



Maria José Mateus Branco

Aplicação Computacional para o processamento e análise de sinais de O_2 recolhidos de modelos animais de epilepsia

Dissertação apresentada à Universidade de Coimbra para cumprimento dos requisitos necessários
à obtenção do grau de Mestre em Engenharia Biomédica

Julho 2018



UNIVERSIDADE DE COIMBRA



C •

FCTUC FACULDADE DE CIÊNCIAS
E TECNOLOGIA
UNIVERSIDADE DE COIMBRA

Aplicação Computacional para o processamento e análise de sinais de O₂ recolhidos de modelos animais de epilepsia

Maria José Mateus Branco

Dissertação apresentada à Universidade de Coimbra para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Engenharia Biomédica

Orientadores:

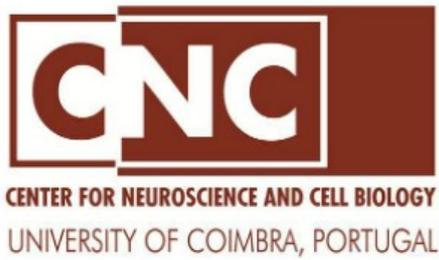
Professor Doutor César Teixeira

Professora Doutora Ana Ledo

Coimbra, Julho de 2018

Este trabalho foi realizado em colaboração com:

 **CISUC**



Esta cópia da tese é fornecida na condição de que, quem a consulta, reconhece que os direitos de autor são pertença do autor da tese e que nenhuma citação ou informação obtida a partir dela pode ser publicada, sem a referência apropriada.

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognize that is copyright rests with its author and that no quotation from the thesis and no information derived from it may be published without proper acknowledge

“The greatest enemy of knowledge is not ignorance, it’s the illusion of knowledge.”

Stephen Hawking

Agradecimentos

Gostaria de agradecer ao Professor Doutor César Teixeira pela orientação científica, apoio e disponibilidade durante a realização do projeto e à Professora Doutora Ana Ledo por toda a ajuda e experiência partilhada.

Quero agradecer a todos os meus amigos, mas de forma especial aos que me acompanharam ao longo dos últimos 5 anos e que estiveram presentes em todos os momentos importantes.

Quero agradecer à minha família, mas de forma muito especial aos meus pais por todo o apoio incondicional e por todos os sacrifícios que fizeram para me poderem proporcionar sempre o melhor. Sem eles este trabalho não seria possível.

A todos muito obrigada.

Resumo

A epilepsia é uma doença cerebral crónica não transmissível, que afeta aproximadamente 50 milhões de pessoas de todas as idades em todo o mundo. Estima-se que, por ano, são diagnosticadas cerca de 2,4 milhões de pessoas com epilepsia. Como forma de melhorar a vida destas pessoas é importante perceber as alterações ocorridas durante e principalmente antes das crises, para possibilitar a sua previsão no futuro e consequentemente a melhoria da qualidade de vida.

Esta tese descreve o desenvolvimento de uma aplicação computacional com a capacidade de analisar e processar sinais eletroquímicos recolhidos através de microelctrodos de platina em matriz microfabricados num suporte cerâmico e que possuem as propriedades eletroanalíticas adequadas para a medição de oxigénio *in vivo*, no espaço extracelular no cérebro, com elevada resolução espacial e temporal, depois de implantados no cérebro de ratos tratados com pilocarpina como modelo de epilepsia do lobo temporal.

No estudo é realizada a extração de características univariadas dos sinais recolhidos com base em estudos prévios realizados com eletroencefalogramas (EEG). A aplicação desenvolvida implementa uma cadeia de processamento composta por: pré-processamento; extração de características; redução/seleção de características; classificação e análise de resultados.

Foram estudados quatro ratos Wistar com 8 semanas de vida, mostrando que é possível identificar os momentos em que ocorrem crises através das características analisadas. A sensibilidade e especificidade médias em dados de teste de deteção foram de 66,41% e 50,04%, respetivamente. No entanto, para alguns animais obtiveram-se sensibilidades e especificidades acima de 80%.

Foi possível perceber que diferentes opções de processamento introduzem alterações no desempenho da classificação dos dados, apesar de não ser visível um padrão concreto, provavelmente devido ao reduzido número de animais analisados.

No melhor do nosso conhecimento parece não haver estudos de previsão/deteção de crises usando sinais relativos à pO_2 , o que torna os resultados apresentados nesta tese inovadores.

Palavras-chave: processamento de sinais eletroquímicos, previsão e detecção de crises, modelo animal de epilepsia induzido por pilocarpina, epilepsia do lobo temporal, *arrays* de microelétrodos de platina em suporte cerâmico.

Abstract

Epilepsy is a non-communicable brain disease that affects approximately 50 million people of all ages worldwide. It is estimated that about 2.4 million people are diagnosed with epilepsy per year. As a way to improve their lives, it is important to understand the changes occurred during and especially before crises to enable their correct prediction.

The purpose of this study was the development of a computational application able to analyze and process electrochemical signals recorded *in vivo* using platinum microelectrodes arrays microfabricated in ceramic support chronically implanted into the brain of rats treated with pilocarpine as temporal lobe epilepsy model and displaying suitable electroanalytical properties for the measurement of oxygen with high resolution in the brain extracellular space.

In this study, extraction of univariate features of the collected signal is executed based on the analysis of electroencephalograms (EEG). The developed application implements a processing chain composed of: pre-processing; extraction of characteristics; reduction / selection of characteristics; classification and analysis of results.

