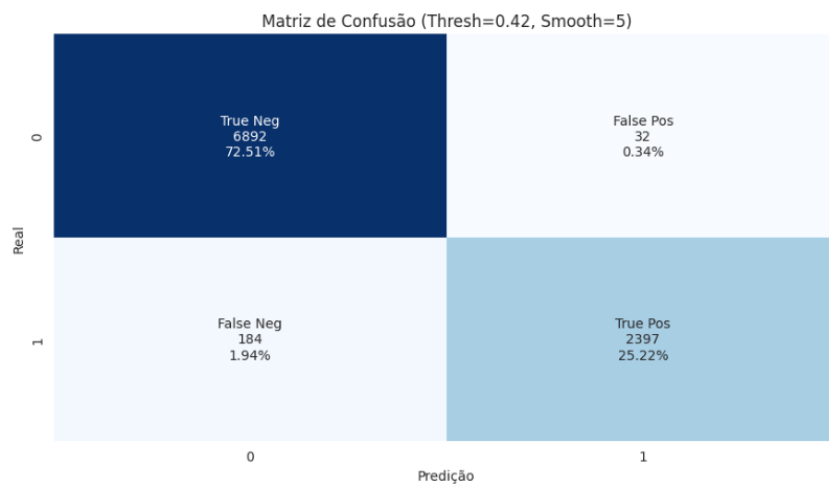
A Figura 18 demonstra isso claramente, quando pode-se notar a ausência de predição na validação com a base real em alguns eventos de crise.

Figura 18 – Validação amostrada em Barcode para o modelo CRNN-Att



Fonte: elaboração do autor com base nos dados coletados

Figura 19 – Matriz de Confusão para o modelo CRNN-Att



Fonte: elaboração do autor com base nos dados coletados

4.0.0.4 K-Nearest Neighbors (Baseline)

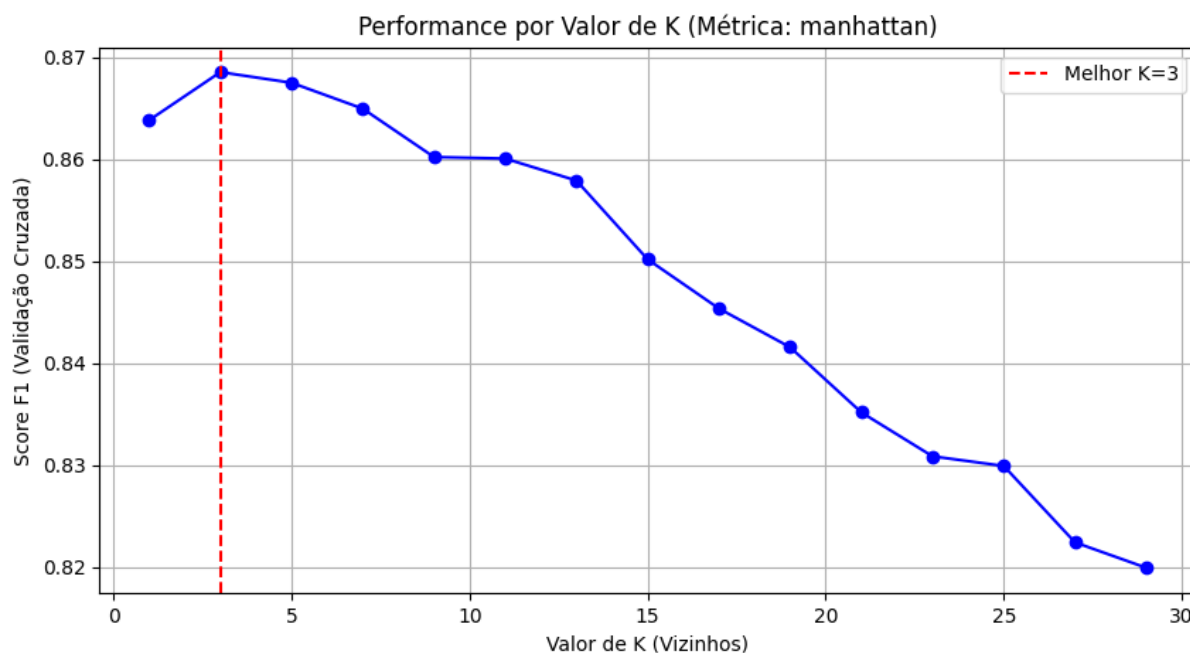
Os resultados do KNN serviram para validar a necessidade de modelos complexos para este problema. Mesmo após a otimização de hiperparâmetros ($K=3$, distância Manhattan), como mostra a Figura 20, o modelo atingiu apenas 84,50% de sensibilidade.

