UNIVERSIDADE DE SÃO PAULO
FFCLRP - DEPARTAMENTO DE FÍSICA
PÓS-GRADUAÇÃO EM FÍSICA APLICADA À MEDICINA E BIOLOGIA

FABRICIO HENRIQUE SIMOZO

**Desenvolvimento de método de análise não linear
do EEG de alta frequência em indivíduos saudáveis
e em pacientes portadores de epilepsia refratária**

Dissertação apresentada à Faculdade de
Filosofia, Ciências e Letras de Ribeirão
Preto da Universidade de São Paulo, como
parte das exigências para a obtenção do
título de Mestre em Ciências, Área: Física
aplicada à Medicina e Biologia.

Ribeirão Preto - SP
2013

FABRICIO HENRIQUE SIMOZO

**Desenvolvimento de método de análise não linear
do EEG de alta frequência em indivíduos saudáveis
e em pacientes portadores de epilepsia refratária**

Dissertação apresentada à Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto da Universidade de São Paulo, como parte das exigências para a obtenção do título de Mestre em Ciências.

Área de Concentração:

Física aplicada à Medicina e Biologia.

Orientador:

Luiz Otávio Murta Junior.

Versão corrigida

Versão original disponível na FFCLRP - USP

Ribeirão Preto - SP

2013

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

FICHA CATALOGRÁFICA

Simozo, Fabricio Henrique

Desenvolvimento de método de análise não linear do EEG de alta frequência em indivíduos saudáveis e em pacientes portadores de epilepsia refratária / Fabricio Henrique Simozo; orientador Luiz Otávio Murta Junior. Ribeirão Preto - SP, 2013.

100 f.:il.

Dissertação (Mestrado - Programa de Pós-graduação em Física aplicada à Medicina e Biologia) - Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto da Universidade de São Paulo, 2013.

1. Eletroencefalografia. 2. Epilepsia. 3. Oscilações de alta frequência. 4. Análise não linear. 5. Processamento de sinais.

Nome: SIMOZO, Fabricio Henrique

Título: Desenvolvimento de método de análise não linear do EEG de alta frequência
em indivíduos saudáveis e em pacientes portadores de epilepsia refratária

Dissertação apresentada à Faculdade de Filosofia,
Ciências e Letras de Ribeirão Preto da
Universidade de São Paulo, como parte das
exigências para a obtenção do título de Mestre em
Ciências.

Aprovado em: ____/____/____.

Banca Examinadora

Prof(a). Dr(a). : _____ Instituição: _____

Julgamento: _____ Assinatura: _____

Prof(a). Dr(a). : _____ Instituição: _____

Julgamento: _____ Assinatura: _____

Prof(a). Dr(a). : _____ Instituição: _____

Julgamento: _____ Assinatura: _____

Agradeço a todos os membros do laboratório CSIM, Antônio Carlos, Juliano, Luiz Eduardo, Isaías, Gabriel e Gustavo, que me receberam com muito companheirismo e sempre auxiliaram nos desenvolvimentos do estudo com um ambiente ao mesmo tempo descontraído e profissional. Em especial, agradeço ao Prof. Dr. Luiz Otávio Murta Júnior, que em muito contribuiu para o trabalho não só com a orientação e ideias de um excelente pesquisador, como também por todo o apoio pessoal dado ao longo destes dois anos.

Agradeço aos membros do CIREP-HCRP, em especial ao Dr. Tonicarlo Rodrigues Velasco, por toda a contribuição dada ao meu conhecimento sobre os assuntos estudados e por todo o tempo e trabalho dedicados ao auxílio no desenvolvimento das ferramentas que seriam utilizadas nas análises.

Agradeço ao Prof. Dr. João Batista Destro Filho, por estar sempre em contato acompanhando e contribuindo com o trabalho, com toda a sua disposição e todo o seu carisma.

É incrível como a formação pessoal tem influência na capacidade de se atuar profissionalmente. A personalidade desenvolvida ao longo dos anos define todas as iniciativas e molda como a solução de cada problema é buscada. Por isso, também dedico este trabalho a meus pais, Marlene e José, a meus irmãos, Juliano e André, meus primos, tios, sobrinhos, avós, e a todos os meus amigos. A influência de seus atos sempre se refletirão em meus trabalhos e conquistas. Sempre serei muito grato pela pessoa que vocês ajudaram a criar.

Cabe também um agradecimento especial à uma pessoa que recentemente entrou em minha vida. Sua confiança, consideração e carinho me acolheram e me deram mais forças do que normalmente teria. Também dedico este trabalho, então, à minha melhor amiga e companheira, Marcela.

“ As oportunidades se multiplicam
à medida que sao agarradas.”

Sun Tzu

RESUMO

SIMOZO, F. H. **Desenvolvimento de método de análise não linear do EEG de alta frequência em indivíduos saudáveis e em pacientes portadores de epilepsia refratária.** 2013. 100 f. Dissertação (Mestrado - Programa de Pós-graduação em Física aplicada à Medicina e Biologia) - Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto - SP, 2013.

Por epilepsia, são denominadas diversas condições neuronais patológicas caracterizadas pela ocorrência de atividade elétrica anormal de uma região do córtex que se difunde formando a crise epiléptica. Tais crises recorrentes possuem efeito relevante sobre a qualidade de vida, além de frequentemente possuírem efeito deletério gradativo sobre a capacidade cognitiva do paciente, sendo necessário tratamento e acompanhamento médico. Dentro das diversas condições epiléticas existe o que é conhecido como epilepsia refratária, que é a condição em que o paciente não responde a nenhum medicamento anticonvulsivante. A alternativa para estes pacientes é a cirurgia. Porém, diversas dificuldades como a localização do foco epiléptico e determinação da extensão da zona epileptogênica podem prejudicar as chances de sucesso deste procedimento. O eletroencefalograma (EEG) sempre foi muito utilizado na clínica neurológica, pois existem padrões de atividade neuronal facilmente observados que podem auxiliar no tratamento de diversas condições, incluindo a epilepsia. Nos últimos anos, vem se demonstrando interesse em cima de um certo padrão que ocorre em frequências antes ignoradas na análise clínica. Estes padrões recebem o nome de oscilações de alta frequência (HFOs). Este estudo teve por objetivo analisar sinais de EEG de pacientes de epilepsia adquiridos

com frequência de aquisição suficientemente alta para a obtenção de informações a cerca da atividade em alta frequência através da detecção automática de HFOs, além da implementação de um método de análise temporal utilizando a Detrended Fluctuation Analysis (DFA) na tentativa de diferenciar características do sinal na transição para o estado ictal e avaliar as aplicações de tais características para prever a ocorrência de crises epiléticas. Os resultados mostram que segmentos distantes da ocorrência de uma crise apresentam média de coeficiente $\alpha_1 = 1.3587$ e $\alpha_2 = 0.6165$, enquanto segmentos durante a ocorrência da crise apresentam coeficientes $\alpha_1 = 1.4061$ e $\alpha_2 = 0.5107$. Apesar das limitações do estudo, os resultados indicam ser possível identificar presença de atividade elétrica anormal, e o método utilizado pode evoluir para auxiliar o diagnóstico e planejamento de cirurgia em epilepsia.

Palavras-chave: 1. Eletroencefalografia. 2. Epilepsia. 3. Oscilações de alta frequência. 4. Análise não linear. 5. Processamento de sinais.

ABSTRACT

SIMOZO, F. H. **Development of nonlinear analysis methods for high frequency EEG in healthy individuals and in patients with refractory epilepsy.** 2013. 100 f. Dissertation (M.Sc. - Postgraduate program in Physics applied to Medicine and Biology) - Faculty of Philosophy, Sciences and Literature, University of São Paulo, Ribeirão Preto - SP, 2013.

By epilepsy, a diversity of pathological neuronal conditions characterized by the occurrences of abnormal electric activity in a cortex region which diffuses causing epileptic seizures. Such recurrent seizures have relevant effects in life quality, and frequently are associated with irreversible damages in cognitive capacity, therefore treatment and medical caring is needed. Inside the classification of epilepsy there is an specific condition known as refractory epilepsy, in which the patient does not respond to anticonvulsant medications. The alternative to these patients is surgery, but the difficulties in determining the exact location of the epileptic focus and the extension of the epileptogenic zone may lead to lower chances of success. The electroencephalogram (EEG) is widely used on neurological clinic, as there are certain patterns of neuronal activity that are easily observed and may help treat a series of pathologies, including epilepsies. Over the last decade, a certain interest has been developed on a specific pattern occurring in frequency bands once ignored on clinical analysis. These patters are called high frequency oscillations (HFOs). The objective of this study was to analyze EEG signals from patients with epilepsy, sampled with a high frequency, so that it is possible to obtain information on high frequency activity through the automatic detection of the HFOs, in addition to the implementation of a nonlinear analysis using Detrended Fluctuation Analysis

(DFA) in an attempt to differentiate signal characteristics during the transition to the ictal state and evaluate the application of said characteristics to predict the occurrence of epileptic seizures. The results shows that segments distant to the occurrence of seizures have a mean value for $\alpha_1 = 1.3587$ and $\alpha_2 = 0.6165$, while segments during the occurrence of seizures have a mean value for $\alpha_1 = 1.4061$ and $\alpha_2 = 0.5107$. Despite study limitations, results indicates it is possible to identify presence of abnormal electric activity and the proposed method may evolve to help epilepsy diagnostic and surgical planning in epileptic patients.

Key-words: 1. Eletroencephalography. 2. Epilepsy. 3. High frequency oscilations. 4. Nonlinear analysis. 5. Signal processing.

LISTA DE FIGURAS

1.1	Sinal de EEG epiléptico em período inter-ictal.	8
1.2	Posicionamento de eletrodos baseado no sistema internacional 10-20. . .	9
1.3	Exemplo de sinal de EEG com início de crise epiléptica.	11
1.4	Exemplo de aumento de atividade de alta frequência em EEG epiléptico. 13	
1.5	Exemplo de segmentações em DFA.	16
1.6	Exemplo do cálculo do expoente através do método da DFA.	17
1.7	Ilustração do efeito de <i>crossover</i> no método da DFA.	18
1.8	Exemplificação de filtros de frequência.	21
1.9	O Fenômeno de Gibbs.	22
1.10	Diferentes janelas e suas respostas em frequência.	26
2.1	Resposta ao impulso e a frequência dos filtros.	33
3.1	Alterações nos espectros de potência de sinal EEG após filtragem. . .	40
3.2	Teste de detecções de HFOs para sinal simulado.	42
3.3	Detecção de HFO com alteração do valor de threshold.	43
3.4	Teste de detecções de HFOs em função do nível de ruído.	45
3.5	Teste de detecções de HFOs em função da taxa de HFO por segundo. 47	
3.6	Médias e desvios do expoente α em função do tamanho do sinal. . . .	50
3.7	Influência do tamanho do segmento nos valores de α em um sinal de EEG.	51
3.8	Exemplos de dados obtidos pelo método da DFA para diferentes pacientes.	56
3.9	Dados de α_2 vs α_1 para os segmentos de EEG epiléptico.	57
3.10	Média e desvio padrao para os dados de α_2 vs α_1 para os segmentos de EEG epiléptico.	58

3.11	Impacto do tamanho da janela nas distribuições de α_1 e α_2	64
3.12	Exemplos de dados obtidos pela detecção de HFOs para diferentes pacientes.	68
3.13	Média e desvio padrao da taxa de ocorrência de detecções de HFO para os segmentos de EEG epiléptico.	69
3.14	Impacto do tamanho da janela nas distribuições de taxa de detecção de HFOs.	72

LISTA DE TABELAS

2.1	Parâmetros da resposta em frequência dos filtros implementados.	32
2.2	Descrição dos grupos para análise estatística dos resultados.	36
2.3	Tabela de especificação das ferramentas utilizadas na construção do software.	37
3.1	Tabela de dados da taxa de detecção em função da amplitude do ruído.	44
3.2	Tabela de dados da taxa de detecção em função da taxa de HFOs por segundo.	46
3.3	Tabela de dados de DFA com sinais simulados.	49
3.4	Dados de média, desvio padrão e teste de significância das distribuições de α_1 e α_2 para todos os sinais de EEG.	60
3.5	Número de segmentos em cada grupo de dados para a análise dos resultados do teste de DFA.	61
3.6	Dados de média, desvio padrão e teste de significância das distribuições de taxa de ocorrência de detecções para todos os sinais de EEG.	70
A.1	Arquivos do banco de dados MIT-CHB utilizados.	85

LISTA DE ABREVIATURAS

EEG	Eletroencefalografia (<i>Eletroencephalography</i>).
HFO	Oscilações de alta frequência (<i>High frequency oscillations</i>).
DFA	Análise de Flutuação Retificada (<i>Detrended Fluctuation Analysis</i>).
MRI	Imageamento por Ressonância Magnética (<i>Magnetic Resonance Imaging</i>).
PET	Tomografia por Emissão de Pósitron (<i>Positron Emission Tomography</i>).
SPECT	Tomografia Computadorizada por emissão de fóton único (<i>Single Photon Emission Computed Tomography</i>).
FIR	Resposta ao Impulso Finita (<i>Finite Impulse Response</i>).
IIR	Resposta ao Impulso Infinita (<i>Infinite Impulse Response</i>).
RMS	Média quadrática (<i>Root Mean Square</i>).