Four Wistar rats were studied at eight weeks of age, showing that it is possible to identify the moments in which seizures occur through analyzed features. The mean sensitivity and specificity in detection test data were 66.41% and 50.4%, respectively. However, for some animals a sensitivity and specificity above 80% have been detected.

Through the application, it was possible to observe that different pre-processing options introduce changes on data classification performance, although a concrete pattern is not visible, probably due the small number of analyzed animals.

To the best of our knowledge it appears to be no seizure prediction/detection studies using pO₂ signals, which makes the results presented in this thesis innovative.

Keywords: electrochemical signals processing, seizures prediction and detection, pilocarpine induced epilepsy animal model, temporal lobe epilepsy, platinum microelectrode arrays in ceramic support

Abreviaturas

ECG – Eletrocardiograma

EEG – Eletroencefalograma

ELT – Epilepsia do Lobo Temporal

FDR/h – *False Detection Rate*

FIR – *Finite Impulse Response*

FN – Falso Negativo

FP – Falso Positivo

FPR/h – *False Prediction Rate*

ICA – *Independent Component Analysis*

IIR – *Infinite Impulse Response*

SVM – *Support Vector Machine*

VN – Verdadeiro Negativo

VP – Verdadeiro Positivo

Glossário

Array – arranjo

False Detection Rate – Taxa de falsas deteções por hora

False Prediction Rate – Taxa de falsas previsões por hora

Finite Impulse Response – Resposta finita ao impulso

In vivo – Que ocorre dentro de um organismo

Independent Component Analysis - Análise de Componentes Independentes

Infinite Impulse Response – Resposta finita ao impulso

Kurtosis - Curtose

Machine Learning – Aprendizagem máquina

pO₂ – Pressão parcial de O₂

Ranking – Ordenação

Skewness – Obliquidade ou assimetria

Spike-and-Wave – Padrão tipicamente observado no EEG durante crises epiléticas

Support Vector Machine – Máquinas de vetores suporte

Valor p – Probabilidade de ser obtida uma estatística de teste igual ou mais extrema do que a observada numa amostra sob a hipótese nula

Wavelet - Onduleta

Lista de Figuras

Figura 2.1: Imagem da ponta cerâmica do array de microelétrodos de platina S2, onde é possível observar a posição dos sítios de platina, linhas condutoras e revestimento isolador em poli-imida.	9
Figura 2.2: Representação simbólica da matriz de confusão.	10
Figura 4.1: Resposta em frequência do filtro Butterworth passa-baixo de ordem 10. ...	25
Figura 4.2: Resposta em frequência dos filtros Chebyshev passa-baixo tipo 1 e tipo 2, respectivamente, de ordem 10.....	26
Figura 4.3: Resposta em frequência do filtro Elíptico passa-baixo de ordem 10.....	26
Figura 4.4: Esquema representativo do processo de decomposição implementado pela DWT para sinais com frequência de amostragem igual a 200Hz.	30
Figura 4.5: Exemplo de mapa tempo-frequência com a representação esquemática da divisão das bandas de frequência pré-definidas e a série temporal associada.	32
Figura 5.1: Interface de importação de dados da aplicação com o canal 1 e 2 do conjunto de dados WO2_005 representados e o instante correspondente à injeção de pilocarpina assinalado.....	43
Figura 5.2: Interface da aplicação correspondente ao painel de pré-processamento, com a representação dos sinais correspondentes ao canal 1 e 2 do conjunto de dados WO2_055 após a aplicação de um filtro Butterworth passa-banda de ordem 10, com banda de passagem de 1 a 49,5Hz, e o instante em que ocorreu a injeção de pilocarpina assinalado.	44
Figura 5.3: Interface da aplicação correspondente ao painel de extração de características, com a representação do kurtosis ao longo do tempo e os instantes da injeção de pilocarpina e da ocorrência de automatismos faciais assinalados.	45
Figura 5.4: Interface da aplicação correspondente ao painel de análise de várias características em simultâneo, com a representação da potência relativa da banda gama correspondentes ao canal 2 em função da potência acumulada dos canais 1 e 2 e do plano de separação dos dois grupos (seizure e other), definidos através da opção de definição do target, obtida através do treino de um classificador linear.	46

Figura 6.1: Representação da acurácia em função do número de características utilizadas na separação linear através do método Fisher LDA, no conjunto de dados WO2_005, sem pré-processamento.....	48
Figura B.1	85
Figura B.2	86
Figura B.3	87
Figura B.4	88
Figura B.5	89
Figura B.6	90
Figura B.7	91
Figura B.8	91
Figura B.9	92
Figura B.10.....	92
Figura B.11.....	93
Figura B.12.....	94
Figura B.13.....	95