As limitações das features manuais (estatísticas simples e potência de banda) ficaram evidentes nas métricas clínicas: com 47,92 FP/h e uma alta taxa de falsos negativos, o KNN mostrou-se incapaz de separar confiavelmente os padrões ictais de artefatos complexos. A matriz de confusão (Figura 21) revela um número expressivo de erros em ambas as classes, confirmando que a relação não-linear entre os canais de EEG exige a extração hierárquica de características provida pelas redes neurais.

É importante ressaltar que o modelo KNN, após a divisão do conjunto de treino/teste, dispõe apenas de 1.15 horas para validação. Isso significa que os números de FN/h e FP/h foram mensurados dentro deste intervalo de tempo, o que provoca mais atenção para o uso

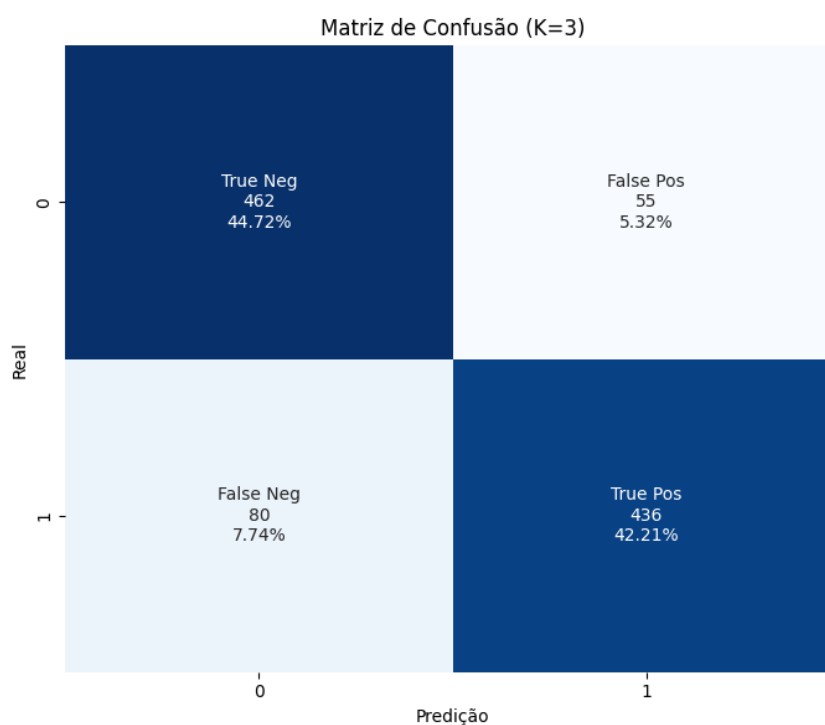
do KNN neste contexto, uma vez que, se tratando das 10.6 horas da base real e mantendo as taxas de erros elevadas, o modelo encontraria ainda mais problemas para detectar sinais de crise.