SUMÁRIO

Lista de Figuras	x
Lista de Tabelas	xii
Lista de Abreviaturas	xiii
1 Introdução	1
1.1 Epilepsia	1
1.2 Sinais Biológicos	3
1.2.1 Caracterização de Sinais	4
1.3 Eletroencefalografia	5
1.3.1 Propriedades dos sinais de EEG e não linearidade	6
1.3.2 Montagem de eletrodos e canais - escalpo	7
1.3.3 EEG na epilepsia	10
1.4 Oscilações de alta frequência - HFOs	12
1.4.1 Os padrões de atividade de alta frequência	12
1.5 Objetivos	14
1.6 Detrended Fluctuation Analysis	14
1.7 Filtros Digitais	19
1.7.1 Representação em tempo e frequência	19
1.7.2 Definição de Filtros Digitais	20
1.7.3 Filtros Ideais e Fenômeno de Gibbs	21
1.7.4 Convolução	23
1.7.5 Filtros Digitais FIR e Desenho por Janelamento	24
1.8 Detecção de HFOs	26

2	Metodologia	30
2.1	Sinais utilizados	30
2.1.1	Sobre a frequência de aquisição dos sinais	31
2.2	Planejamento dos Filtros Digitais	31
2.3	Detrended Fluctuation Analysis	34
2.4	Detecção de HFOs	35
2.5	Software Utilizado	36
3	Resultados e discussões	39
3.1	Testes	39
3.1.1	Testes de Filtros	39
3.1.2	Testes de Detecção de HFOs	41
3.1.3	Testes de DFA	48
3.2	Resultados	52
3.2.1	DFA	53
3.2.2	Detecção de HFOs	65
4	Conclusões	73
	Referências Bibliográficas	76
	Apêndice A - Dados Utilizados CHB-MIT	82

INTRODUÇÃO

1.1 Epilepsia

A palavra “epilepsia” vem do grego “*epilepsia*”, significando algo próximo de “tomar posse de”. Este conceito vem da antiga crença de que um indivíduo sofrendo ataques de convulsão poderia estar sendo possuído por deuses ou espíritos malignos [1]. Hipócrates considerava a epilepsia como uma condição que deveria ser tratada, mas as crenças religiosas impediram investigações a respeito até o século XIX [2].

Hoje utiliza-se o termo “epilepsia” para designar não uma doença, mas qualquer condição neurológica crônica que se caracteriza pela predisposição a gerar convulsões. Fisher e colaboradores definem a epilepsia como “uma desordem do cérebro caracterizada por predisposição a gerar crises epiléticas e pelas consequências neurobiológicas, cognitivas, psicológicas e sociais desta condição. A definição de epilepsia requer a ocorrência de ao menos uma crise epilética”, enquanto a definição dada para o conceito de crise epilética é de “uma ocorrência transiente de sinais e/ou sintomas devido à atividade neuronal anormalmente excessiva ou sincronizada do cérebro” [3].

Existem organizações (incluindo a ILAE - *International League Against Epilepsy*, ou Liga Internacional Contra a Epilepsia) que promove eventos e artigos, dentre outros, como formas de conscientizar e informar a população mundial sobre diversos aspectos da doença como causas das crises, necessidades que portadores podem possuir, quais os possíveis tratamentos e até mesmo como agir em um evento de crise [4].

A própria ILAE se propõe a divulgar e atualizar definições para classificação das condições epiléticas. Estas definições são revisadas pela chamada Comissão de Classificação e Terminologia (*Classification and Terminology Commission*), e o mais recente relato da comissão foi divulgado em 2010 [5]. Por se tratar de um termo amplo, as definições para os diferentes tipos de epilepsia, apesar de definidas, são muito discutidas [6]. Alguns autores compartilham da opinião de que, atualmente, começam a surgir um grande número de definições que descrevem condições com pequenas diferenças.

Cerca de 1% da população mundial sofre de alguma condição causadora de crises epiléticas [7], sendo que em 30 a 40 % dos casos ocorre remissão espontânea, 30 a 40% dos casos ocorre remissão após o tratamento com drogas antiepiléticas (AEDs - *Antiepileptic Drugs*) e em 30 a 40% dos casos as crises persistem mesmo sob o efeito de de AEDs. Nesta última condição, encontra-se o que chamamos de epilepsia refratária [8]. A condição em que um paciente não responde à medicamentos é especialmente grave. Estudos mostram que pacientes com epilepsia refratária possuem limitações em aspectos profissionais e sociais e redução na qualidade de vida, devido a diversos danos causados pela epilepsia, tanto neurofisiológicos quanto psiquiátricos [9].

A única intervenção médica capaz de eliminar ou atenuar os sintomas da epilepsia destes pacientes seria a cirurgia. A ideia da intervenção cirúrgica vem das evidências de que haveriam regiões no córtex que seriam causadoras das crises epiléticas, chamadas de zonas epileptogênicas.

Atualmente existem várias ferramentas que podem auxiliar na localização e delimitação da zona epileptogênica. A melhor forma de se efetuar esta localização é através de exames de MRI, quando há lesões visíveis. Ainda assim, o uso de outras técnicas como análise de semiologia, V-EEG e imagens de PET e SPECT é fundamental para a confirmação dos achados em MRI ou, nos casos em que não há lesões visíveis, tentar fornecer alternativas para esta localização [10].

O maior problema encontra-se no fato de que a zona epileptogênica é um conceito teórico que incluiria tanto a região causadora das crises epiléticas quanto todas as outras regiões potenciais causadoras, que poderiam gerar crises quando a primeira fosse removida [11].

A incapacidade de se excluir potenciais regiões causadoras de crises torna a delimitação precisa da zona epileptogênica uma tarefa extremamente complicada. Logo, há a necessidade de novas pesquisas e tecnologias desenvolvidas e testadas para auxiliar tanto a localização espacial e delimitação da zona epileptogênica, como também o diagnóstico e acompanhamento da epilepsia.

Assim como muitas condições neurológicas, a epilepsia reflete alteração no funcionamento das células cerebrais, resultando na alteração da atividade elétrica. Por este motivo, o EEG é um exame muito recomendado para estudos, já que, além do custo dos exames ser relativamente baixo, nos casos em que a alteração na atividade elétrica pode ser visualizada em escalpo, o exame é feito de forma não invasiva.

Ainda não se conhecem de forma detalhada os mecanismos causadores das crises epilépticas. Porém, alguns estudos mostram que, ao contrario do que se pensava há um tempo, as crises podem não ser iniciadas de forma abrupta. Mormann e colaboradores relatam ser possível a observação de alterações na sincronização entre diferentes regiões através de análise do sinal de EEG antes de uma crise [12], além de outras publicações que sugerem que seja possível prever a ocorrência de crises epilépticas [13].

O fato de que seja possível prever uma crise indica que os mecanismos geradores entram em ação antes de que a crise possa ocorrer. Assim, pode ser possível obter informações e parâmetros extraídos de sinais de EEG que ajudem a esclarecer o funcionamento destes mecanismos, podendo futuramente auxiliar no desenvolvimento de novas terapias para epilepsia.

1.2 Sinais Biológicos

Por sinal, entendemos toda sequência de dados que pode ser definida. Os sinais estão por toda a parte, desde os mais óbvios (sinais de celular, televisão, rádio - através de ondas eletromagnéticas) até os menos perceptíveis (dados de tamanho de uma planta em função do tempo, variação do valor do dólar, etc).

Normalmente é definido como sinal biológico a manifestação de atividade

fisiológica de algum órgão que pode ser adquirida através da utilização de sensores especiais (transdutores), capazes de converter os sinais analógicos em uma série de dados digital. Nessa categoria, se enquadra o EEG (Eletroencefalograma), que é o objeto de estudo deste trabalho.

O estudo dos sinais biológicos se mostrou, com o passar do tempo, de grande importância na medicina, seja para diagnóstico, tratamento e acompanhamento de pacientes com diversas condições patológicas.

A necessidade da digitalização do sinal vai muito além da simples adaptação à era tecnológica. É fácil pensar que a conversão dos sinais analógicos para representações digitais seja utilizada para facilitar a visualização e o armazenamento dos dados através do uso de computadores e servidores. Porém, a extração de informações dos sinais pode ser uma tarefa complicada e trabalhosa, sendo as vezes até impossível de ser feita sem o auxílio de métodos de computação. Tendo em vista este aspecto, a digitalização permite que métodos computacionais sejam desenvolvidos para análise e extração automática de características que podem auxiliar a quantificação e qualificação de vários fatores relacionados à geração dos sinais.

1.2.1 Caracterização de Sinais

Dentre as características importantes para a caracterização de um sinal, podem ser citadas as classificações entre linear e não linear, estacionário e não-estacionário, determinístico e estocástico, periódico e transiente.

A linearidade está relacionada à resposta do sinal quando aplicada alguma operação à este sinal. Isso quer dizer que linearidade não é, na verdade, uma classificação de sinal, mas sim de sistema (Um sistema corresponde, basicamente, à um sinal ou conjunto de sinais a que se está sendo aplicada uma operação ou conjunto de operações). O sinal é linear se satisfaz o princípio da superposição, que define que a soma ponderada de sinais de entrada é igual à mesma soma ponderada de sinais de saída.

A estacionariedade está relacionada à características como média, variância, ou outros momentos de ordem superior ou características espectrais mudam com o

tempo. Diversos métodos de processamento utilizados para extrair informações de sinais assumem que a série temporal estudada é estacionária, devendo haver cuidado na hora da aplicação de métodos e da análise dos resultados obtidos. Mais sobre análise de sinais não lineares é discutida em 1.3.1.

A importância na classificação de sinais como lineares ou não-lineares está justamente ligada à quais os métodos de análise (operações) podem ser aplicadas a determinados sinais obtendo-se resultados confiáveis.

A aleatoriedade do sinal está relacionada à natureza intrínseca da geração deste sinal. Um sinal é determinístico quando não há nenhuma incerteza em relação à seu valor em qualquer instante de tempo.

Já um sinal aleatório contém incertezas sobre sua ocorrência. Estes sinais são comumente estudados através de análises probabilísticas.

A periodicidade de um sinal está relacionada a existência de um valor constante positivo T que satisfaz

$$x(t) = x(t + T), \quad (1.1)$$

ou seja, o sinal se repete a cada intervalo de tempo T .

Enquanto um sinal transiente (ou um trecho de sinal transiente) pode ser definido como não sendo periódico. Um trecho ou padrão transiente em um sinal pode até vir a se repetir, mas não obedece a relação de sinal periódico, ou seja, não pode-se definir um período para o surgimento deste padrão.

1.3 Eletroencefalografia

Eletroencefalografia é o nome dado ao exame que consiste, basicamente, da aquisição dos sinais elétricos provenientes do cérebro de um indivíduo ou animal. Esta aquisição é feita através de um certo número de eletrodos, que podem ser posicionados no escalpo (EEG de superfície) ou inseridos diretamente ou à uma certa profundidade no córtex (EEG invasivo). O primeiro relato de medição da atividade elétrica cerebral foi feito por Caton em 1875 [14] e muito se evoluiu desde

então. Hoje, o exame de EEG é uma poderosa ferramenta para auxiliar o diagnóstico de uma série de patologias, síndromes ou condições neurológicas.

Os sinais que são captados por um exame de EEG são potenciais provocados pela atividade pós-sináptica dos neurônios. É importante notar que os sinais captados não são provenientes de apenas um neurônio, mas de um grupo deles.

Para um EEG de superfície, a soma das atividades elétricas dos neurônios próximos a um determinado eletrodo ainda são atenuadas pelas várias camadas de tecidos entre o córtex e o eletrodo. Estas adversidades ocasionam a falta de resolução espacial além de poderem mascarar atividades que seriam mais facilmente observadas no EEG invasivo.

1.3.1 Propriedades dos sinais de EEG e não linearidade

Sinais de EEG mostram comportamentos de não-linearidade e não-estacionariedade muito fortes. Estas características tornam o processamento dos sinais uma tarefa delicada, pois os resultados obtidos da aplicação de métodos de análise que não são recomendados para sinais não-lineares ou não-estacionários podem não caracterizar de fato o sinal.

Alguns autores consideram que a não-linearidade dos sinais é um reflexo direto da complexidade dos sistemas, visto que esta contribui, sendo talvez necessária, para a adaptação dos organismos vivos no meio dinâmico em que são envolvidos [15].

Como já citado na seção 1.2.1, um sistema linear é um sistema que obedece o princípio da superposição. Logo, podem-se definir sistemas não lineares como os sistemas que não satisfazem tal princípio. Muitas vezes, dinâmicas não lineares são associadas a sistemas complexos, mas as definições de sistemas complexos ainda são alvo de algumas discussões e afirmações conflitantes entre diferentes publicações.

A tentativa de extração de informações de um sinal com determinadas propriedades pode ocasionar em resultados aproximados e/ou até mesmo distorcidos pelo fato de que os métodos de análise não se adequam à características que representam o sinal. Ou seja, os métodos podem nem sempre representar a dinâmica que caracteriza o sinal estudado [16].

Portanto, a justificativa para se utilizar análises não lineares está na exigência de se utilizar métodos que reflitam as características do sinal estudado de forma robusta, evitando a interpretação equívoca de resultados que podem não estar corretos.

A motivação para a utilização dos métodos não lineares pode estar na observação de que o sinal possui características que seriam melhor estudadas com tais métodos, ou no conhecimento de que o sistema que tenha gerado tais sinais possui característica complexa. Esta última suposição pode ser perigosa devido ao fato de que não há garantia de que a aquisição de sinais provenientes de sistemas que apresentam características não lineares resultará em sinais com características não lineares [17].

Na análise de sinais de EEG, muitas vezes a justificativa para a utilização de métodos não lineares está nos indícios de existência de atratores estranhos, conceito proveniente da teoria do Caos, e muitas análises foram feitas com sucesso utilizando estimativas para expoentes de Lyapunov, que possui ligação com teoria dos fractais [18].

Diversos métodos não-lineares foram utilizados com sucesso para análise de sinais de EEG, seja na modelagem de sinais ou na obtenção novas informações para auxílio de classificação de patologias [19]. De fato, é comumente aceito que a aplicação de métodos não lineares em análise de sinais de EEG abre novas fronteiras para caracterização de diversas condições e potencialmente pode trazer novas informações a respeito das dinâmicas de geração da atividade elétrica cerebral [18, 20]

Ainda assim, a preocupação com a estacionariedade deve estar sempre presente. Blanco e colaboradores fizeram um estudo sobre a estacionariedade do EEG e citaram alguns problemas decorrentes da aplicação de certos algoritmos a trechos de sinais de EEG não-estacionários [21].

1.3.2 Montagem de eletrodos e canais - escalpo

Os sinais de EEG de escalpo geralmente obedecem a regras de posicionamento de eletrodos. O sistema mais utilizado é chamado de sistema internacional 10-20,

ilustrado na figura 1.2, nome que se refere às relações de distância entre os eletrodos.

Os dados de atividade neuronal coletados pelos eletrodos são utilizados para a montagem dos canais a serem visualizados e analisados. Existem várias montagens que podem ser utilizadas, sendo as mais comuns

- O mesmo eletrodo de referência é utilizado para todos os canais, sendo cada canal o sinal de um eletrodo ativo diferenciado pelo sinal do eletrodo de referência.
- Cada canal é resultado da diferenciação entre dois eletrodos ativos.