Lista de Tabelas

Tabela 4.1: Descrição dos conjuntos de dados utilizados no estudo	24
Tabela 6.1: Medidas médias de desempenho da separação de cada um dos conjuntos de dados.	51
Tabela 6.2: Desempenho médio obtido na classificação baseada na distância euclidiana mínima de cada conjunto de dados após a aplicação do algoritmo fastICA na fase de pré-processamento.	52
Tabela 6.3: Descrição dos conjuntos usados para treinar e testar classificadores lineares Fisher LDA.	53
Tabela 6.4: Medidas médias de desempenho obtidas nas fases de treino e teste de classificadores lineares Fisher LDA, para cada um dos conjuntos de dados utilizando diferentes opções de pré-processamento.	55
Tabela 6.5: Desempenho médio obtido no treino e teste de classificadores lineares Fisher LDA utilizando os conjuntos 1, 2, 3 e 4	56
Tabela A.1: Desempenho obtido na classificação linear dos dados WO2_005 utilizando 2, 3, 69 e 76 características extraídas sem pré-processamento do sinal original.	71
Tabela A.2: Desempenho obtido na classificação linear dos dados WO2_007 utilizando 2, 3 e 76 características extraídas sem pré-processamento do sinal original.	71
Tabela A.3: Desempenho obtido na classificação linear dos dados WO2_017 utilizando 2, 3, 67 e 76 características extraídas sem pré-processamento do sinal original.	71
Tabela A.4: Desempenho obtido na classificação linear dos dados WO2_020 utilizando 2, 3, 66 e 76 características extraídas sem pré-processamento do sinal original.	72
Tabela A.5: Desempenho obtido na classificação linear dos dados WO2_005 filtrados, utilizando 2, 3, 67 e 76 características extraídas.	72
Tabela A.6: Desempenho obtido na classificação linear dos dados WO2_005 filtrados, utilizando 2, 3, 73 e 76 características extraídas, após a aplicação do algoritmo fastICA na fase de pré-processamento.	72

Tabela A.7: Desempenho obtido na classificação linear dos dados WO2_007 filtrados utilizando 2, 3, 71 e 76 características extraídas.	73
Tabela A.8: Desempenho obtido na classificação linear dos dados WO2_007 filtrados utilizando 2, 3, 74 e 76 características extraídas, após a aplicação do algoritmo fastICA na fase de pré-processamento.	73
Tabela A.9: Desempenho obtido na classificação linear dos dados WO2_017 filtrados utilizando 2, 3, 74 e 76 características extraídas.	73
Tabela A.10: Desempenho obtido na classificação linear dos dados WO2_017 filtrados utilizando 2, 3, 70 e 76 características extraídas, após a aplicação do algoritmo fastICA na fase de pré-processamento.	74
Tabela A.11: Desempenho obtido na classificação linear dos dados WO2_020 filtrados utilizando 2, 3, 65 e 76 características extraídas.	74
Tabela A.12: Desempenho obtido na classificação linear dos dados WO2_020 filtrados utilizando 2, 3, 75 e 76 características extraídas, após a aplicação do algoritmo fastICA na fase de pré-processamento.	74
Tabela A.13: Resultados obtidos na classificação através do método Fisher LDA utilizando as 2 características com melhor capacidade de discriminação para cada um dos casos.	75
Tabela A.14: Resultados obtidos na classificação através do método Fisher LDA utilizando as 3 características com melhor capacidade de discriminação para cada um dos casos.	76
Tabela A.15: Resultados obtidos na classificação através do método Fisher LDA utilizando todas as características cada um dos casos.	77
Tabela A.16: Resultados obtidos na classificação através do método Fisher LDA utilizando as N características com melhor capacidade de discriminação, que permitem obter o desempenho ótimo para cada um dos casos.	78
Tabela A.17: Desempenho obtido na classificação baseada na distância euclidiana mínima dos dados WO2_005 utilizando 2, 3 e 76 características extraídas, após a aplicação do algoritmo fastICA na fase de pré-processamento.....	79
Tabela A.18: Desempenho obtido na classificação baseada na distância euclidiana mínima dos dados WO2_007 utilizando 2, 3, 37 e 76 características extraídas, após a aplicação do algoritmo fastICA na fase de pré-processamento.....	79

Tabela A.19: Desempenho obtido na classificação baseada na distância euclidiana mínima dos dados WO2_017 utilizando 1, 2, 3 e 76 características extraídas, após a aplicação do algoritmo fastICA na fase de pré-processamento.....	79
Tabela A.20: Desempenho obtido na classificação baseada na distância euclidiana mínima dos dados WO2_020 utilizando 2, 3, 18 e 76 características extraídas, após a aplicação do algoritmo fastICA na fase de pré-processamento.....	80
Tabela A.21: Desempenho obtido no treino e teste dos diferentes conjuntos de dados sem pré-processamento.	81
Tabela A.22: Desempenho obtido no treino e teste dos diferentes conjuntos de dados após a aplicação do filtro Butterworth.....	82
Tabela A.23: Desempenho obtido no treino e teste dos diferentes conjuntos de dados após a aplicação do filtro Butterworth e do algoritmo fastICA.	83

Índice

Agradecimentos.....	vii
Resumo	ix
Abstract.....	xi
Abreviaturas.....	xiii
Glossário.....	xiv
Lista de Figuras.....	xv
Lista de Tabelas	xvii
1. Introdução	1
1.1. Contextualização.....	1
1.2. Motivação.....	2
1.3. Objetivos	3
2. Conceitos base	5
2.1. A epilepsia.....	5
2.1.1. Epileptogénese	5
2.1.2. Epilepsia do Lobo Temporal.....	6
2.1.3. Modelo de Epilepsia do Lobo Temporal Induzido por Pilocarpina 7	
2.1.4. Crises Epiléticas	7
2.2. <i>Array</i> de microelétrodos de platina.....	8
2.3. Detecção e Previsão de crises epiléticas	9
3. Estado da arte.....	13
4. Materiais e Métodos.....	23
4.1. Dados.....	23
4.2. Filtros	24
4.3. Análise de Componentes Independentes (ICA).....	26
4.4. Extração de características	27