Figura 20 – Melhor K com base na métrica de Manhattan



Fonte: elaboração do autor com base nos dados coletados

Figura 21 – Matriz de Confusão para o modelo KNN



Fonte: elaboração do autor com base nos dados coletados

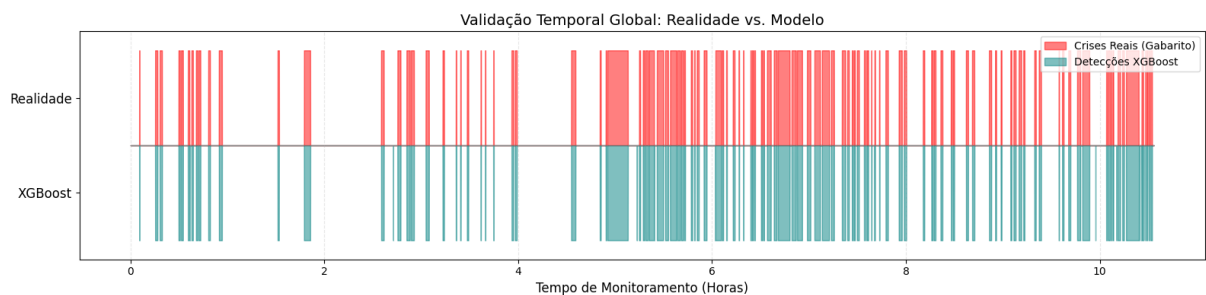
4.0.0.5 XGBoost (Modelo Otimizado)

O modelo XGBoost apresentou o melhor equilíbrio global entre todas as métricas avaliadas, combinando a alta sensibilidade das CNNs com a especificidade da CRNN. Atendendo uma sensibilidade de 98,72% e uma especificidade de 99,51%, o modelo demonstrou a eficácia da engenharia de características avançada.

A Figura 23 ilustra o comportamento da probabilidade de saída durante uma crise: o modelo responde quase instantaneamente ao início do evento ictal, mantendo a confiança alta durante toda a duração. A análise de importância de atributos (Figura 25) revela que o modelo priorizou as bandas de alta frequência (Theta e Delta), que são marcadores fisiológicos conhecidos do início de crises focais, validando a interpretabilidade clínica do algoritmo.

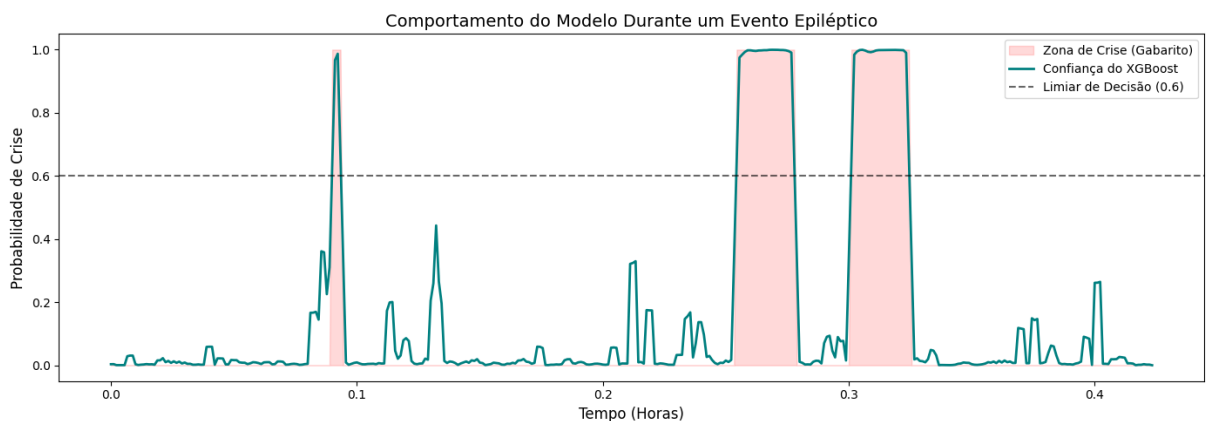
Com uma taxa de alarmes falsos controlada em 3,22 FP/h e a maior acurácia global (99,30%), o XGBoost destacou-se como a solução mais viável para implementação em *hardware* embarcado, dado seu baixo custo computacional na inferência comparado às redes profundas.