Figura 1.1: Sinal de EEG epiléptico em período inter-ictal. Repare os nomes dos canais indicando a montagem a partir dos sinais dos eletrodos.

No caso deste estudo, os sinais utilizados seguem o segundo tipo de

montagem: cada canal é derivado a partir da diferenciação entre dois eletrodos vizinhos, sem a utilização de referência fixa.

Figura 1.2: Posicionamento de eletrodos baseado no sistema internacional 10-20. As letras representam as regiões do cérebro (F - frontal, T - temporal, P - parietal, O - occipital, C - central). Z representa o eixo sagital e os números são associados às distâncias dos eletrodos ao eixo Z, com números ímpares representando eletrodos à esquerda e números pares representando eletrodos à direita.

1.3.3 EEG na epilepsia

A epilepsia é a condição que mais faz uso da análise de EEG e muito da análise clínica já é bem estabelecido [22]. Normalmente é utilizada a análise de diversos ritmos que são divididos pela frequência das oscilações, apesar de as definições entre estas oscilações não ser muito exata, podendo variar entre fontes. Usualmente a divisão dos ritmos pode ser aceita como

- Delta: $< 4Hz$.
- Teta: 4 a $7Hz$.
- Alfa: 8 a $12Hz$.
- Beta: 13 a $30Hz$.
- Gama: $> 30Hz$.

A consideração de um sinal EEG como normal, ou seja, sem a presença de indicadores de patologia, também não é exatamente definida. O que existem são relatos de características de ritmos e padrões que ocorrem em determinadas situações normais de funcionamento do cérebro (sono, vigília, etc). A distribuição topográfica e simetria destes ritmos e padrões é considerada importante, e pequenas variações do que é esperado nestas condições normais pode indicar alguma anormalidade.

Para considerar um EEG como epiléptico, é necessária a identificação de uma série de variações nos ritmos cerebrais como ocorrência de *spikes* (alterações bruscas em amplitude com duração entre 20 a 70 ms), por exemplo [23]. Estas alterações caracterizam o que pode ser chamado de marcador biológico (ou biomarcador) para as condições de epilepsia, ou seja, através da identificação destes padrões específicos da epilepsia, é possível não só a indicar a presença da condição patológica como também estimar a localização da zona epileptogênica e classificar a gravidade da condição. Outras conclusões sobre marcadores em epilepsia podem ser observadas em [24] e [23].

Figura 1.3: Exemplo de sinal de EEG com início de crise epiléptica. A linha vertical vermelha indica o início da crise, marcado por um especialista. Repare como o sinal elétrico capturado pelos eletrodos indica maior atividade durante uma crise. O aumento na amplitude das oscilações é facilmente distinguido do chamado sinal de base, que é o sinal do paciente em condições normais (no período chamado inter-ictal, ou entre crises). Muito do que é capturado pelos eletrodos durante uma crise epiléptica é artefato proveniente de contrações musculares. Sinal adquirido do banco de dados do Physionet (detalhado em 2.1)

1.4 Oscilações de alta frequência - HFOs

Dentro da rotina clínica, usualmente eram analisados os traços de EEG correspondentes à atividades de 0.5 à 40 Hz. Com o avanço tecnológico nas composições instrumentais e o potencial computacional dos softwares desenvolvidos, surgiu a possibilidade de aquisição e processamento de sinais digitalizados de EEG com frequências de aquisição cada vez mais altas. Assim, há cerca de 10 anos, vem se criando um grande interesse aos estudos de sinais de EEG com foco nas atividades de alta frequência.

1.4.1 Os padrões de atividade de alta frequência

Desde então, as observações e estudos vêm destacando um tipo de padrão específico que potencialmente poderia ser utilizado como biomarcador de epilepsia. Esse padrão recebe o nome de Oscilação de Alta Frequência (OAF, do inglês *High Frequency Oscillation* - HFO).

O primeiro registro de padrões de HFO ocorrendo de forma espontânea em pacientes de epilepsia foi obtido em 1999 [25]. A partir de então, diversos pesquisadores vêm observando grande potencial na análise da relação entre as HFOs e as crises de epilepsia.

A detecção das HFOs foi capaz de diferenciar a localidade da região causadora das crises epilépticas de pacientes com imagens de MRI normal [26] e mesmo registros feitos no escalpo foram capazes de fornecer dados promissores para uso de HFOs em comparação ao uso de *spikes* como biomarcadores para localização da zona epileptogênica [27], desde que se tome cuidado especial para diferenciar atividade elétrica cerebral de artefatos musculares.

Porém, existem certas situações normais em que podem ser detectadas HFOs em indivíduos saudáveis, mostrando que ainda existe espaço para propostas e desenvolvimentos de métodos que possam classificar os padrões de alta frequência de forma mais eficiente [28, 29].

Jacobs e colaboradores fizeram uma revisão de vários estudos que sugerem o potencial da utilização de HFOs como biomarcadores concluindo que, atualmente,

somente os pacientes sem resposta a medicamentos que são candidatos à cirurgia se beneficiam dos estudos de HFOs, devido a maior facilidade de se observar estes padrões em registros intra-cranianos [30]. Porém, sugere que a identificação dos padrões de HFO em registros de escalpo poderiam auxiliar o acompanhamento de um maior número de pacientes por ser não invasiva, necessitando de melhorias na parte médica (melhor compreensão dos padrões para facilitar a identificação e qualificação de HFOs como patológicas ou fisiológicas) e na parte da engenharia (melhorias nos métodos de identificação dos padrões).

Figura 1.4: Exemplo de aumento de atividade de alta frequência em EEG epiléptico. No topo, um sinal de um canal de EEG sem nenhum filtro digital. Em baixo, o sinal é passado por um filtro passa banda, com frequências de corte de $F_{c1} = 80$ e $F_{c2} = 118\text{Hz}$. Repare como há um aumento visível na atividade em alta frequência nos trechos delimitados pelas linhas verticais vermelhas. A escala de amplitude (canto superior esquerdo) no sinal sem filtro corresponde à uma amplitude de $75\mu\text{V}$ enquanto a escala de amplitude no sinal filtrado corresponde à uma amplitude de $5\mu\text{V}$.

1.5 Objetivos

O objetivo do estudo foi extrair características de sinais de EEG de pacientes diagnosticados com epilepsia através da aplicação um método não linear baseado na *Detrended Fluctuation Analysis* (DFA, ou Análise de Flutuações Destendenciada) e da detecção do aumento local de atividade de alta frequência através da aplicação de um método de detecção de HFOs baseado em um método proposto na literatura, e investigar as alterações destas características durante o sinal de EEG e quais as relações entre a dinâmica de tais com a atividade ictal dos pacientes.

Em especial, a observação de alterações nas características medidas antes que a crise possa ocorrer pode indicar que seria possível expandir a utilização dos métodos aplicados para análise em tempo real de sinais de pacientes com epilepsia, possibilitando a detecção prévia das crises durante a aquisição do sinal.

A detecção de crises pode causar grande impacto nas investigações relacionadas ao diagnóstico e tratamento em epilepsia. Possibilidades de exames SPECT pré-ictal, intervenção médica para evitar a ocorrência dos disparos anormais de neurônios e até possibilitar os pacientes a se prepararem para a ocorrência da crise, facilitando a execução de tarefas do cotidiano com menor riscos para a saúde, resultando no aumento da qualidade de vida dos mesmos.

1.6 Detrended Fluctuation Analysis

A Análise de Flutuações Retificada (DFA - *Detrended Fluctuation Analysis*) é uma poderosa ferramenta capaz de fornecer um parâmetro simples para representar as propriedades de correlação de leis de potência à longos prazos em um sinal. As vantagens deste método sobre vários outros é a capacidade de se detectar correlações mesmo quando a série temporal analisada não é estacionária [31]. Descrito primeiramente para análise de séries de batimentos cardíacos não estacionárias [32], o método já foi utilizado com sucesso nos mais variados estudos, como identificação nas chances de sucesso de desfibrilações [33], identificação de relações entre incêndios em florestas e parâmetros climáticos [34], monitoramento da atividade sísmica de vulcões [35], análise de descargas neuronais [36], estudos sobre apneia do sono [37], entre outros.

Para a ilustração do método, considera-se a série temporal $x[i]$ ($i = 1, \dots, N$). Calcula-se a série temporal integrada de $x[i]$,

$$x_{int}[i] = \sum_{n=0}^i (x[n] - \langle x \rangle), \quad (1.2)$$

onde $\langle x \rangle$ é a média dos valores de $x[i]$,

$$\langle x \rangle = \frac{1}{N} \sum_{j=1}^N x[j]. \quad (1.3)$$

A série integrada é dividida então em K segmentos de tamanho L e, para cada segmento, ajusta-se a série utilizando uma função polinomial $y[i]$.

No caso deste trabalho, foi utilizada a DFA de 1ª ordem. Neste caso, a função $y[i]$ é de 1ª ordem, sendo da forma

$$y[i] = a_j i + b_j. \quad (1.4)$$

onde $j = (1 \dots K)$.

Calcula-se então o desvio quadrático total entre o sinal integrado e as aproximações lineares,

$$F[L] = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_{int}[i] - y[i])^2}. \quad (1.5)$$

O procedimento de 1.5 é repetido para diferentes valores de tamanho de segmento (L), obtendo-se então uma relação entre os valores de $F[L]$ e L .

Figura 1.5: Exemplo de segmentações em DFA. Um sinal composto por ruído branco de tamanho $N = 5000$ foi utilizado para a demonstração. No topo, o sinal é dividido em 50 segmentos. No meio, o sinal é dividido em 10 segmentos e no fundo o sinal é dividido em 5 segmentos. Para cada segmento, é feita uma regressão linear. Para o cálculo do expoente de correlação da DFA, o valor do erro entre a aproximação linear e o sinal é calculado para cada valor de tamanho de segmento L .

Usualmente pode-se descrever a relação obtida como uma relação de lei de potência

$$F[L] \propto L^\alpha, \quad (1.6)$$

O coeficiente α é o fator de escala, capaz de indicar o nível da correlação a longo prazo na série $x[i]$. Fazendo a construção de um gráfico log-log dos dados de $F[L]$ e L , pode-se extrair α como o coeficiente angular da reta obtida por regressão linear.

Figura 1.6: Exemplo do cálculo do expoente através do método da DFA. Os valores do tamanho do intervalo (L) e do RMS para cada tamanho ($F(L)$) são plotados em escala logarítmica. Os resultados podem ser aproximados para uma reta, e o valor do coeficiente angular desta reta é o coeficiente resultante do método. Na figura, são mostradas as retas resultantes da aplicação do método em um ruído branco e em um ruído browniano, com seus respectivos coeficientes sendo indicados.

Porém, nem sempre os dados dos valores de $\log(F[L])$ vs $\log(L)$ podem ser ajustados à apenas uma reta. Desde o artigo original do método, existe uma atenção ao fenômeno chamado de *Crossover Phenomena* (ou “fenômeno de cruzamento”). Este fenômeno tem o nome derivado do fato de que em certos casos, pode ocorrer a alteração do coeficiente angular da reta obtida a partir de um valor de escala L . O ponto em que ocorre a variação de coeficiente angular foi pensado como ponto de cruzamento entre as duas retas, logo, fenômeno de cruzamento.

A ocorrência deste fenômeno foi originalmente associada com a utilização de séries temporais de tamanho pequeno (menor que 10 mil pontos), que poderiam apresentar vantagens na análise de dados em diversas aplicações, como nas áreas de ciências médicas.

Figura 1.7: Ilustração do efeito de crossover no método da DFA. É possível perceber que há um ponto (com $L \simeq 1.76$) em que o coeficiente angular da reta que melhor ajusta os pontos se altera. Este é o chamado fenômeno de crossover. Por isso, muitas vezes a análise pelo método da DFA retorna dois valores de coeficiente de correlação em lei de potência: α_1 e α_2 .

1.7 Filtros Digitais

O uso de filtros de frequência abre muitas possibilidades para o estudo de sinais, desempenhando papel importante tanto no processamento dos sinais para a extração de informações sob aspecto de diferentes bandas de frequência, ou até mesmo na própria aquisição do sinal, na implementação de um filtro eletrônico para eliminar componentes de frequência que não satisfazem o Teorema de Nyquist, ou seja, impedir que o sinal adquirido contenha componentes de frequência que sejam maiores que 2 vezes a frequência de aquisição.

Os filtros eletrônicos são amplamente aplicados na área de processamento de sinais, mas a relativa dificuldade de sua implementação e a exigência de repetidas análises mesmo após a aquisição inicial do sinal tornam o uso de filtros digitais essenciais para várias pesquisas.

1.7.1 Representação em tempo e frequência

Os sinais registrados por EEG são o que é chamado de série no tempo. Cada amostra (ou cada ponto) do sinal corresponde à atividade elétrica em um determinado instante de tempo. Embora seja possível o desenvolvimento de métodos de análise em cima das propriedades temporais de um sinal, praticamente todos os métodos desenvolvidos para manipulação e extração de informações de sinais de EEG incluem alguma etapa envolvendo análise em frequência.

Joseph Fourier foi capaz de mostrar como representar as séries temporais através de funções de senos e cossenos, cujo as frequências seriam as frequências elementares desta série.

Seja $x[t]$ uma série temporal periódica com período $T = 1/f$, sua representação em termos de seno e cosseno é

$$x[t] = \frac{a_0}{2} + \sum_{n=1}^{\infty} [a_n \cos 2\pi n f t + b_n \sin 2\pi n f t]. \quad (1.7)$$

Esta representação é chamada de série de Fourier. Os coeficientes a_n e b_n são utilizados para representar a energia e a potência de um sinal, sendo a energia

a soma dos quadrados dos coeficientes e a potência seria a energia dividida pelo tamanho do intervalo de tempo em que a série foi calculada.

Representando a potência em função da frequência, obtêm-se o que é definido como o espectro de potências da série $x[t]$.

A representação por series de Fourier se limita apenas a sinais periódicos, e o espectro de potências possui componentes discretas, múltiplas das componentes de frequência fundamentais do sinal. Porém, sinais não periódicos podem possuir um espectro de potências contínuo na frequência.

Para a análise destes sinais, utiliza-se a generalização das séries de Fourier. Esta generalização é chamada de Transformada de Fourier.