4.4.1. Medidas estatísticas	27
4.4.2. Coeficientes <i>wavelet</i>	29
4.4.3. Mapa Tempo-Frequência	31
4.4.4. Potência Relativa	31
4.4.5. Parâmetros de Hjorth	32
4.4.6. Potência acumulada	34
4.4.7. Modelo Quantitativo	34
4.5. Normalização dos dados	35
4.6. Teste U de Mann-Whitney e Teste de Kruskal-Wallis	36
4.7. Classificadores	36
4.7.1. Classificador <i>Fisher LDA</i>	36
4.7.2. Classificador baseado na distância euclidiana mínima	38
5. Aplicação computacional para processamento de sinais de O₂	41
6. Resultados e discussão	47
7. Conclusão e Trabalho Futuro	59
Bibliografia	61
A. Medidas de desempenho obtidas nos diferentes testes	71
B. Manual do Utilizador	85
B.1. Iniciar a aplicação	85
B.2. Painel 1 - <i>Load</i>	85
B.3. Painel 2 – <i>PreProcessing</i>	86
B.4. Painel 3 – <i>Processing</i>	90
B.5. Painel 4 – <i>Analysis</i>	92

1. Introdução

Este projeto foi desenvolvido no âmbito da unidade curricular Projeto do Mestrado Integrado de Engenharia Biomédica da Faculdade de Ciências e Tecnologias da Universidade de Coimbra, no ano letivo de 2017/2018. O documento descreve o processo de elaboração de uma aplicação computacional que permite analisar sinais originados pela monitorização eletroquímica da pressão parcial de O_2 (pO_2) através da utilização de microelétrodos de platina em matriz – *arrays* - implantados no cérebro de ratos tratados com pilocarpina como modelo crónico de Epilepsia do Lobo Temporal (ELT).

O mesmo documento é submetido a avaliação para a obtenção do grau de Mestre em Engenharia Biomédica.

1.1. Contextualização

A epilepsia é uma doença cerebral crónica não transmissível, que afeta aproximadamente 50 milhões de pessoas de todas as idades em todo o mundo. Estima-se que, por ano, são diagnosticadas cerca de 2,4 milhões de pessoas com epilepsia [1].

Entre 30% e 40% dos pacientes com epilepsia não podem ser tratados com medicamentos antiepiléticos ou cirurgicamente [2] e a incidência de epilepsia refratária é muito relevante (40%) em pacientes com ELT [3]. Para este grupo de doentes, a previsão e deteção de crises epiléticas é muito importante para melhorar de forma significativa a sua qualidade de vida

O trauma inicial associado com ELT induz alterações do foro bioquímico, molecular, fisiológico ou estrutural que contribuem para o desenvolvimento de epileptogénese e desencadeiam crises espontâneas e recorrentes [4]. As consequências socioeconómicas resultantes da perda de produtividade e diminuição da função cognitiva nos doentes com epilepsia são significativas [5].

O eletroencefalograma (EEG) tem sido a maior fonte para obter informação no que diz respeito à identificação de alterações da atividade cerebral resultante de um episódio epilético [6], contudo os algoritmos que se baseiam apenas na análise deste não podem ser usados em aplicações clínicas, pois ainda apresentam um desempenho fraco no que

respeita à detecção e previsão de crises. Desta forma, torna-se necessário recorrer a outras fontes de informação para melhorar estes métodos e possibilitar a previsão e detecção de crises em ambiente real.

Sabe-se que o metabolismo energético cerebral se encontra comprometido em pacientes com epilepsia [7], [8]. Estudos de imagiologia no período livre de crises revelaram a ocorrência de hipoperfusão, redução do aporte de glucose, bem como um déficit energético. Contudo, ainda não é bem compreendida a dinâmica temporal e espacial do acoplamento neurovascular e neurometabólico que ocorre durante a evolução para uma crise: a alteração extrema de excitabilidade neuronal altera a necessidade energética, traduzida numa diminuição transiente da concentração local de O_2 [9]. Estudos demonstram que na sequência de uma atividade neuronal, o tecido se torna ligeiramente isquémico até que a resposta vascular reponha o aporte de nutrientes adequados à necessidade energética.

O desenvolvimento deste projeto surgiu da necessidade, de um grupo de investigadores do Centro de Neurociências e Biologia Celular de Coimbra (CNC), em ter uma ferramenta para processar e analisar os sinais de pO_2 recolhidos nos seus estudos com modelos animais de epilepsia.

1.2. Motivação

Uma vez que a epilepsia é uma doença que afeta um grande número de pessoas em todo o mundo e que alguns casos não podem ser tratados recorrendo a medicamentos antiepiléticos ou cirurgia, existe uma necessidade de melhorar a qualidade de vida das pessoas que sofrem diariamente com a natureza imprevisível e a recorrência de crises epiléticas. Ao longo dos anos foram realizados vários estudos baseados na medição da atividade elétrica do sistema nervoso, na sua maioria com base na análise de EEG, mas nunca conjugados com a monitorização de eventos químicos quer associados à neurotransmissão, quer associados ao metabolismo.

Surge então a necessidade de analisar outros bio-sinais, ambicionando que fundidos com informação do EEG levem a uma melhoria da detecção/previsão das crises e consequentemente a uma melhoria da qualidade de vida dos pacientes.