Figura 22 – Validação amostrada em Barcode para o modelo XGB



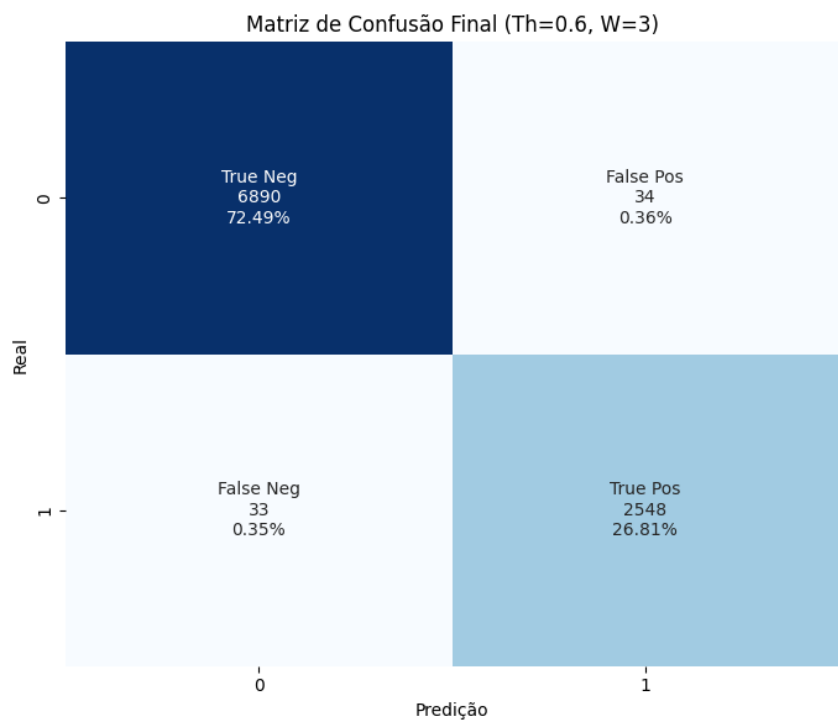
Fonte: elaboração do autor com base nos dados coletados

Figura 23 – Comportamento do modelo durante crise (Probabilidade vs Tempo)



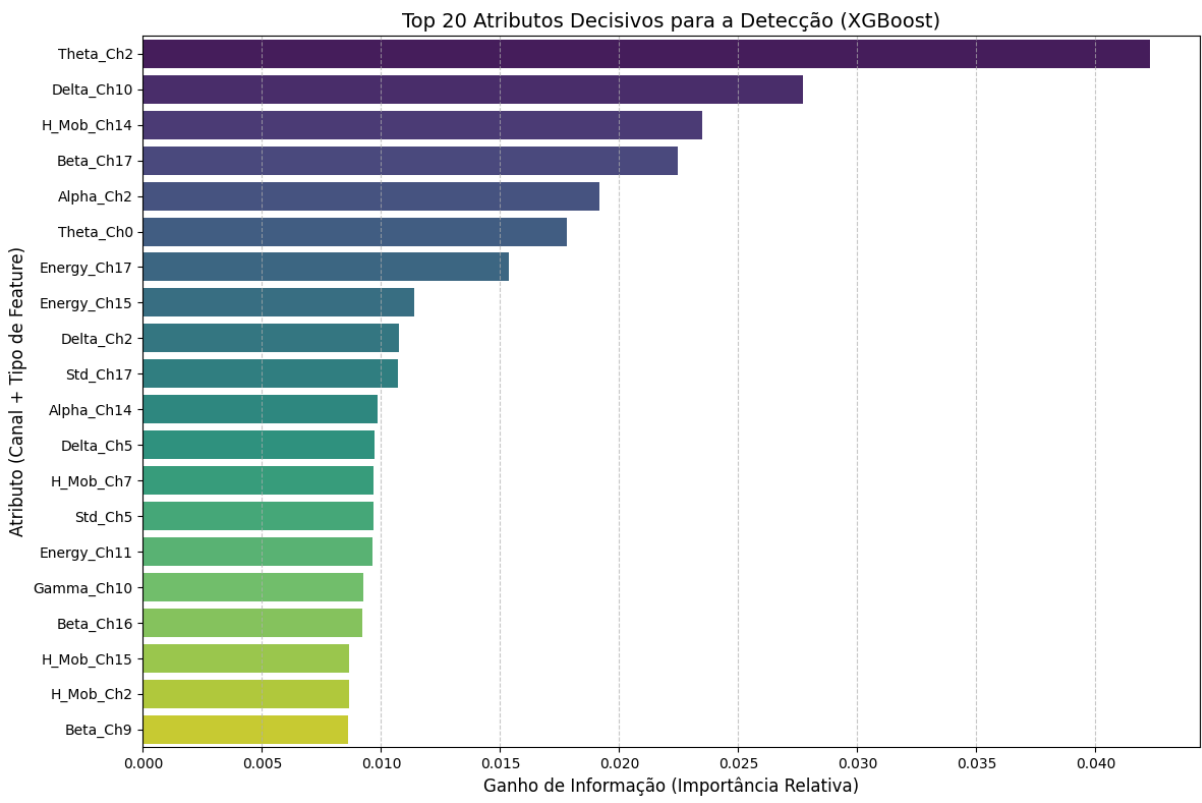
Fonte: elaboração do autor com base nos dados coletados