Para uma série de tempo discreto $x[t]$, a representação em frequência seria calculada por

$$X(e^{i\omega}) = \sum_{t=-\infty}^{\infty} x[t]e^{-i\omega t}. \quad (1.8)$$

1.7.2 Definição de Filtros Digitais

O objetivo final de um filtro é que as componentes de frequência indesejadas para a análise (as chamadas “bandas de rejeição”) sejam atenuadas, enquanto as componentes de frequência desejadas sofram o mínimo de alteração possível (o nível de alteração é comumente chamado de *ripple* da banda de passagem). Os pontos de separação entre uma banda de passagem e uma banda de rejeição é chamada “frequência de corte” F_c , e o intervalo em que ocorre a transição entre a banda de frequência e a banda de rejeição é chamado de “banda de transição”.

Um filtro consiste então de uma função de resposta em frequência ($H(e^{i\omega})$) que ao ser multiplicada pela representação em frequência de uma série temporal $x[t]$ retorna a representação da série filtrada de x ($X(e^{i\omega})$). Em outras palavras, H deve ser composta por coeficientes de ponderação H_ω , que serão aplicados às componentes de frequência descritas em $X(e^{i\omega})$.

Um filtro ideal possuiria atenuação máxima na banda de rejeição, nenhum *ripple* na banda de passagem e a transição entre a banda de rejeição e a banda de

passagem, na frequência de corte, seria abrupta.

Figura 1.8: Exemplificação de filtros de frequência. (A) Exemplo de resposta em frequência ($H(e^{i\omega})$) de um filtro passa alta com $F_c = 200\text{ Hz}$. (B) Filtro FIR passa alta obtido através do método de janelamento. Em (B1) os traços azuis delimitam a banda de transição e, em (B2), os traços azuis delimitam o ripple na banda de passagem.

1.7.3 Filtros Ideais e Fenômeno de Gibbs

A implementação de um filtro ideal é, entretanto, impossível. A tentativa de se representar transições abruptas (ou descontinuidades) resulta no surgimento

de uma distorção próxima à frequência de corte. Esta distorção é conhecida como fenômeno de Gibbs.

A ocorrência do fenômeno está relacionada ao fato de que uma série de funções contínuas uniformemente convergente necessariamente representa uma função contínua [38], não havendo, portanto, como representar a transição abrupta.

Este fenômeno deve ser evitado, pois o surgimento de oscilações na banda de passagem faz com que o sinal resultante da aplicação do filtro de frequências sofra grandes distorções em seu espectro de frequências.

Figura 1.9: O Fenômeno de Gibbs. Este filtro possui os mesmos parâmetros do filtro em 1.8(B2), porém, não passa pelo efeito de janelamento (ver seção “Método de Janelamento”). Repare que o ripple na banda de passagem é muito maior ($\sim 2\text{db}$), enquanto a largura da banda de transição é menor. Usualmente existe um comprometimento entre estas duas características no filtro. Mesmo com técnicas para diminuir o fenômeno de Gibbs, sempre que tenta-se reduzir o ripple na banda de passagem, há um aumento na largura da banda de transição, e vice e versa.

Outra propriedade inconveniente do fenômeno de Gibbs é que a amplitude do ripple próximo à F_c' não diminui com o aumento do tamanho do sinal. Considera-se então que a única forma de se livrar da ocorrência do fenômeno é, realmente, evitar a sua ocorrência.

Já que o fenômeno é decorrente da incapacidade de representar uma função com transição abrupta a partir de uma função contínua, faz-se com que a transição da banda de passagem para a banda de corte ocorra de uma forma mais suave (resultando na banda de passagem).

A aceitação de que tanto a largura na banda de transição quanto a pequenas distorções na banda de passagem são inevitáveis nos limita a apenas obter a melhor forma de atenuar estes efeitos indesejáveis.

Um estudo mais aprofundado sobre o desenho de filtros mostra que existe um comprometimento entre a largura da banda de transição e a distorção na banda de passagem. Portanto, o ajuste de parâmetros geralmente ocorre de forma que os filtros obtidos se adaptem às necessidades de cada estudo.

1.7.4 Convolução

A convolução é uma operação matemática formal que parte de dois sinais e gera um terceiro sinal. É comumente utilizada em várias áreas envolvendo matemática, como probabilidade, estatística, processamento de imagens e sinais [39].

A definição utilizada para descrever uma convolução entre dois sinais ($x[t]$ e $h[t]$) é, usualmente,

$$y[t] = \sum_{n=0}^{\infty} h[n]x[t-n]. \quad (1.9)$$

Uma propriedade interessante da convolução, que torna esta operação extremamente útil para a aplicação de filtros digitais, tem relação com a transformada de Fourier de um sinal obtido através da convolução.

Seja o sinal $y[t]$ obtido como a convolução de um sinal $x[t]$ com uma função $h[t]$ (como na equação 1.9). Substituindo o valor da transformada inversa para $x[t-n] = \frac{1}{2\pi} \sum_{\omega=-\infty}^{\infty} X[i\omega]e^{i\omega(t-n)}$ na equação 1.9, chega-se a

$$y[t] = \sum_{\omega=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} h[n] \frac{1}{2\pi} X[i\omega] e^{i\omega t} e^{i\omega n} = \frac{1}{2\pi} \sum_{\omega=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} h[n] e^{i\omega n} X[i\omega] e^{i\omega t}$$

$$= \frac{1}{2\pi} \sum_{\omega=-\infty}^{\infty} H[i\omega]X[i\omega]e^{i\omega t} \quad (1.10)$$

Ou seja, a convolução de duas séries temporais corresponde à transformada inversa de Fourier da multiplicação de seus espectros de potência.

Assim como uma série temporal $x[t]$ pode ser representada por um espectro de frequências $X[e^{i\omega}]$, um espectro de frequência de um filtro $H[e^{i\omega}]$ pode ser representado por uma série temporal $h[t]$. Esta representação é chamada de “resposta ao impulso” do filtro digital.

Como a aplicação do filtro consiste na multiplicação da resposta em frequência deste filtro ao espectro de frequência de uma série temporal, de acordo com a equação 1.10, a convolução da resposta ao impulso do filtro com a série temporal também resultará em um sinal filtrado.

Com exceção dos casos em que se pode aplicar a chamada Transformada Rápida de Fourier, a filtragem por convolução é geralmente mais rápida do que a filtragem pela aplicação da resposta em frequência do filtro no espectro de frequências da série, além de possuir uma implementação muito mais simples.

1.7.5 Filtros Digitais FIR e Desenho por Janelamento

É possível classificar os filtros digitais, em relação à resposta ao impulso, como FIR (resposta ao impulso finita) ou IIR (resposta ao impulso infinita).

Em comparação com os filtros FIR, os filtros IIR conseguem satisfazer as mesmas especificações de resposta em frequência com uma ordem menor. Isso resulta em computações mais rápidas do sinal filtrado, potencialmente aumentando o desempenho do algoritmo de filtragem. Porém, os filtros IIR são potencialmente indesejáveis por não possuir o aspecto de fase linear, o que pode ocasionar distorções de fase no sinal filtrado.

Em compensação, os filtros FIR são mais fáceis de se implementar, e possuem respostas em frequência com fase linear, não causando distorções de fase no sinal filtrado [40].

O processo de desenho de um filtro envolve definir a resposta em frequência desejada $H_d(e^{i\omega})$. Basta agora obter os coeficientes da resposta ao impulso $h[t]$ de

forma que a resposta em frequência obtida $H(e^{i\omega})$ possua uma boa aproximação com $H_d(e^{i\omega})$.

Considerando que $h_d[t]$ é a resposta ao impulso da resposta em frequência desejada pelo filtro, pode-se considerar que o erro entre o filtro obtido e o filtro desejado pode ser expresso por:

$$E = \sum_{t=-\infty}^{\infty} |h_d[t] - h[t]|^2. \quad (1.11)$$

É possível notar que a melhor forma de se minimizar o erro, utilizando um número finito de coeficientes N , é fazendo

$$h[t] = \begin{cases} h_d[t], & 0 \leq t \leq N \\ 0, & \text{caso contrário} \end{cases} \quad (1.12)$$

A equação 1.12 é resultado da aplicação de uma janela retangular à resposta ao impulso do filtro desejado. Uma janela é uma função de tamanho N em que os pontos contidos no intervalo $[0 \ N]$ possuem valor diferente de zero, e os valores do que está fora do intervalo é zero. A multiplicação de uma função janela com uma função h_d limita os valores de h_d no intervalo $[0 \ N]$.

A janela retangular é a janela mais simples que existe, possuindo valor 1 dentro do intervalo e 0 fora do intervalo. A aplicação da janela retangular em h_d pode ser escrita como

$$h[t] = w[t]h_d[t], \quad (1.13)$$

onde

$$w[t] = \begin{cases} 1, & 0 \leq t \leq N \\ 0, & \text{caso contrário} \end{cases} \quad (1.14)$$

Porém, como explicado na seção 1.7.3, a utilização de transições abruptas (neste caso, a janela retangular) para obter representações entre os domínios do tempo e da frequência não é uma boa ideia.

Por isso, normalmente é utilizada uma janela com transições suaves para a obtenção dos coeficientes do filtro. Existem diferentes janelas utilizando de diferentes funções para obter uma transição suave e melhorar a representação entre os domínios tempo-frequência. As mais comuns são exemplificadas na figura 1.10.

Neste trabalho, os filtros foram obtidos pelo método de janelamento utilizando a janela de Hamming,

$$w[t] = \begin{cases} \alpha - (1 - \alpha) \cos \frac{2\pi t}{N}, & 0 \leq t \leq N \\ 0, & \text{caso contrário} \end{cases} \quad (1.15)$$

com $\alpha = 0.54$.

Figura 1.10: Diferentes janelas e suas respostas em frequência. No topo, uma comparação entre as funções de janelamento no tempo. No meio, a comparação entre as respostas em frequência de cada janela. No fundo, uma aproximação para melhor visualização das janelas na banda de transição. A janela de Hamming foi escolhida pois, apesar de a resposta em frequência em frequências distantes da F_c não ser tão atenuada, próximo à F_c pode-se observar uma melhor atenuação, possuindo também a transição no menor intervalo de frequência, ao lado da janela de Hann.

1.8 Detecção de HFOs

A observação dos padrões de alta frequência certamente abriu um novo caminho para a análise dos registros de EEG em pacientes de epilepsia. Porém

podemos citar a marcação das oscilações como um aspecto de dificuldade para os estudos. A marcação de forma manual destes padrões, feita por um especialista em neurologia, demandaria as seguintes etapas principais:

- Visualização dos sinais, filtrados para mostrarem apenas as frequências baixas (até 40Hz), em uma certa escala de tempo e amplitude, para eliminação dos trechos contendo ruídos.
- Visualização dos sinais filtrados para mostrar apenas as frequências altas (acima de 40Hz até a frequência desejada/possível), necessitando nova escala de tempo e amplitude, para visualização de possíveis oscilações que se destaquem do *background*.
- Comparação entre as duas visualizações no trecho onde há a observação do padrão em alta frequência, para a eliminação de oscilações que possam ocorrer por qualquer outra atividade elétrica que não fisiológica do cérebro.

Sabendo que estas etapas devem ser feitas em registros multicanal (até 21 canais para registros em escalpo e mais de 40 canais para registros intracorticais) de mais de 5 horas de duração, com frequência de aquisição relativamente alta, é fácil perceber que o processo seria extremamente demorado e delicado, demandando muito tempo e esforço do profissional. Além disso, marcações manuais sempre são subjetivas, e o resultado final pode e será diferente entre diferentes especialistas.

Assim, há a necessidade da criação de métodos automáticos ou semi-automáticos para auxiliar os especialistas a obter as localizações das HFOs nos registros de EEG.

Alguns métodos surgiram visando a detecção automática ou semi-automática de HFOs para diferentes bandas de frequência ou ate diferentes locais.

Grande parte dos métodos funciona através da comparação de energia e/ou potência com um valor limite, calculado para cada trecho do sinal. Zelman e colaboradores fizeram uma comparação entre o desempenho de diferentes detectores, mostrando que é possível a ocorrência de casos em que detectores mais complexos possuem melhor desempenho do que os outros, com o exemplo de canais onde há

muitas ocorrências de HFOs, mas o desempenho é parecido em canais com algumas ocorrências (condição que geralmente é satisfeita) [41].

O método desenvolvido para as detecções descritas neste trabalho é baseado no método proposto por von Ellenrieder e colaboradores [42].

Foi escolhido este método pois a banda de frequência para que este foi desenvolvido - 40 a 200 Hz - inclui da banda de frequência de interesse para as análises dos dados do Physionet (seção 2.1) - 80 a 128 Hz. Além disso, o método foi considerado relativamente fácil de se implementar, aparentando ser robusto o suficiente para satisfazer as necessidades deste estudo.

Para a implementação deste método considera-se que, como definido na proposta inicial do método, os padrões de HFOs possuem quatro ciclos em sua frequência. Busca-se, então, detectar as HFOs através de aumentos na potência em trechos de duração de $4N_k$, onde N_k é o número de amostras em um ciclo da frequência central da banda de interesse.

No caso dos sinais do banco de dados CHB-MIT (Physionet), como a frequência de aquisição dos dados é de 256Hz, pode-se trabalhar na faixa de 80 a 128 Hz. Esta é, portanto, nossa banda de interesse.

Os sinais são então filtrados para frequências altas, com frequência de corte de 80 Hz, e logo após são filtrados novamente com um filtro passa-banda de 80 a 128 Hz.

Calcula-se então, para todo o sinal, os valores de RMS para janelas móveis de tamanho $4N_k$.

$$RMS(x) = \sqrt{\frac{1}{4N_k + 1} \sum_{m=-2N_k}^{2N_k} y_k^2[x - m]}, \quad (1.16)$$

onde N_k é o número de amostras em um ciclo da frequência central da banda de interesse.

Além do valor do RMS, é calculado também o valor do *background* móvel

$$B_k(x) = c \frac{1}{N_b} \sum_{m=0}^{N_b-1} z_k[x - 2N_k - m], \quad (1.17)$$

onde N_b é o número de amostras na janela móvel do *background*, c é a constante de proporcionalidade do limiar e $z_k[x]$ é o valor do RMS na posição x , limitado pelo valor do limiar móvel

$$z_k[x] = \min(B_k(x), RMS[x]), \quad (1.18)$$

Visando facilidade de implementação computacional, o valor de B_k é calculado da seguinte forma

$$B_k(x) = B_k(x-1) + c\left(\frac{1}{N_b}(z_k[x-2N_k] - z_k[x-N_b-2N_k])\right), \quad (1.19)$$

É feita então a comparação, ponto a ponto, entre os valores calculados em 1.16 e 1.19, obtendo-se um vetor de resultado da forma

$$R(x) = \begin{cases} 1, & RMS[x] > B_k[x] \\ 0, & \text{caso contrário} \end{cases} \quad (1.20)$$

Nota-se, então, que o resultado é dado em forma de um vetor de dados com valores 0 e 1.