1.3. Objetivos

Este estudo tem como objetivo a criação de uma aplicação capaz de analisar e processar sinais obtidos através de conjuntos de microelétrodos de platina, com características eletroquímicas básicas, usados para realizar medições de oxigénio *in vivo*, no espaço extracelular no cérebro, com alta resolução, implantados no cérebro de ratos.

Este projeto tem também como foco explorar a relação entre a pressão de oxigénio no cérebro e a ocorrência das crises epiléticas e, analisar a sua capacidade de discriminação dos diferentes estados epiléticos, através da aplicação de diferentes filtros aos sinais iniciais, da análise de componentes independentes (ICA), da extração e análise de características dos sinais e classificação linear das mesmas.

2. Conceitos base

2.1. A epilepsia

A epilepsia é uma doença crónica neurológica que afeta pessoas em todo mundo e que é caracterizada pela ocorrência de crises epiléticas espontâneas e recorrentes. Estas são episódios com intensidade e duração variável e resultam de descargas elétricas sucessivas numa população de neurónios. Dependendo da localização do foco da crise, dos padrões de propagação, da medicação, do estado de maturação do cérebro, da fase do ciclo de sono, do estado emocional podem surgir vários sintomas, durante o período em que ocorre a crise, como mudanças de humor, perda de memória, distúrbios sensoriais, problemas na execução de movimentos, convulsões e até perda de consciência [10].

Há 40 a 50 casos de epilepsia em 100.000 pessoas por ano nos países desenvolvidos e entre 122 a 190 casos em 100.000 pessoas nos países em desenvolvimento, devido aos baixos níveis socioeconómicos que favorecem a ocorrência de fatores que desencadeiam a epilepsia [11].

A epilepsia é uma doença que pode adquirir diferentes classificações, de acordo com a identificação de vários sintomas e causas, entre as quais se encontram as epilepsias refratárias, que são aquelas que não apresentam resposta a tratamentos com medicamentos antiepiléticos e que podem ou não ser controladas através de intervenções cirúrgicas [12].

2.1.1. Epileptogénese

O processo pelo qual um circuito neuronal transita de um estado normal para um estado de hiperexcitabilidade, com capacidade de gerar crises epiléticas espontaneamente denomina-se epileptogénese. Vários estudos realizados mostraram que este processo é dinâmico e envolve uma cascata de alterações moleculares, celulares e de redes neuronais [13].

No córtex cerebral há cerca de 10^{10} neurónios e as sinapses são as interfaces eletroquímicas que permitem que os neurónios comuniquem entre si. Numa sinapse estão

envolvidos um neurónio emissor ou pré-sináptico e um neurónio recetor ou pós-sináptico que pode responder de forma inibitória ou excitatória. O que define as sinapses são os neurotransmissores e respetivos recetores membranares envolvidos. O resultado da ligação de cada tipo de neurotransmissor a um tipo de recetor membranar poderá resultar na polarização do neurónio pós-sináptico – efeito inibitório – ou na despolarização do neurónio recetor – efeito excitatório, sendo que a alteração do potencial de membrana resulta do fluxo de iões como sódio (Na^+), potássio (K^+) e cálcio (Ca^{2+}) através de canais iónicos distribuídos na membrana plasmática em resultado da ligação de ligandos e/ou em função do potencial de membrana [14].

Na ocorrência de crises epiléticas, a atividade elétrica de um grupo de neurónios é anormal (excessiva) e síncrona (envolve vários neurónios ao mesmo tempo) e há uma desregulação entre a inibição e excitação dos neurónios. Esta despolarização anormal pode envolver neurónios localizados apenas de um lado do cérebro (crises epiléticas focais) ou nos dois lados (crises epiléticas generalizadas), podendo ainda dar-se uma despolarização com início em apenas um lado e a sua propagação para neurónios no outro lado (crises epiléticas focais com generalização secundária).

2.1.2. Epilepsia do Lobo Temporal

A epilepsia do lobo temporal é a forma mais comum de epilepsia focal ou relacionada à localização, assinalada por convulsões recorrentes com origem no sistema límbico, particularmente do hipocampo, da amígdala e do córtex entorrinal. Este tipo de epilepsia, associada a esclerose do hipocampo, tem uma grande relevância clínica, pois apresenta alta incidência e gravidade, para além de ser a forma mais comum adquirida na fase adulta e de ser caracterizada pela ocorrência de convulsões imprevisíveis e recorrentes, e pela resistência a fármacos antiepiléticos. A ELT surge, normalmente, precedida de uma convulsão febril, de um traumatismo craniano ou, ocasionalmente, de um estado epilético seguido de um período latente, no qual a atrofia unilateral do hipocampo desencadeia convulsões espontâneas [15].

2.1.3. Modelo de Epilepsia do Lobo Temporal Induzido por Pilocarpina

Ao longo dos anos foi possível reproduzir doenças humanas em modelos animais, o que produziu um grande avanço na medicina.

A ELT apresenta quatro características principais, que são a localização dos focos da crise no sistema límbico, particularmente no hipocampo, no córtex entorrinal e na amígdala; a conclusão frequente da “lesão precipitante inicial” que precede o aparecimento da ELT; o intervalo sem ocorrência de crises conhecido como período latente, seguindo a crise espontânea, e uma alta incidência de esclerose mesial ou no *Cornu Ammonis* [4]. A maioria destas características pode ser reproduzida em modelos animais crônicos de ELT, nomeadamente o modelo de pilocarpina.