Figura 24 – Matriz de Confusão para o modelo XGB



Fonte: elaboração do autor com base nos dados coletados

Figura 25 – Importância dos Atributos (Feature Importance) no XGBoost



Fonte: elaboração do autor com base nos dados coletados

5 Conclusão

O presente trabalho dedicou-se à investigação e desenvolvimento de arquiteturas computacionais para a detecção automática de crises epiléticas em sinais de eletroencefalograma (EEG). A motivação central residiu na necessidade de transpor as barreiras dos ambientes controlados de laboratório para cenários de monitoramento contínuo, onde o equilíbrio entre a sensibilidade (segurança do paciente) e a especificidade (redução de alarmes falsos) é o fator determinante para a viabilidade clínica.

A avaliação comparativa entre abordagens clássicas e métodos de aprendizagem profunda revelou que não existe uma solução única perfeita, mas sim compromissos de engenharia distintos. O modelo *baseline* (KNN), embora útil para estabelecer um limite inferior de desempenho, mostrou-se inapto para a complexidade estocástica do EEG, apresentando taxas de erro proibitivas para uso real.

As arquiteturas de *Deep Learning* trouxeram avanços significativos. A introdução da Transformada Wavelet Contínua (CWT) na rede convolucional elevou a sensibilidade a patamares esperados (98,53%), provando que a análise tempo-frequência é superior à análise temporal pura para a captura de padrões ictais sutis. No entanto, a alta sensibilidade veio acompanhada de uma taxa elevada de alarmes falsos, o que poderia gerar fadiga de alarme em cuidadores e pacientes.

Por outro lado, a arquitetura C-RNN com mecanismo de Atenção demonstrou como a modelagem de dependências temporais de longo prazo pode limpar a predição, reduzindo drasticamente os falsos positivos. Contudo, essa filtragem agressiva resultou em uma perda de sensibilidade, elevando o risco de não detecção de eventos críticos.

Desta forma, conclui-se que o modelo XGBoost, otimizado com engenharia de características baseada em parâmetros de Hjorth e potência espectral, representou a solução mais equilibrada e robusta. Ao atingir uma acurácia global de 99,30%, combinada com uma taxa de falsos negativos controlada (3,12 FN/h) e alta eficiência computacional, o XGBoost valida a hipótese de que modelos de conjunto bem ajustados podem superar redes neurais profundas em tarefas com dados tabulares complexos, sendo a escolha ideal para implementação em hardware embarcado de baixo consumo.

Síntese dos Achados

A Tabela 2 sumariza as principais conclusões obtidas a partir dos experimentos realizados e da análise dos resultados.