METODOLOGIA

2.1 Sinais utilizados

Foram utilizados dados de EEG coletados de pacientes pediátricos no *Children's Hospital Boston* (Hospital de Crianças de Boston).

Os dados são de 5 pacientes de sexo masculino com idades de 3 a 22 anos e 17 pacientes do sexo feminino com idade de 1.5 a 19 anos. Cada paciente possui de 9 a 42 arquivos, contendo cerca de 1 hora contínua de registro de EEG. Há intervalos entre os registros de cada arquivo, devido a limitações de equipamentos (os intervalos são de 10 segundos ou menos para a maioria dos casos, mas existem alguns intervalos maiores ocasionalmente). Os dados pessoais e datas das aquisições foram substituídos por informação *surrogate* para preservar a privacidade dos paciente, mas as relações de tempo entre os arquivos individuais de cada paciente foram mantidas.

Todos os sinais foram adquiridos à uma frequência de aquisição de 256 Hz com resolução de 16 bits, utilizando o sistema internacional 10-20 para posicionamento e nomenclatura dos eletrodos (figura 1.3.2). Foi retirada qualquer medicação anti-convulsão dias antes da coleta dos dados para melhor caracterizar as crises de forma natural.

A maioria dos arquivos possui cerca de 1 hora de duração. Porém, alguns arquivos podem possuir 2 ou 4 horas de duração. A frequência de aquisição e resolução, porém, não são alteradas.

Todas as crises observadas durante as aquisições foram registradas, existindo marcações sobre o início e o término de cada crise. Estas marcações possibilitaram as

análises dos parâmetros extraídos do sinal e comparações em relação à proximidade da crise.

Todos os dados são descritos com mais detalhes por Shoeb em sua tese [43], e estão disponíveis no banco de dados do PhysioNet [44].

2.1.1 Sobre a frequência de aquisição dos sinais

Vale ressaltar que, como a frequência de aquisição dos sinais utilizados é $F_a = 256Hz$, de acordo com o Teorema de Nyquist, podem ser analisadas apenas componentes de frequência até $F_{max} = 128Hz$. Esta frequência máxima está dentro do que é considerado na literatura de HFOs em epilepsia como faixa de *Ripple*. Para a análise completa da atividade de alta frequência, devem ser utilizados dados com frequências de aquisição maiores, permitindo incluir análises dentro da faixa de *Fast-Ripple*. Alguns estudos indicam que a atividade na faixa de *Fast-ripple* teria maiores relações com a ocorrência de patologias, como a presença de foco epilético. Porém, atualmente esta determinação já não é tão bem aceita, podendo-se encontrar indícios de HFOs relacionadas à presença do foco epilético mesmo dentro da faixa de *Ripple*.

2.2 Planejamento dos Filtros Digitais

Como explicado na seção 1.7, todos os filtros utilizados neste estudo foram obtidos pelo método de janelamento utilizando uma janela de Hamming. As equações para obtenção da resposta ao impulso dos filtros e indicada abaixo.

$$h_{lp}[n] = \begin{cases} \frac{w_1}{\pi}(\alpha - (1 - \alpha) \cos(\frac{2\pi n}{N-1})), & \tau = n \\ \frac{\sin((n - \tau)w_1)}{\pi * (n - \tau)}(\alpha - (1 - \alpha) \cos(2\pi n/(N-1))), & \tau \neq n \end{cases} \quad (2.1)$$

$$h_{hp}[n] = \begin{cases} (1 - \frac{w_1}{\pi})(\alpha - (1 - \alpha) \cos(\frac{2\pi n}{N-1})), & \tau = n \\ \frac{\sin((n - \tau)\pi) - \sin((n - \tau)w_1)}{\pi(n - \tau)}(\alpha - (1 - \alpha) \cos(2\pi n/(N-1))), & \tau \neq n \end{cases} \quad (2.2)$$

$$h_{bp}[n] = \begin{cases} \frac{w_2 - w_1}{\pi}(\alpha - (1 - \alpha) \cos(\frac{2\pi n}{N - 1})), & \tau = n \\ \frac{\sin((n - \tau)w_2) - \sin((n - \tau)w_1)}{\pi(n - \tau)}(\alpha - (1 - \alpha) \cos(2\pi n/(N - 1))), & \tau \neq n \end{cases} \quad (2.3)$$

$$h_{bs}[n] = \begin{cases} (1 - \frac{w_2 - w_1}{\pi})(\alpha - (1 - \alpha) \cos(\frac{2\pi n}{N - 1})), & \tau = n \\ \frac{\sin((n - \tau)w_1) - \sin((n - \tau)w_2)}{\pi(n - \tau)}(\alpha - (1 - \alpha) \cos(2\pi n/(N - 1))), & \tau \neq n \end{cases} \quad (2.4)$$

Na figura 2.1, é possível observar a resposta ao impulso e resposta em frequência dos filtros obtidos por este método. Os parâmetros importantes que descrevem a resposta em frequência podem ser observados na tabela 2.1.

	Ripple (db)	Atenuação (db)	largura transição (Hz)
Passa Baixa	0.02	-58	9.5
Passa Alta	0.02	-57	9.5
Passa Banda	0.02	-56	9.5
Rejeita Banda	0.02	-20	6.5

Tabela 2.1: Parâmetros da resposta em frequência dos filtros implementados. Os parâmetros apresentados mostram aproximadamente o resultado menos satisfatório obtido na configuração de frequências de corte em que os filtros seriam utilizados, ou seja, garante-se que todos os filtros utilizados possuem parâmetros de melhor ou igual qualidade do que os dados apresentados. As frequências de corte são: Passa baixa - $F_c = 128Hz$, Passa alta - $F_c = 80Hz$, Passa banda - $F_{c1} = 80Hz$, $F_{c2} = 118Hz$, Rejeita banda - $F_{c1} = 57Hz$, $F_{c2} = 63Hz$.

Os resultados apresentados na tabela mostram que a resposta em frequência dos filtros implementados é boa, sendo a única falha dos filtros passa baixa, passa alta e passa banda a largura da banda de transição. Mesmo assim, a largura ainda está dentro do que espera-se de um filtro de alta qualidade.

A exceção é o desempenho do filtro rejeita banda de atuar como um filtro *notch* (filtro que corta uma frequência ou faixa de frequência estreita específica) em

60 Hz. Este filtro possui atenuação relativamente baixa, e mesmo com a definição das frequências de corte para rejeitar uma banda de 6Hz , a largura da banda rejeitada pelo filtro resultante é de $\sim 13\text{Hz}$. Este filtro não foi utilizado neste estudo. Os dados provenientes do banco de dados do Physionet (ler seção 2.1) já possuem filtragem para eliminar o ruído proveniente da rede elétrica.

Figura 2.1: Resposta ao impulso e a frequência dos filtros. Em ordem, superior-esquerdo: Filtro passa baixa, $F_c = 250\text{Hz}$; superior-direito: Filtro passa alta, $F_c = 250\text{Hz}$; inferior-esquerdo: Filtro passa banda, $F_{c1} = 200\text{Hz}$, $F_{c2} = 250\text{Hz}$; inferior-direito: Filtro rejeita banda, $F_{c1} = 200\text{Hz}$, $F_{c2} = 250\text{Hz}$. Todos os filtros são obtidos com $N = 401$

2.3 Detrended Fluctuation Analysis

A aplicação do método de DFA neste trabalho envolve uma iniciativa não tão comum. Trata-se da tentativa de se estimar o coeficiente α em uma janela móvel no tempo, obtendo-se então uma função de α pelo tempo.

Como relatado na literatura, nos períodos próximos ao ictal, podem ser observadas alterações no sinal de EEG. Neste aspecto, o objetivo dessa estimativa é poder analisar a variação da correlação em lei de potência em diferentes situações de um paciente de epilepsia, focando principalmente na análise em segmentos próximos à uma crise epiléptica.

Uma vez calculados os dados da variação do coeficiente α em função do tempo, é possível também a comparação desta variação com outras análises temporais, como a própria detecção de HFOs.

Destaca-se também que, devido à ocorrência do fenômeno de *crossover* (seção 1.6) com o tamanho N utilizado, foi necessária a adaptação dos algoritmos para trabalhar não com um coeficiente α , mas com dois - um para cada ajuste obtido. Os dois coeficientes são chamados, neste texto, de α_1 e α_2 .

Para a criação da série de variação temporal dos coeficientes, os sinais foram primeiramente divididos em segmentos de tamanho N . Para cada segmento do sinal, foram obtidos os valores de α_1 e α_2 aplicando-se o método da DFA. Cada ponto destas séries corresponde, então, aos valores de α_1 e α_2 para o segmento de tamanho N correspondente àquela janela de tempo.

Uma vez com a série temporal obtida, os segmentos foram divididos em 4 categorias: segmentos temporalmente distantes de uma crise, segmentos próximos antes de uma crise, segmentos próximos após uma crise e segmentos durante uma crise 2.2.

Com esta divisão, podem ser feitas comparações das distribuições de todos os segmentos de cada um dos conjuntos, para assim analisar o quão possível é diferenciar cada uma das distribuições através de um teste de significância.

2.4 Detecção de HFOs

Como explicado em 1.8, o método de detecção de HFO implementado visa a detecção de aumentos locais de potência para o sinal filtrado na banda de interesse e retorna como resultado um vetor de dados contendo todas as ocorrências de HFOs.

Este método tem ótima resolução temporal, e cada ponto do vetor de dados resultante corresponde à uma aquisição no sinal de entrada. Com isso, é possível analisar o resultado em função do tempo com resolução igual a $1/f_a$, onde f_a é a frequência de aquisição do sinal.

De acordo com a literatura, a ocorrência de HFOs pode ser mais intensa nas proximidades de uma crise, apesar de também ocorrer no período inter ictal. Neste estudo, a análise dos resultados deste método é feita a partir de dados de taxas de ocorrência de detecções de HFOs em uma janela móvel no tempo de tamanho N .

A segmentação dos resultados, apesar de causar perda na resolução temporal, facilita a análise por reduzir a quantidade de dados a serem processados além de facilitar a comparação com outros métodos, como a análise proposta pela DFA. Além disso, a janela de tempo utilizada possui tamanho relativamente curto. É esperado que, mesmo com a perda da resolução temporal, o aumento da taxa de ocorrência de HFOs seja suficiente para identificar qualquer alteração na atividade em alta frequência do sinal EEG.

Uma vez feita a conversão de dados, cada ponto na série resultante mostra a taxa de ocorrência de detecções de HFOs por número de pontos. Resta então classificar os segmentos em 4 categorias: segmentos temporalmente distantes de uma crise, segmentos próximos antes de uma crise, segmentos próximos após uma crise e segmentos durante uma crise (tabela 2.2).

Com esta divisão, podem ser feitas comparações das distribuições de todos os segmentos de cada um dos conjuntos, para assim analisar o quão possível é diferenciar cada uma das distribuições através de um teste de significância.

Como os dois métodos testados possuem mesmo tamanho de segmentos e utilizam as mesmas classificações, uma comparação entre os resultados dos dois métodos pode ser feita de maneira muito mais fácil e simples.

Nome do grupo	Descrição do grupo
Distante	Segmentos distantes de uma crise. Todos os segmentos que não se encaixam em nenhuma das outras três descrições.
Antes	Segmentos próximos antes de uma crise. Todos os segmentos que representam dados de até alguns minutos antes da marcação de início de uma crise.
Após	Segmentos próximos após uma crise. Todos os segmentos que representam dados de até alguns minutos após a marcação de fim de uma crise.
Durante	Segmentos durante uma crise. Todos os segmentos que incluem pelo menos um ponto adquirido após a marcação de início ou antes da marcação de fim de uma crise.

Tabela 2.2: *Descrição dos grupos para análise estatística dos resultados. Esta classificação foi utilizada nas análises tanto no método de DFA quanto no método de detecções de HFO. Foram utilizados nomes ao invés de números para facilitar a compreensão durante a leitura dos resultados.*

2.5 Software Utilizado

Para toda a etapa de processamento dos sinais de EEG deste estudo, foi utilizado um software de desenvolvimento próprio. A ideia da criação deste software veio da necessidade de se obter uma ferramenta suficientemente adaptável para diversos métodos de processamento, alto desempenho e que satisfizesse as necessidades básicas de um especialista em neurologia para visualização dos sinais EEG. O software foi desenvolvido em linguagem C++ para a plataforma Qt, utilizando uma biblioteca para Qt nas construções de gráficos (Qwt) e uma biblioteca para C++ para análises baseadas na transformada de Fourier (FFTW). As versões utilizadas para a compilação do código encontram-se na tabela

Algumas das características do software são

	Versão	Detalhes
MinGW builds	4.7.2	x64 - posix - sjlj
Qt	4.8.2	Compilado em x64
Qwt	6.0.1	Compilado em x64
FFTW	3.3.3	Compilado em x64

Tabela 2.3: Tabela de especificação das ferramentas utilizadas na construção do software.

- Suporte para sinais em formato .edf, facilitando o armazenamento dos sinais através de arquivos binários em um formato open source.
- Seleção de sinais de eletrodos e montagem de canais através de referência única que pode ser alterada em tempo de visualização.
- Visualização e navegação através de 6 canais de forma simultânea, permitindo alteração de escalas em tempo de visualização.
- Aplicação de filtros digitais (FIR: passa baixa, passa alta, passa banda e rejeita banda) em tempo de visualização.
- Renderizações dos sinais e filtragens na visualização são feitas em paralelo na CPU, permitindo uma maior fluidez na utilização das ferramentas de visualização do software.
- Adição de métodos de processamento feita através de plugins, possibilitando maior facilidade no desenvolvimento dos métodos sem a necessidade de alteração do código do software.
- Métodos são aplicados aos canais em paralelo na CPU, utilizando melhor os recursos de computadores mais potentes.
- Resultados são exportados em formato ASCII, facilitando a comunicação com outros softwares de análise, principalmente para análise estatística dos resultados.