A pilocarpina é um agonista colinérgico muscarínico, que quando é administrado em altas doses em roedores provoca um quadro de alterações comportamentais com manifestação de crises motoras límbicas [16]. O efeito epileptogénico deste agente colinérgico resulta de descargas originadas em neurónios do hipocampo devido ao bloqueio das correntes de potássio transmembranar. As características principais deste modelo são a indução do estado epilético agudo mais rápido do que com outro fármaco usado para reproduzir ELT em animais; a presença de um período latente seguido da ocorrência de convulsões espontâneas recorrentes; a ocorrência de lesões generalizadas, algumas delas localizadas nas mesmas áreas cerebrais de doentes com ELT e associadas a reorganização da rede neuronal em regiões do hipocampo e do giro para-hipocampal e o facto das convulsões não serem controladas de forma eficiente em pacientes e roedores epiléticos tratados com pilocarpina. Deste modo, é um modelo com elevado grau de isomorfismo com a ELT em humanos [17].

2.1.4. Crises Epiléticas

Cada crise epilética causa um grande aumento da demanda metabólica. Estudos recentes elucidaram uma relação espaciotemporal entre atividade neuronal, volume sanguíneo cerebral e oximetria *in vivo*. Após os períodos livres de crises e durante as crises ocorre uma diminuição focal na oxigenação tecidual e um aumento focal na hemoglobina desoxigenada que pode persistir durante o período da crise, sugerindo que há uma incapacidade de responder à necessidade metabólica. Uma crise é também

acompanhada por um aumento focal rápido do fluxo sanguíneo cerebral e no volume sanguíneo cerebral [18].

Com a realização de vários estudos do cérebro e da sua atividade, foram definidos quatro estados cerebrais relacionados com a epilepsia: inter-ictal, pré-ictal, ictal e pós-ictal. O estado inter-ictal corresponde ao intervalo de tempo entre crises ou convulsões e é caracterizado pela ausência destas e pela presença de padrões *spike-and-wave*.

O estado pré-ictal corresponde ao estado que ocorre imediatamente antes da ocorrência clínica da crise, seguidamente ao inter-ictal.

O estado ictal corresponde ao período em que se verifica a ocorrência de crises.

O estado pós-ictal corresponde ao intervalo de tempo delimitado pelo fim de um episódio epilético e pela restituição da *baseline* que diz respeito ao período inter-ictal [19].

2.2. Array de microelétrodos de platina

Neste estudo, são usados *arrays* de microelétrodos de platina em suporte cerâmico que apresentam excelentes propriedades eletroanalíticas para a medição de oxigénio *in vivo* no espaço extracelular no cérebro com elevada resolução espacial e temporal. Estes microelétrodos foram desenvolvidos no CenMet, University of Kentucky, Lexington, KY, U.S.A. e são fabricados por técnicas fotolitográficas, permitindo uma elevada reprodutibilidade nas características morfológicas e eletroanalíticas. No presente trabalho utilizaram-se arranjos miniatura da tipologia S2 que contêm 2 pares de sítios ativos de platina ($15 \times 333 \mu\text{m}^2$) dispostos em linha (Figura 2.1).

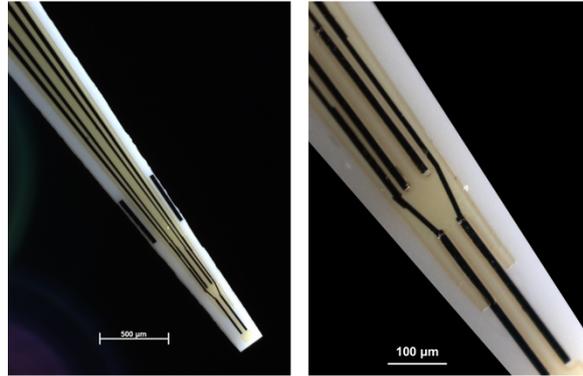


Figura 2.1: Imagem da ponta cerâmica do *array* de microelétrodos de platina S2, onde é possível observar a posição dos sítios de platina, linhas condutoras e revestimento isolador em poli-imida.

A platina enquanto superfície eletroativa de elétrodo, apresenta uma condutividade elétrica alta que é ideal para estimulação e registo na interface neuronal. Por outro lado, é reconhecidamente o metal mais ativo na redução eletrocatalítica do O_2 e é altamente biocompatível e inerte, permitindo uma implantação de longo prazo. A utilização de um suporte cerâmico permite aumentar a biocompatibilidade da matriz, para além de oferecer resistência mecânica e isolamento elétrico.

Estes *arrays* de microelétrodos, além de permitirem medir alterações tónicas e de fase rápidas dos níveis de glutamato, lactato e glucose, podem ser configurados para uma análise múltipla, como foi descrito para a colina e acetilcolina [20], [21]. A sua capacidade para medir pO_2 cerebral no espaço extracelular em ratos anestesiados foi demonstrada [22].

2.3. Detecção e Previsão de crises epiléticas

A deteção de crises epiléticas refere-se à deteção do estado ictal. Por outro lado, a previsão de crises epiléticas refere-se à identificação do estado pré-ictal, isto é o estado imediatamente anterior à ocorrência clínica de um episódio epilético. Através da deteção da fase pré-ictal é possível antecipar a ocorrência de uma crise dentro de um determinado intervalo de tempo. Apesar deste ser o período de maior importância no que respeita ao bem-estar dos pacientes, também levanta algumas dificuldades, uma vez que a sua definição não é concreta, isto é, não é definido clinicamente.