Tabela 2 – Quadro síntese das conclusões do estudo.

Aspecto avaliado	Conclusão
Relevância do pré-processamento	A aplicação da Transformada Wavelet Contínua (CWT) demonstrou superioridade sobre a análise no domínio do tempo (CNN 1D), permitindo a detecção de padrões oscilatórios de alta frequência que passariam despercebidos, resultando na maior sensibilidade do estudo (98,53%).
Trade-off Sensibilidade vs. Especificidade	Enquanto a C-RNN com Atenção priorizou a especificidade (reduzindo falsos alarmes), a CNN+Wavelet priorizou a sensibilidade. O XGBoost ofereceu o melhor compromisso, maximizando a área sob a curva ROC (AUC 0,9995).
Viabilidade em hardware embarcado	O XGBoost destacou-se não apenas pela performance estatística, mas pela eficiência computacional na inferência. Diferente das redes profundas que exigem aceleração via GPU, o XGBoost é viável para processadores de borda (<i>Edge AI</i>).
Limitações das features manuais	O baixo desempenho do KNN (F1-Score e Sensibilidade < 85%) confirmou que estatísticas simples são insuficientes para capturar a dinâmica não-linear da epilepsia, justificando o uso de descritores avançados como Parâmetros de Hjorth.
Melhor modelo global	O XGBoost consolidou-se como a solução definitiva, apresentando a menor taxa de erro global e a robustez necessária para operar em cenários de monitoramento contínuo de longa duração (10,6 horas).

5.0.1 Trabalhos Futuros

Apesar dos resultados promissores obtidos, especialmente com a arquitetura baseada em XGBoost, o desenvolvimento de um dispositivo vestível de grau clínico apresenta desafios que transcendem o escopo deste trabalho. Com base nas limitações identificadas e nas tendências recentes da literatura, sugerem-se as seguintes direções para pesquisas futuras:

- **Implementação em hardware FPGA ou Microcontroladores de borda:** O próximo passo lógico é a transposição do modelo XGBoost, que demonstrou a melhor eficiência computacional, para um ambiente embarcado real. Recomenda-se a implementação em microcontroladores de baixo consumo (como a família STM32 ou ESP32) ou em FPGAs (*Field-Programmable Gate Arrays*), avaliando o consumo energético em *milliwatts* e a latência de inferência em tempo real, visando a autonomia de bateria necessária para uso ambulatorial prolongado.

- **Validação em bases de dados externas (Cross-Dataset Validation):** Para garantir a generalização do modelo e evitar o *overfitting* às características específicas do conjunto CHB-MIT, sugere-se a validação cruzada com outras bases de dados públicas, como a *Siena Scalp EEG Database* ou a *TUH EEG Corpus* (Temple University Hospital). Isso testará a robustez dos algoritmos frente a diferentes protocolos de aquisição, impedâncias de eletrodos e demografias de pacientes.
- **Exploração de arquiteturas transformers para séries temporais:** O sucesso do mecanismo de Atenção na C-RNN sugere que arquiteturas baseadas inteiramente em Atenção, como os *Transformers* (ex: *Vision Transformers* aplicados a espectrogramas ou *Temporal Fusion Transformers*), podem superar as limitações das redes recorrentes (LSTMs) em capturar dependências de longuíssimo prazo, potencialmente reduzindo ainda mais a taxa de falsos positivos sem sacrificar a sensibilidade.
- **Aprendizado federado (*Federated Learning*):** Considerando a privacidade dos dados médicos, uma abordagem promissora seria o treinamento distribuído, onde o modelo aprende localmente no dispositivo de cada paciente e apenas os pesos (atualizações) são compartilhados com o servidor central. Isso permitiria a criação de um modelo global robusto sem que os dados brutos de EEG jamais saiam do dispositivo do usuário.