O software ainda continua passando por revisões e melhorias, com o ideal de que seja liberado em forma de software open source para utilização geral em processamento de sinais de EEG.

RESULTADOS E DISCUSSÕES

3.1 Testes

Nesta seção, serão mostrados os resultados obtidos para os testes feitos de todos os métodos implementados.

Sendo os métodos e software desenvolvidos por conta própria, esta comparação é importante para a validação da implementação e desenvolvimento da plataforma, garantindo que os dados obtidos não possuem qualquer tipo de erro proveniente de algoritmos, implementação e arquitetura tanto do software quanto dos métodos em si.

3.1.1 Testes de Filtros

Os testes com os filtros implementados serviram simplesmente para a validação da aplicação dos filtros em sinais, obtendo os resultados desejados.

Os parâmetros da resposta em frequência obtida na implementação (tabela 2.1) se mostraram satisfatórios e dentro do que é esperado de um filtro digital de alta qualidade. Porém testes com sinais reais foram feitos para que fosse observada a alteração dos espectros de frequência de sinais de EEG.

Figura 3.1: Alterações nos espectros de potência de sinal EEG após filtragem. As linhas tracejadas horizontais mostram onde foram ajustadas as frequências de corte para o filtro passa baixa e para o passa alta. Ocorre mínima distorção na banda de passagem, e altíssima atenuação nas frequências das bandas de corte.

3.1.2 Testes de Detecção de HFOs

Foram feitos alguns testes simples para validar a capacidade do detector implementado de efetivamente detectar aumentos em amplitude para as atividades de alta frequência. Além disso, os testes serviram para adquirir informação sobre os diferentes parâmetros iniciais do método de detecção e seus efeitos sobre os resultados.

Foi criado um sinal simulado contendo componentes senoidais de frequência baixa (10, 20, 30 e 40Hz) e amplitude $A_1 = 100$. Ao sinal resultante, foram adicionados pequenos trechos contendo componentes senoidais de alta frequência.

As HFOs simuladas foram adicionadas em intervalos de tempo diferentes e possuem diferentes frequências, variando de 80 a 120Hz, de 2 em 2Hz.

O método foi configurado como indicado na seção 2.4, com a banda de interesse fixada de 80 a 120Hz. A figura 3.2 mostra os valores do *threshold* móvel, do RMS calculado e indica os valores de detecção.

É possível observar na figura 3.2 que todas as senóides de alta frequência foram detectadas, confirmando a capacidade que o detector possui de identificar o aumento da potência na banda de interesse. Porém, é possível notar que o valor de *threshold* precisa de algumas iterações para estabilizar, sendo que enquanto esta estabilização ocorre, são feitas várias detecções de falsos positivos.

Isto ocorre porque o valor inicial para o *threshold* foi fixado em $B_k[0] = RMS[0]$ (valor inicial de *threshold* é igual o valor inicial de RMS). Com isso, sempre existirá a possibilidade de que, enquanto o valor de *threshold* cresce, pequenas oscilações no RMS poderão ser indicadas como HFOs.

Para reduzir o número de detecções falsas no início de cada sinal, adotou-se a atribuição do valor inicial para $B_k[0] = 2RMS[0]$, ou seja, o valor inicial do *threshold* será atribuído já acima do valor inicial do RMS.

Pode-se perceber que o valor do *threshold* continua apresentando um comportamento crescente antes de se estabilizar, sendo que a estabilização ocorre praticamente no mesmo ponto no tempo (cerca de $3,5 \cdot 10^4$), mas o valor para o qual o *threshold* tende é maior. Isso indica que a alteração no valor inicial do *threshold* não altera o número de iterações necessárias para a estabilização, mas pode alterar

Figura 3.2: Teste de detecções de HFOs para sinal simulado. Em (A), pode-se visualizar os valores calculados para o RMS e o threshold para cada amostra (ponto) do sinal, com as detecções indicadas pela linha vermelha. Em (B), uma aproximação visual dos primeiros pontos do gráfico. Percebe-se que, enquanto o valor de threshold não alcança o equilíbrio, as pequenas oscilações de RMS são indicadas como HFOs. Este comportamento faz com que as detecções no início de cada sinal analisado precisem ser ignoradas.

significativamente os valores para o *threshold*. Com isso, concluímos que a alteração dos valores iniciais pode também alterar a sensibilidade do algoritmo.

Após o ajuste no valor inicial do *threshold*, foi feito um teste com a adição de ruído branco no sinal simulado.

Apesar de se tratar de ruído branco, com média zero e distribuição uniforme, este teste pode ajudar a prever o comportamento do detector em um situações onde a informação a ser extraída encontra-se embutida em um comportamento menos ordenado.

Foram feitas 5 detecções para cada nível de amplitude de ruído, e os

Figura 3.3: Detecção de HFO com alteração do valor de threshold. Em (A), percebe-se que as detecções continuam ocorrendo, porém não há falsos positivos no início do sinal. Este comportamento segue o que era esperado pela alteração do valor inicial de threshold. Em (B), percebe-se que os valores para o threshold quando $B_k[0] = \text{RMS}[0]$ são menores do que os valores de $B_k[0] = 3\text{RMS}[0]$. Com isso, é possível não só evitar algumas detecções com pequenas oscilações de RMS como também pode-se alterar a sensibilidade do algoritmo. Quanto maior o valor inicial de $B_k[0]$, maior os valores para o threshold e, consequentemente, maior deverá ser o valor de RMS para ocorrer uma detecção.

resultados da taxa de detecção média ($\frac{N^{\circ} \text{detecções}}{N^{\circ} \text{oscilações geradas}}$) e os desvios das taxas de detecção são visualizados na tabela 3.1.

Amplitude Ruído	SNR	Taxa Detecção Média (%)	Desvio Padrão
1	20.000	1.0000	0.0000
2	13.979	1.0000	0.0000
3	10.458	1.0000	0.0000
4	7.959	1.0000	0.0000
5	6.021	0.8857	0.1801
6	4.437	0.8952	0.0782
7	3.098	0.8667	0.0621
8	1.938	0.7048	0.0522
9	0.915	0.5238	0.1615
10	0.000	0.1333	0.0398
15	-3.522	0.0476	0.0583
20	-6.021	0.0000	0.0000

Tabela 3.1: Tabela de dados da taxa de detecção em função da amplitude do ruído. SNR é a relação sinal-ruído calculada como $SNR = 20 \log \left(\frac{Amp_{Sinal}}{Amp_{Ruído}} \right)$. foram feitas 5 repetições em cada caso.

Figura 3.4: Teste de detecções de HFOs em função do nível de ruído. SNR é a relação sinal-ruído calculada como $SNR = 20 \log \left(\frac{Amp_{Sinal}}{Amp_{Ruido}} \right)$. Foram feitas 5 repetições, e as barras de erro no gráfico representam o desvio padrão. Repare que começa a ocorrer queda significativa do número de detecções para valores de SNR 5, ou seja, quando a amplitude do ruído chega a metade da amplitude das HFOs geradas.

Como pode-se perceber, a partir de um certo valor de amplitude de ruído adicionado no sinal ($Amp_{ruído} \approx 0.5 Amp_{sinal}$), há uma queda considerável na capacidade do detector de identificar as HFOs. Isso era esperado devido ao fato de que a presença das componentes de alta frequência do ruído adicionado contribuem para o aumento do valor de *threshold*. O importante deste teste foi obter justamente o valor aproximado de razão entre amplitude de sinal e amplitude de *background* em que a queda significativa de performance ocorre.

Foram executados também testes de desempenho do detector em relação à quantidade de oscilações presentes. É possível imaginar que a presença de uma grande quantidade de HFOs pode aumentar o valor do *threshold*. Com isso, a presença de várias HFOs em um curto intervalo de tempo poderia resultar na ocorrência de falsos negativos, ou seja, o detector deixaria de indicar a presença de um aumento de potência na banda de interesse.

Taxa HFOs	Taxa Detecção (%)	Taxa HFOs	Taxa Detecção (%)
0.1	1.000	2.4	0.525
0.5	1.000	2.5	0.496
1.0	1.000	2.6	0.473
1.5	1.000	2.7	0.441
1.7	1.000	2.8	0.400
1.8	0.889	2.9	0.386
1.9	0.779	3.0	0.340
2.0	0.705	3.1	0.297
2.1	0.648	3.2	0.200
2.2	0.600	3.2	0.058
2.3	0.561	3.4	0.000

Tabela 3.2: Tabela de dados da taxa de detecção em função da taxa de HFOs por segundo. Não foi feita nenhuma repetição no caso, pois a detecção não envolveu geração de ruído aleatório.

Figura 3.5: *Teste de detecções de HFOs em função da taxa de HFO por segundo. Repare que existem dois pontos de flexão do gráfico. Para taxas de HFO a partir de 1.7HFO/s , começa a haver um decaimento exponencial na taxa de detecção. Para taxas a partir de $\sim 3.1\text{HFO/s}$, o coeficiente de decaimento aumenta. Para taxas a partir de HFO/s , o número de HFOs por segundo é tão alto que o valor de threshold é alterado de forma que nenhuma detecção é feita. A partir deste ponto, o algoritmo considera as HFOs como parte do background, e não mais se destacando deste.*

Como era esperado, com o aumento do valor de oscilações por segundo adicionadas ao sinal, ocorre uma queda de desempenho do algoritmo de detecção. O que surpreende desses resultados é a maneira com que a queda ocorre.

Existe um platô para valores abaixo de 1.7 oscilações por segundo. A partir deste valor, a taxa de detecção decai exponencialmente. O mais curioso é que, após outro certo valor de oscilações por segundo (3.1), a queda ocorre de maneira mais rápida.

O acontecimento do platô e a queda se aproxima muito do que era esperado. O valor de *threshold* sofre aumento com o aumento da taxa de HFOs, porém este aumento só é suficientemente alto para prejudicar as detecções a partir de $1.7HFO/s$.

É importante citar que as detecções deixam de ocorrer das ultimas para as primeiras. Ou seja, nos primeiros valores de taxa de HFOs que são suficientes para atrapalhar as detecções, apenas as últimas oscilações geradas deixam de ser detectadas. Isto também era esperado e indica que o valor de *threshold* aumenta gradativamente a cada HFO, até ultrapassar o valor de RMS.

3.1.3 Testes de DFA

Os testes de DFA tiveram como objetivo validar a implementação, comparando os resultados obtidos com sinais simulados com o conhecimento já registrado na literatura. Foram também ajustados parâmetros para melhor adaptar o método às necessidades do trabalho.

Primeiramente, foram gerados ruídos branco, browniano e $1/f$ (*flicker*, ruído rosa) de tamanho $N = 100000$ pontos, e o coeficiente α foi obtido para cada sinal gerado. Pode-se comparar os resultados obtidos com os resultados esperados, como feito na tabela 3.3. Os dados obtidos são muito próximos dos dados esperados, validando a implementação do método. Apenas um α foi obtido, pois se tratando de um tamanho de sinal elevado, a ocorrência do *crossover* não é significativa.

Foram então feitos testes variando-se o tamanho dos sinais. A importância deste teste está na forma como o método da DFA deverá ser aplicado neste estudo. Quanto menor o tamanho dos segmentos usados para calcular a DFA, melhor a

	Branco	$1/f$	Browniano
Média	0.50138	0.98363	1.48455
Desvio Pad.	0.01168	0.01693	0.02129
Valor esperado para Média	0.5	1	1.5

Tabela 3.3: Tabela de dados de DFA com sinais simulados. A tabela apresenta dados obtidos com 200 sinais gerados de cada tipo, com tamanho $N = 100000$ pontos. Os valores encontram-se muito próximos dos valores esperados, validando o método. Não foram obtidos dois coeficientes pois, se tratando de sinais com tamanho grande, a ocorrência do crossover é pequena.

resolução temporal do método. Porém, imagina-se que a utilização de tamanhos pequenos de sinal pode prejudicar a capacidade de obter valores precisos para estimativa do coeficiente α .

Foram gerados sinais simulados de diferentes tamanhos e foram comparados os valores de média e desvio padrão para cada resultado. Esta comparação pode ser vista na figura 3.6.

Mas não basta analisar apenas os resultados obtidos com um sinal simulado. É preciso também observar como o tamanho da segmentação de um sinal de EEG para obtenção dos valores de α poderia influenciar a sensibilidade do método. Tamanhos maiores garantiriam a precisão da estimativa do coeficiente, mas é possível que, ao analisar um sinal de maior tamanho, alterações importantes na característica do sinal, com curta duração, possam ser mascaradas.

Na imagem 3.7, é perceptível que o aumento do tamanho dos sinais reduz a diferença observada na presença da crise epiléptica. Com isso, deve-se levar em conta o compromisso precisão-sensibilidade temporal para fixar um tamanho que melhor se ajuste às necessidades do estudo.

Figura 3.6: Médias e desvios do expoente α em função do tamanho do sinal. Percebe-se que o valor das médias se aproximam do valor esperado (linhas sólidas) quando o tamanho do sinal é maior. Além disso, observa-se que o desvio padrão aumenta com a diminuição do tamanho do sinal, conforme o esperado. Foram gerados 200 sinais para cada tipo de ruído e valor de tamanho.

Figura 3.7: *Influência do tamanho do segmento nos valores de α em um sinal de EEG. Da esquerda para a direita, de cima para baixo, respectivamente: $N = 1000$, $N = 5000$, $N = 10000$, $N = 15000$, $N = 20000$, $N = 25000$. Observa-se que, com o aumento do tamanho de N , fica mais difícil identificar as variações que parecem ocorrer próximas ao momento da crise. Os dados mostrados neste gráfico sugerem que realmente seria possível identificar uma crise com o valor de DFA em intervalos no tempo, uma vez ajustado o valor de N .*

Com os resultados obtidos dos dois testes (ruídos simulados e sinal EEG), fixou-se o valor do tamanho de segmento $N = 5000$ para a análise dos dados.

3.2 Resultados

Nesta seção, serão apresentados os resultados obtidos das aplicações dos métodos nos sinais de EEG de pacientes com epilepsia (dados descritos em 2.1).

De todos os arquivos disponíveis no banco de dados, somente alguns foram selecionados. A seleção dos arquivos do banco de dados é apresentada no **apêndice A**.

3.2.1 DFA

Como explicado na seção 2.3, os dados de EEG foram divididos em segmentos de tamanho $N = 5000$, e para cada um deste segmento, foram obtidos os valores de α_1 e α_2 .