No que diz respeito ao desempenho dos sistemas de previsão e deteção de crises, este é avaliado recorrendo a diferentes métricas de acordo como se aborda o problema. Se o problema for abordado amostra a amostra, onde uma amostra corresponde a um instante temporal ou a uma janela temporal deslizante, com ou sem sobreposição, usa-se como medidas a sensibilidade, a especificidade, e a acurácia.

A sensibilidade (equação 2.1) corresponde à razão entre os instantes classificados como pertencente ao período ictal (deteção) ou pré-ictal (previsão) e o total de instantes que na realidade foram classificados como tal.

A especificidade (equação 2.2) corresponde à razão entre os instantes classificados como ausência de crise (deteção) ou identificados como não pré-ictal (previsão) e o total de instantes em que na realidade não ocorre qualquer evento.

A precisão consiste na razão entre os verdadeiros positivos e a soma destes com os falsos positivos.

A acurácia (equação 2.4) consiste na razão entre as amostras classificadas corretamente e o total de amostras.

Por outro lado, se o sistema de previsão/deteção de crises for avaliado do ponto de vista dos eventos discretos, isto é, do ponto das crises, normalmente avalia-se em termos de sensibilidade e a taxa de falsas previsões/deteções por hora (FPR/h ou FDR/h). A sensibilidade neste sentido é definida como a razão entre as crises previstas/detetadas corretamente e a totalidade de crises ocorridas.

		Valor Previsto	
		Positivo	Negativo
Valor Verdadeiro	Positivo	VP	FN
	Negativo	FP	VN

Figura 2.2: Representação simbólica da matriz de confusão.

$$\text{Sensibilidade} = \frac{VP}{VP+FN} \tag{2.1}$$

$$\text{Especificidade} = \frac{VN}{VN+FP} \quad (2.2)$$

$$\text{Precisão} = \frac{VP}{VP+FP} \quad (2.3)$$

$$\text{Acurácia} = \frac{VN+VP}{VN+FN+VP+FP} \quad (2.4)$$

$$\text{FDR/h} = \text{FPR/h} = \frac{\text{Número de falsos alarmes}}{\text{Duração período interictal}} \quad (2.5)$$

3. Estado da arte

A história da previsão de crises epiléticas começou com Viglione e Walsh em 1975 [23], que tentaram extrair precursores das crises a partir de EEG, usando aproximações lineares. Mais tarde, foi reportada a alteração de parâmetros dos modelos, num período de 6 segundos antes da ocorrência da crise (período considerado pré-ictal), utilizando modelos autorregressivos [24], [25]. Em 1982, foram identificadas mudanças em características entre épocas que precediam as crises em 1 minuto e épocas de controle para cada paciente [26].

Outro grupo estudou o valor preditivo da taxa de ocorrência de picos em EEG, enquanto que tinha sido reportado, num estudo recente, uma diminuição da taxa focal de picos com um aumento da taxa de picos bilaterais antes das crises [25]. Contudo, estudos em bases de dados mais extensas não mostraram alterações sistemáticas na taxa de picos antes dos episódios clínicos [27]-[29].

Com a teoria da não linearidade dos sistemas, foram introduzidas novas abordagens com o objetivo de melhorar a caracterização de sistemas dinâmicos com comportamentos complexos em relação ao que era possível através das abordagens lineares convencionais. Iasemidis e seus colaboradores reportaram uma diminuição do comportamento caótico nos sinais de EEG intracraniano de pacientes com epilepsia, nos minutos que antecedem o início das crises [30].

Um grupo de investigadores na França reportou, num grupo de pacientes maior, uma diminuição pré-ictal na complexidade espaciotemporal, através da densidade de correlação antes das crises. Este mesmo grupo desenvolveu ainda outra medida, denominada índice de semelhança dinâmica, para quantificar mudanças dinâmicas em relação a uma janela de referência constante no início de uma gravação pré-ictal, tendo assim identificado uma diminuição dinâmica antes das crises semelhante em sinais de EEG intracranianos [31] e de escalpe [32].

Durante muitos anos, pensou-se que as crises epiléticas começavam apenas alguns segundos antes dos ataques clínicos, mas através do estudo de EEG intracranianos de longa duração de diferentes doentes avaliados para cirurgias de epilepsia, foi possível observar evidências que sugerem que as crises se desenvolvem vários minutos ou até

horas antes do início clínico. Alguns autores reportam a alteração do fluxo de sangue no cérebro, a alteração da quantidade de oxigênio no sangue e da frequência cardíaca antes dos ataques. Estas evidências sugerem a existência de um estado que precede as crises designado pré-ictal, o que por sua vez infere a possibilidade de previsão dos episódios clínicos [33]–[39].

Na sequência destes trabalhos onde se procurou perceber que características poderiam evidenciar a existência do estado pré-ictal, surgiram os primeiros estudos com o objetivo de prever crises [40], [41]. Estes estudos basearam-se em medidas de dimensão de correlação, obtendo uma taxa de ocorrência de falsos positivos nula e uma sensibilidade de 94%, num teste realizado com um total de 16 pacientes, com um registo de EEG intracraniano de 21 horas e ao qual não realizam nenhum tipo de validação estatística do desempenho.