Referências

- ANUSHA, K. S.; MATHEWS, M. T.; PUTHANKATTIL, S. D. Classification of normal and epileptic EEG signal using time & frequency domain features through artificial neural network. In: *2012 International Conference on Advances in Computing and Communications (ICACC)*. Cochin, India: IEEE, 2012. p. 98–101. Citado 2 vezes nas páginas 23 e 26.
- CHEN, T.; GUESTRIN, C. Xgboost: A scalable tree boosting system. In: ACM. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. [S.l.], 2016. p. 785–794. Citado na página 31.
- DOMINGUES, M. et al. Explorando a transformada wavelet contínua. *Revista Brasileira de Ensino de Física*, v. 38, 09 2016. Citado 4 vezes nas páginas 27, 28, 38 e 39.
- FATMA, N.; SINGH, P.; SIDDIQUI, M. K. Extensive review of epileptic seizure detection techniques: Performance, achievements, and future directions. *International Journal of Image and Graphics*, v. 23, n. 2550013, 2023. Citado 7 vezes nas páginas 15, 16, 17, 21, 22, 24 e 26.
- GOMES, M. da M. et al. A house-to-house survey of epileptic seizures in an urban community of Rio de Janeiro, Brazil. *Arquivos de Neuro-Psiquiatria*, v. 60, n. 2, p. 209–211, 2002. Disponível em: <<https://www.scielo.br/j/anp/>>. Citado na página 15.
- GUTTAG, J. *CHB-MIT Scalp EEG Database*. 2010. PhysioNet. Version 1.0.0. Disponível em: <<https://doi.org/10.13026/C2K01R>>. Citado 2 vezes nas páginas 20 e 32.
- KLEM, G. H. et al. The ten-twenty electrode system of the international federation. the international federation of clinical neurophysiology. *Electroencephalography and clinical neurophysiology. Supplement*, v. 52, p. 3–6, 1999. Disponível em: <<https://api.semanticscholar.org/CorpusID:27778550>>. Citado na página 21.
- LIU, G. et al. Minireview of epilepsy detection techniques based on electroencephalogram signals. *Frontiers in Systems Neuroscience*, v. 15, 2021. Citado 4 vezes nas páginas 15, 16, 17 e 26.
- ROSS, S. M. *A First Course in Probability*. 10. ed. [S.l.]: Pearson, 2018. Citado na página 35.
- SANEI, S.; CHAMBERS, J. A. *EEG Signal Processing*. Chichester, UK: John Wiley & Sons, Ltd, 2007. ISBN 978-0-470-02581-9. Citado 10 vezes nas páginas 15, 16, 18, 19, 20, 22, 23, 24, 26 e 38.
- SARKER, I. H. Machine learning: Algorithms, real-world applications and research directions. *SN Computer Science*, Springer, v. 2, n. 3, p. 160, 2021. Citado 4 vezes nas páginas 28, 29, 30 e 31.
- SCHMIDHUBER, J. Deep learning in neural networks: An overview. *Neural Networks*, Elsevier, v. 61, p. 85–117, 2015. Baseado no Technical Report IDSIA-03-14 / arXiv:1404.7828. Citado 2 vezes nas páginas 28 e 29.

VETRIKANI, R.; BOBBY, T. C. Diagnosis of epilepsy-a systematic review. In: *3rd International Conference on Biosignals, Images and Instrumentation (ICBSII)*. Chennai, India: IEEE, 2017. Citado 6 vezes nas páginas 15, 17, 19, 20, 21 e 27.

WANI, F. M.; KARKI, M. V. EEG signal analysis for epilepsy detection using deep neural networks. In: *5th International Conference on Circuits, Control, Communication and Computing (I4C)*. Bengaluru, India: IEEE, 2024. Citado 9 vezes nas páginas 15, 16, 17, 18, 21, 24, 26, 29 e 39.

ZHAO, C. et al. A double-channel hybrid deep neural network based on cnn and bilstm for remaining useful life prediction. *Sensors*, v. 20, n. 24, 2020. ISSN 1424-8220. Disponível em: <<https://www.mdpi.com/1424-8220/20/24/7109>>. Citado na página 30.