Os valores foram separados, então, dependendo da proximidade com as crises, de acordo com a tabela 2.2.

Na figura 3.8 são apresentados os valores de α_1 e α_2 para as primeiras crises registradas de cada paciente analisado.

Os resultados podem ser vistos de duas formas: No primeiro gráfico de cada imagem, são mostrados os valores de ambos os coeficientes em função do tempo. No segundo gráfico, são mostrados os valores de dispersão de α_2 por α_1 . Este gráfico de dispersão é utilizado para facilitar a visualização das diferenças entre os valores para cada um dos dois coeficientes em relação à classificação dos segmentos que os valores representam. Nos gráficos de coeficientes em função do tempo, as cores mais escuras representam valores de α_2 e as cores mais claras representam valores de α_1 .

Figura 3.8: Exemplos de dados obtidos pelo método da DFA para diferentes pacientes. Aqui é mostrada a primeira (ou primeiras) crises registradas para cada paciente que teve os sinais utilizados. Os gráficos A) representam os valores dos coeficientes de cada segmento em função da posição do segmento no tempo. As cores claras representam α_1 e as cores escuras representam α_2 . Nos gráficos B) são apresentadas as distribuições de α_1 e α_2 para cada segmento. É perceptível que, enquanto em alguns casos os coeficientes próximos e durante as crises são visivelmente diferentes, em outros casos a diferenciação não é feita facilmente. Essa diferença de sinal para sinal e paciente para paciente pode sugerir que não é o mais indicado estudar todos os pacientes de uma forma só, sendo talvez necessária a análise dos dados de forma separada para cada paciente, possivelmente sendo melhor ainda a separação da análise de cada crise, mesmo para o mesmo paciente. Os pacientes são, da esquerda para a direita, de cima para baixo: chb01, chb02, chb03, chb04, chb05, chb06, chb07, chb08, chb09, chb23.

Apesar de meramente ilustrativa, esta imagem já pode indicar a capacidade de diferenciação entre as distribuições dos valores dos coeficientes. É especialmente perceptível nos gráficos de dispersão que os valores de α_1 nos segmentos durante a crise parecem ser levemente maiores que os segmentos distantes à crise, enquanto os

Figura 3.9: Dados de α_2 vs α_1 para os segmentos de EEG epiléptico. A) Segmentos distantes à crise. B) Segmentos antes da crise (até 3 minutos antes). C) Segmentos após a crise (até 3 minutos após). D) Segmentos durante a crise. Para B) e C), foram considerados pontos de até 3 minutos antes do começo ou após o final de cada crise, respectivamente. É possível perceber que, com a proximidade da crise, ocorre uma pequena localização dos valores. Na crise, os valores de α_2 são menores e os valores de α_1 são maiores quando comparados aos segmentos distantes. Os resultados de significância são mostrados na tabela 3.4

Figura 3.10: Média e desvio padrao para os dados de α_2 vs α_1 para os segmentos de EEG epiléptico. É possível notar que cada distribuição possui pequena diferença das demais. Os valores de α_1 tem menor média antes da crise e maior média após a crise. Já os valores de α_2 tem maior média após a crise e menor média durante a crise. O grupo que visualmente mais se destaca dos demais, como era esperado, corresponde aos valores obtidos durante a crise.

As figuras 3.9 e 3.10 ilustram de uma maneira melhor as diferenças observadas entre as distribuições.

Em 3.9 são apresentadas as distribuições para cada uma das classificações de forma separada, juntamente com as médias e os desvios padrões para valores dos dois coeficientes. Em 3.10 são apresentados apenas os valores de média e desvio padrão de cada distribuição. É possível perceber de forma mais clara que realmente existem diferenças entre as distribuições.

Foram feitos testes de significância para comprovar a hipótese de que os conjuntos de dados de cada classificação são gerados por distribuições diferentes. Para isso, foi utilizado um teste t, sem considerar que as variâncias entre os dois grupos seriam iguais (também conhecido como teste t de Welch). A hipótese do teste é que os grupos de dados seriam gerados pela mesma distribuição. Com $p < 0.01$, esta hipótese é rejeitada com 5% de confiança, ou seja, considera-se que os dados possuem distribuições estatisticamente diferentes.

Os resultados de significância são positivos, indicando que todas as classificações são distinguíveis entre si.

			Distante	Antes	Após	Durante
	Média	σ	p<0.01	p<0.01	p<0.01	p<0.01
α_1						
Distante	1.3587	0.3493	-	s	s	s
Antes	1.3383	0.2913	s	-	s	s
Após	1.4234	0.3041	s	s	-	s
Durante	1.4061	0.2850	s	s	s	-
α_2						
Distante	0.6165	0.1918	-	s	s	s
Antes	0.6626	0.1675	s	-	s	s
Após	0.7105	0.1823	s	s	-	s
Durante	0.5107	0.1863	s	s	s	-

Tabela 3.4: *Dados de média, desvio padrão e teste de significância das distribuições de α_1 e α_2 para todos os sinais de EEG. O teste indica que é possível diferenciar todas as distribuições entre si. Esse resultado sugere que podem ser possíveis feitos como a identificação de crises em sinais não marcados ou até a predição de crises com métodos baseados na DFA.*

A facilidade de se distinguir as amostras estatisticamente apesar de as distribuições parecerem possuir variações pequenas de média e desvio padrão podem estar relacionadas ao tamanho de cada amostra. Como foram utilizados todos os dados de todos os registros de todos os pacientes para a análise estatística, o número de dados disponíveis para análise é extremamente grande, como mostrado na tabela 3.5.

Grupo	N^o de segmentos
Distante	322161
Antes	10395
Após	10353
Durante	4221

Tabela 3.5: *Número de segmentos em cada grupo de dados para a análise dos resultados do teste de DFA. O menor número de segmentos corresponde também ao grupo que mais se diferencia dos demais. O tamanho de cada grupo pode justificar a facilidade de se diferenciar as distribuições, mesmo com médias aparentemente não tão significantes.*

Os dados desta tabela são obtidos quando utiliza-se o tamanho de janela (intervalo de tempo máximo para que um segmento seja classificado como antes ou após a crise) de $3min$.

Foi pensada a possibilidade de que diferentes tamanhos desta janela pudessem resultar em diferenças nas distribuições. Os resultados obtidos com diferentes tamanhos de janela podem ser visualizados na figura 3.14.

É possível notar através da figura 3.11 que, quanto maior o tamanho da janela utilizada, mais os dados da distribuição dos segmentos após a crise se destacam. Porém, como já é possível diferenciar as distribuições no aspecto estatístico, nenhuma alteração na metodologia foi feita. Continua-se então com tamanho de janela de $3min$ antes e após a crise. O mesmo teste foi repetido para as taxas de detecção de HFOs e os resultados mostraram que tamanhos menores ajudavam a diferenciar as distribuições de resultados para aquele método (Discussão sobre detecção de HFOs na seção 3.2.2. Dados de tamanho da janela na figura 3.14).

Figura 3.11: Impacto do tamanho da janela nas distribuições de α_1 e α_2 . Da esquerda para a direita, de cima para baixo: 1 minuto, 2 minutos, 3 minutos, 4 minutos, 5 minutos, 6 minutos, 7 minutos, 8 minutos, 9 minutos e 10 minutos antes do início e após o fim de cada crise. Repare que quanto maior o valor da janela de tempo analisado antes e após a crise, melhor se destaca a distribuição dos coeficientes obtidos após.

3.2.2 Detecção de HFOs

As análises das detecções de HFO seguem o mesmo padrão que as análises para a DFA. Porém, apenas um valor é obtido. Este valor reflete a média do aumento na atividade de alta frequência dentro da janela de tamanho $N = 5000$ e os segmentos obtidos são classificados de acordo com a proximidade da ocorrência de uma crise (tabela 2.2).

Figura 3.12: Exemplos de dados obtidos pela detecção de HFOs para diferentes pacientes. Aqui é mostrada a primeira (ou primeiras) crises registradas para cada paciente que teve os sinais utilizados. É visível que na grande maioria dos casos a atividade em alta frequência durante a crise é muito maior quando comparada à atividade para os segmentos distantes. Porém, assim como na DFA, os resultados obtidos podem variar muito de paciente para paciente, indicando que talvez a melhor forma de proceder para análises mais detalhadas seja trabalhar com os dados de forma separada. Os pacientes são, da esquerda para a direita, de cima para baixo: chb01, chb02, chb03, chb04, chb05, chb06, chb07, chb08, chb09, chb23.

A figura 3.12 ilustra as séries de dados obtidas de taxa de detecção de HFOs para cada segmento. Nos graficos A) de cada imagem é apresentada a distribuição dos valores de taxa para cada segmento em função do tempo. Já nos gráficos B) são apresentados os valores de média e desvio padrão para as distribuições de cada classificação dos segmentos.

Figura 3.13: Média e desvio padrao da taxa de ocorrência de detecções de HFO para os segmentos de EEG epiléptico. Para as distribuições de antes e após a crise, foram considerados segmentos de até 3 minutos antes do início ou após o fim da crise, respectivamente. É possível notar que existe uma diferença na média dos valores, em especial para os valores durante a crise. Os desvios padrões possuem comportamento esperado, sendo menor no conjunto com maior número de dados (distante à crise) e maior no conjunto com menor número de dados (durante a crise).

É visível que as distribuições dos conjuntos são facilmente diferenciadas, com exceção das distribuições de segmentos antes e após as crises. Como era esperado, a ocorrência de crises aumenta significativamente a atividade em alta frequência. Muito do que se pode ver poderia ser considerado artefato muscular, durante uma crise, por isso a distribuição dos segmentos durante as crises se destaca mais.

Os testes de significância na diferenciação das distribuições é apresentado na

tabla 3.6. Novamente, todas as distribuições podem ser diferenciadas entre si.

			Distante	Antes	Após	Durante
	Média	σ	p<0.01	p<0.01	p<0.01	p<0.01
Distante	0.0134	0.0503	-	s	s	s
Antes	0.0201	0.0613	s	-	s	s
Após	0.0247	0.0585	s	s	-	s
Durante	0.0747	0.1638	s	s	s	-

Tabela 3.6: *Dados de média, desvio padrão e teste de significância das distribuições de taxa de ocorrência de detecções para todos os sinais de EEG. Assim como os resultados com o método da DFA, todas as distribuições puderam ser diferenciadas entre si.*

A facilidade na diferenciação também pode se explicar pelo tamanho das amostras, como no método de DFA (tabela 3.5). Os dados foram trabalhados de forma que os tamanhos dos segmentos tanto para o método de DFA quanto para o método de detecções de HFO sejam os mesmos, sendo assim, o número de segmentos total e consequentemente o número de segmentos em cada grupo também são os mesmos.

Ainda assim, é perceptível que os valores de média e desvio para as distribuições possuem diferenças mais significativas do que os resultados mostrados com a análise de DFA.

Também foram gerados dados para diferentes tamanhos de janela. Os dados são apresentados na figura 3.14, e indicam que, quanto menor a janela, melhor é a diferenciação aparente entre as distribuições.

Isso pode ser explicado pelo fato de que os mecanismos de geração de HFOs iniciam sua atividade próximo à ocorrência da crise, e um aumento no tamanho da janela simplesmente estaria considerando tanto segmentos onde as oscilações já estariam sendo geradas em associação à crise quanto segmentos em que ainda não há geração de HFOs, fazendo com que a diferença entre as distribuições seja reduzida.

Figura 3.14: Impacto do tamanho da janela nas distribuições de taxa de detecção de HFOs. Da esquerda para a direita, de cima para baixo: 1 minuto, 2 minutos, 3 minutos, 4 minutos, 5 minutos, 6 minutos, 7 minutos, 8 minutos, 9 minutos e 10 minutos antes do início e após o fim de cada crise. É visível que quanto maior a janela, maior a proximidade entre as distribuições de segmentos antes e após a crise com os segmentos distantes da crise. Pode-se concluir que os melhores valores de janela para se analisar estes dados são os valores mais baixos, e valores mais altos da janela misturam dados de distribuições diferentes. O valor de 3min foi adotado para facilitar comparações com o método da DFA.

CONCLUSÕES

De maneira geral, pode-se concluir que a hipótese proposta do estudo foi investigada de tal forma a contribuir para análise de sinais de EEG com aplicações para a epilepsia.

Os resultados sugerem que seja possível que os mecanismos geradores das crises epiléticas podem também ser causadores de variações nas leis de potência que regem a correlação a longo termo dos sinais de atividade elétrica do cérebro. Sendo assim, o desenvolvimento de métodos baseados na obtenção dos expoentes através da Detrended Fluctuation Analysis podem contribuir para áreas de detecção de crises, determinação da zona epileptogênica através da criação de mapas para a localização espacial, dentre outras.

Porém, para que estas aplicações tenham potencial comprovado, são necessários mais estudos para investigar correlações em níveis mais específicos, como o cruzamento de informações de cada eletrodo com a proximidade do eletrodo de uma zona epileptogênica, por exemplo. A necessidade desta especificidade também se justifica na observação de que os dados com conjuntos muito grandes envolvendo muitos pacientes podem refletir dinâmicas que não necessariamente se aplicam a todos os pacientes envolvidos.

Os resultados sugerem que é possível que os valores dos coeficientes de DFA possam ser utilizados para cálculo de um valor de *threshold*, parecido com as detecções de HFO. Assim apenas trechos com variação significantes no valor dos coeficientes seriam levados em consideração, facilitando a visualização e localização de alterações de estado significantes, e as comparações de significância seriam feitas

utilizando apenas os dados passados da serie de cada paciente, garantindo maior especificidade de análise.

Estas aprofundações nas investigações exigiriam diversas condições que não foram disponíveis no desenvolvimento deste estudo. Além da informação sobre a localização da zona epileptogênica, seriam necessárias a seleção de trechos sem nenhum tipo de artefato muscular, informações sobre estado em que se encontram os pacientes (sono, vigília, movimentos, etc), entre outras. Estas informações seriam de extrema utilidade para seleção dos dados a serem processados e comparados.

O desempenho do software desenvolvido para este projeto se mostrou extremamente satisfatório. A utilização de computação paralela contribuiu para que os métodos fossem executados e testados em tempo mais curto e a utilização de *plugins* para adição dos métodos de processamento contribuiu para a maior facilidade nas alterações e adaptações julgadas necessárias durante o desenvolvimento do projeto. Além disso, a implementação dos métodos de processamento contribuiu de forma excelente para que o autor da pesquisa se familiarizasse com o funcionamento dos mesmos, coisa que não ocorreria na utilização de métodos ou bibliotecas já prontos para processamento de sinais.

É importante citar que as oscilações de alta frequência detectadas não passaram por teste de validação através de comparação com marcações de um especialista. Com isso, as detecções feitas simplesmente refletem um aumento na atividade cerebral em alta frequência, e não refletem de forma direta a presença dos padrões de HFO destacados na literatura. Ainda assim os resultados utilizando os dados de detecção são bastante condizentes com o que encontra-se em estudos a respeito de HFOs, com a presença das crises aumentando a ocorrência de atividade de alta frequência.

Para as perspectivas de estudos futuros podem ser citadas investigações de forma individual de sinais obtidos de pacientes, com informações a respeito da localização da zona epileptogênica, seleção de dados feita por especialista para padronizar o estado dos pacientes a serem estudados, além da validação das detecções como padrões de HFO.

Vale a pena destacar que estudos com sinais de EEG de indivíduos saudáveis e sinais de EEG intracortical podem trazer grandes benefícios para as análises

propostas por este estudo. Os sinais de EEG intracorticais podem refletir os mecanismos de atividade cerebral de forma mais direta, podendo conter informações atenuadas ou inexistentes em EEG de escalpo. Já o estudo em indivíduos saudáveis pode auxiliar a validação de métodos como a análise de DFA proposta para a associação de alterações dos coeficientes com a presença de foco epiléptico.

Existe também a necessidade de continuar melhorando o software desenvolvido, para melhor adaptação às necessidades de potenciais usuários e para oferecer uma maior gama de funções, principalmente na visualização e análise de resultados.

REFERÊNCIAS BIBLIOGRÁFICAS*

- [1] REYNOLDS, E. H. The ILAE/IBE/WHO Global Campaign against Epilepsy: Bringing Epilepsy 'Out of the Shadows'. *Epilepsy & behavior : E&B*, v. 1, n. 4, p. S3–S8, ago. 2000.
ISSN 1525-5069.
- [2] TEMKIN, O. *The Falling Sickness: A History of Epilepsy from the Greeks to the Beginnings of Modern Neurology (Softshell Books)*. Johns Hopkins University Press, 1994. 488 p. ISBN 0801848490.
- [3] FISHER, R. S. et al. Epileptic seizures and epilepsy: definitions proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE). *Epilepsia*, v. 46, n. 4, p. 470–2, abr. 2005.
ISSN 0013-9580.
- [4] WHO, W. H. O. WHO | Epilepsy: historical overview.
World Health Organization.
- [5] BERG, A. T. et al. Revised terminology and concepts for organization of seizures and epilepsies: report of the ILAE Commission on Classification and Terminology, 2005-2009. *Epilepsia*, v. 51, n. 4, p. 676–85, abr. 2010.
ISSN 1528-1167.
- [6] PANAYIOTOPOULOS, C. P. The new ILAE report on terminology and concepts for the organization of epilepsies: critical review and contribution. *Epilepsia*, v. 53, n. 3, p. 399–404, mar. 2012.
ISSN 1528-1167.

*De acordo com a Associação Brasileira de Normas Técnicas. NBR 6023.

-
- [7] ENGEL, J. et al. *Epilepsy: A Comprehensive Textbook (3-volume set)*. Lippincott Williams & Wilkins, 2007. 3056 p. ISBN 0781757770.
- [8] KWAN, P.; SANDER, J. W. The natural history of epilepsy: an epidemiological view. *Journal of neurology, neurosurgery, and psychiatry*, v. 75, n. 10, p. 1376–81, out. 2004.
- ISSN 0022-3050.
- [9] BRODIE, M. J. Diagnosing and predicting refractory epilepsy. *Acta neurologica Scandinavica. Supplementum*, v. 181, p. 36–9, jan. 2005.
- ISSN 0065-1427.
- [10] BELEZA, P. Refractory epilepsy: a clinically oriented review. *European neurology*, v. 62, n. 2, p. 65–71, jan. 2009.
- ISSN 1421-9913.
- [11] ROSENOW, F.; LUDERS, H. Presurgical evaluation of epilepsy. *Brain*, v. 124, n. 9, p. 1683–1700, 2001.
- [12] MORMANN, F. et al. Epileptic seizures are preceded by a decrease in synchronization. *Epilepsy research*, v. 53, n. 3, p. 173–85, mar. 2003.
- ISSN 0920-1211.
- [13] DURAND, D.; BIKSON, M. Suppression and control of epileptiform activity by electrical stimulation: a review. *Proceedings of the IEEE*, v. 89, n. 7, p. 1065–1082, jul. 2001.
- ISSN 00189219.
- [14] CATON, R. The electric currents of the brain. *British Medical Journal*, v. 2, p. p. 278, 1875.
- [15] KOCH, C. Complexity and the Nervous System. *Science*, v. 284, n. 5411, p. 96–98, abr. 1999.
- ISSN 00368075.

- [16] LUDANK, S. Testing for nonlinearity in time series: the method of surrogate data. v. 58, p. 77–94, 1992.
- [17] SCHREIBER, T.; SCHMITZ, A. Surrogate time series. *Physica D: Nonlinear Phenomena*, v. 142, n. 3-4, p. 346–382, ago. 2000.
ISSN 01672789.
- [18] KUNHIMANGALAM, R.; JOSEPH, P.; SUJITH, O. Nonlinear analysis of EEG signals: Surrogate data analysis. *Irbm*, v. 29, n. 4, p. 239–244, set. 2008.
ISSN 19590318.
- [19] STAM, C. J. Nonlinear dynamical analysis of EEG and MEG: review of an emerging field. *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology*, v. 116, n. 10, p. 2266–301, out. 2005.
ISSN 1388-2457.
- [20] LEHNERTZ, K. et al. Nonlinear EEG Analysis in Epilepsy :. v. 18, n. 3, p. 209–222, 2001.
- [21] BLANCO, S. et al. Stationarity of the EEG series. *IEEE Engineering in Medicine and Biology Magazine*, v. 14, n. 4, p. 395–399, 1995.
ISSN 07395175.
- [22] NIEDERMEYER, E.; SILVA, F. L. da. *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. Philadelphia, PA 19106 USA: Lippincott Williams & Wilkins, 2005. ISBN 0781751268.
- [23] SUBHA, D. P. et al. EEG signal analysis: a survey. *Journal of medical systems*, v. 34, n. 2, p. 195–212, abr. 2010.
ISSN 0148-5598.
- [24] FINGELKURTS, A. A.; FINGELKURTS, A. A.; KAPLAN, A. Y. Interictal EEG as a physiological adaptation. Part II. Topographic variability of composition of brain oscillations in interictal EEG. *Clinical neurophysiology : official journal of*

the International Federation of Clinical Neurophysiology, v. 117, n. 4, p. 789–802, abr. 2006.

ISSN 1388-2457.

- [25] BRAGIN, a. et al. Hippocampal and entorhinal cortex high-frequency oscillations (100–500 Hz) in human epileptic brain and in kainic acid-treated rats with chronic seizures. *Epilepsia*, v. 40, n. 2, p. 127–37, fev. 1999.

ISSN 0013-9580.

- [26] ANDRADE-VALENCA, L. et al. Interictal high frequency oscillations (HFOs) in patients with focal epilepsy and normal MRI. *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology*, v. 123, n. 1, p. 100–5, jan. 2012.

ISSN 1872-8952.

- [27] ANDRADE-VALENCA, L. P. et al. Interictal scalp fast oscillations as a marker of the seizure onset zone. *Neurology*, v. 77, n. 6, p. 524–31, ago. 2011.

ISSN 1526-632X.

- [28] BUZSÁKI, G.; SILVA, F. L. da. High frequency oscillations in the intact brain. *Progress in neurobiology*, v. 98, n. 3, p. 241–9, set. 2012.

ISSN 1873-5118.

- [29] ENGEL, J.; SILVA, F. L. da. High-frequency oscillations - where we are and where we need to go. *Progress in neurobiology*, v. 98, n. 3, p. 316–8, set. 2012.

ISSN 1873-5118.

- [30] JACOBS, J. et al. High-frequency oscillations (HFOs) in clinical epilepsy. *Progress in neurobiology*, v. 98, n. 3, p. 302–15, set. 2012.

ISSN 1873-5118.

- [31] CHEN, Z. et al. Effect of nonstationarities on detrended fluctuation analysis. *Physical review. E, Statistical, nonlinear, and soft matter physics*, v. 65, n. 4 Pt 1, p. 041107, abr. 2002.

ISSN 1539-3755.

- [32] PENG, C. K. et al. Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. *Chaos (Woodbury, N.Y.)*, v. 5, n. 1, p. 82–7, jan. 1995.

ISSN 1054-1500.

- [33] LIN, L.-Y. et al. Detrended fluctuation analysis predicts successful defibrillation for out-of-hospital ventricular fibrillation cardiac arrest. *Resuscitation*, v. 81, n. 3, p. 297–301, mar. 2010.

ISSN 1873-1570.

- [34] ZHENG, H.; SONG, W.; WANG, J. Detrended fluctuation analysis of forest fires and related weather parameters. *Physica A: Statistical Mechanics and its Applications*, v. 387, n. 8-9, p. 2091–2099, mar. 2008.

ISSN 03784371.

- [35] LACHOWYCZ, S. M. et al. Long-range correlations identified in time-series of volcano seismicity during dome-forming eruptions using detrended fluctuation analysis. *Journal of Volcanology and Geothermal Research*, Elsevier B.V., v. 264, p. 197–209, ago. 2013.

ISSN 03770273.

- [36] BLESÍĆ, S. et al. Detecting long-range correlations in time series of dorsal horn neuron discharges. *Annals of the New York Academy of Sciences*, v. 1048, p. 385–91, jun. 2005.

ISSN 0077-8923.

- [37] LEE, J. M. et al. Detrended fluctuation analysis of EEG in sleep apnea using MIT/BIH polysomnography data. *Computers in biology and medicine*, v. 32, n. 1, p. 37–47, jan. 2002.

ISSN 0010-4825.

- [38] BOCHER, M. Introduction to the Theory of Fourier 's Series. *Annals of Mathematics*, v. 7, n. 3, p. 81–152, 1906.

- [39] SMITH, S. W. *The Scientist and Engineer's Guide to Digital Signal Processing*.
- [40] ROY, S. D.; KUMAR, B.; JAIN, B. Fir notch filter design: A review. *Facta universitatis-series: Eletronics and Energetics*, v. 14, n. 3, p. p295–327, 2001.
- [41] ZELMANN, R. et al. A comparison between detectors of high frequency oscillations. *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology*, International Federation of Clinical Neurophysiology, v. 123, n. 1, p. 106–16, jan. 2012.
ISSN 1872-8952.
- [42] ELLENRIEDER, N. von et al. Automatic detection of fast oscillations (40-200 Hz) in scalp EEG recordings. *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology*, v. 123, n. 4, p. 670–80, abr. 2012.
ISSN 1872-8952.
- [43] SHOEB, A. H. *Application of Machine Learning to Epileptic Seizure Onset Detection and Treatment Ali Hossam Shoeb*. 162 p. Tese (PhD Thesis) — Massachusetts Institute of Technology, 2009.
- [44] GOLDBERGER, A. L. et al. PhysioBank, PhysioToolkit, and PhysioNet : Components of a New Research Resource for Complex Physiologic Signals. *Circulation*, v. 101, n. 23, p. e215–e220, jun. 2000.
ISSN 0009-7322.

DADOS UTILIZADOS CHB-MIT

Como discutido na apresentação dos resultados, nem todos os dados do banco de dados CHB-MIT foram considerados para estudo. A seleção dos dados foi feita levando-se em consideração os seguintes fatores:

- Durante a aquisição dos dados do arquivo, deve ter ocorrido ao menos uma crise.
- Não há eletrodos extra, adicionados durante a aquisição para fins de testes específicos.

A justificativa para esta seleção está em obter maior facilidade para carregar, processar os dados e comparar os resultados entre os diferentes arquivos, eliminando todos os arquivos que não possuem crise para a redução na quantidade de dados a ser processada. A eliminação dos arquivos sem crises não é um problema, pois o interesse deste estudo é analisar o sinal EEG nas proximidades das crises.

Os dados utilizados são apresentados na tabela A.1, bem como dados da quantidade de crises por arquivo e as marcações de início e fim de crises.

Nome do arquivo	Número de crises	Início	Final
chb01_03	1	2996	3036
chb01_04	1	1467	1494
chb01_15	1	1732	1772
chb01_16	1	1015	1066
chb01_18	1	1720	1810
chb01_21	1	327	420
chb01_26	1	1862	1963
chb02_16+	1	2972	3053
chb02_19	1	3369	3378
chb03_01	1	362	414
chb03_02	1	731	796
chb03_03	1	432	501
chb03_04	1	2162	2214
chb03_34	1	1982	2029
chb03_35	1	2592	2656
chb03_36	1	1725	1778
chb04_05	1	7804	7853
chb04_08	1	6446	6557
chb04_28	2	1679	1781
		3782	3898
chb05_06	1	417	532
chb05_13	1	1086	1196
chb05_16	1	2317	2413
chb05_17	1	2451	2571
chb05_22	1	2348	2465

Nome do arquivo	Número de crises	Inicio	Final
chb06_01	3	1724	1738
		7461	7476
		13525	13540
chb06_04	2	327	347
		6211	6231
chb06_09	1	12500	12516
chb06_10	1	10833	10845
chb06_13	1	506	519
chb06_18	1	7799	7811
chb06_24	1	9387	9403
chb07_12	1	4920	5006
chb07_13	1	3285	3381
chb07_19	1	13688	13831
chb08_02	1	2670	2841
chb08_05	1	2856	3046
chb08_11	1	2988	3122
chb08_13	1	2417	2577
chb08_21	1	2083	2347

Nome do arquivo	Número de crises	Início	Final
chb09_06	1	12231	12295
chb09_08	2	2951	3030
		9196	9267
chb09_19	1	5299	5361
chb23_06	1	3962	4075
chb23_08	2	325	345
		5104	5151
chb23_09	4	2589	2660
		6885	6947
		8505	8532
		9580	9664

Tabela A.1: Arquivos do banco de dados MIT-CHB utilizados.