A capacidade de prever crises através de medidas de dimensão de correlação foi posteriormente questionada [42], [43], tendo-se demonstrado a incapacidade de prever crises usando este tipo de característica, uma vez que os valores de sensibilidade obtidos foram muito baixos.

Nos primeiros anos do século XXI, foram analisados outros tipos de métodos de previsão de crises, entre os quais se destacaram, devido ao seu desempenho positivo, os que tinham em conta medidas como a sincronização de fase de diferentes zonas do córtex [44]–[46], a densidade de Lerner [47], o arrastamento dinâmico (medida baseada em expoentes de Lyapunov) [48], a energia cumulativa do sinal [49], [50] e simulação de modelos neuronais [51]. A baixa eficácia destes métodos foi, no entanto, evidenciada por outros, levando a concluir-se que, na fase de treino, quando os parâmetros usados na previsão de crises são ajustados como forma de otimizar o desempenho do algoritmo para um determinado conjunto de dados, não reproduzem os mesmos resultados quando aplicados a um conjunto de teste diferente. Esta observação torna necessário proceder à aplicação do método num conjunto de teste diferente da amostra inicial, como forma de conhecer o seu verdadeiro desempenho [52].

Com o aumento da capacidade de armazenamento de dados, os centros de epilepsia tornaram-se capazes de adquirir dados da monitorização pré-cirúrgica completa, sem a necessidade de selecionar amostras [53]. Em 2005, vários grupos publicaram séries de estudos realizados num conjunto de cinco gravações contínuas de vários dias por

diferentes centros de epilepsia para o “First International Collaborative Workshop on Seizure Prediction” [54], com o objetivo de ter diferentes grupos de teste e de comparar o desempenho dos diferentes métodos, aplicando-os ao conjunto dos dados.

Em 2005 e no seguimento do evento anterior, estudos levados a cabo por diferentes grupos revelaram, em grande parte, um desempenho fraco das medidas univariadas [55]–[59], sendo assim reportado um melhor desempenho por medidas bivariadas e multivariadas [59]–[61].

Com a melhoria dos valores de sensibilidade e especificidade, tornou-se necessário mostrar que os algoritmos de previsão de crises, quando aplicados a dados fora do conjunto de treino, apresentam um desempenho melhor do que um processo de previsão aleatório. Assim, chegou-se à ideia de que antes de se questionar se um método apresenta um desempenho suficiente para ser aplicado em ambiente clínico, deve ser analisado se este é mais eficaz do que o acaso, o que levou à inevitabilidade da aplicação de métodos de validação estatística. Estes métodos baseiam-se em simulações de Monte Carlo [59] ou na comparação com resultados analíticos derivados de esquemas de previsão aleatórios ou periódicos [62].

A abundância de ferramentas matemáticas, estatísticas e de processamento de sinal, torna o conjunto de possíveis métodos de previsão de crises muito amplo, o que deu origem a diferentes estudos ao longo do tempo.

Sackellares e os seus colaboradores realizaram um estudo com EEG de longa duração de um conjunto de 10 doentes que continha 130 crises e 2100 horas de gravação. Neste estudo, para um espaço de previsão de 30 minutos, foi possível obter um algoritmo que apresentou uma sensibilidade de 80% e 0,566 FPR/h [63].

Mais tarde, foi feito um estudo baseado em 6 características lineares e não lineares bivariadas que reportou resultados ótimos, sugerindo que a combinação de características lineares e não lineares derivadas do domínio temporal e espectral poderiam conduzir a resultados de alta previsão, mesmo para um horizonte de previsão grande (2 horas). Através da utilização de redes de convolução combinadas com coerência de ondulas, foi possível prever todas as crises fora da amostra de treino, em 15 dos 21 pacientes, sem ocorrência de alarmes falsos, apresentando uma sensibilidade de 71% e 0 falsos positivos [64].

Um estudo baseado na coerência da fase média bivariada, reportou desempenhos razoáveis para a maioria dos pacientes, sendo que, para cada paciente, a melhor sensibilidade variou entre 50% e 88% e as taxas de ocorrência de previsões falsas entre 0,64/h e 4,69/h, o que se tornou num obstáculo para a utilização deste método em aplicações clínicas, uma vez que ultrapassam a taxa máxima proposta (0,15/h) [65].

Em 2011, foi apresentado o EPILAB, um pacote de *software* para auxiliar os investigadores no estudo de previsão de crises epiléticas, com uma interface gráfica intuitiva e conveniente. Esta ferramenta é capaz de processar dados de EEG e ECG (eletrocardiograma) armazenados em diferentes formatos, de extrair mais de 35 medidas no domínio de tempo e frequência com base na análise de dados univariada e multivariada, que podem ser pós-processados e usados para prever crises epiléticas. Na previsão de crises podem ser usados limiares otimizados ou métodos de classificação, como redes neuronais artificiais, redes neuronais celulares e *Support Vector Machines* (SVM). Para além da eficiência na previsão, o EPILAB demonstrou ser uma forma de comunicar, avaliar e comparar resultados e dados entre a comunidade de previsão de crises [66].

Um método baseado em características espectrais de EEG intracraniano e *machine learning* foi treinado e testado em dados de 18 pacientes, alcançando uma sensibilidade média de 98,3% e uma taxa média de ocorrência de falsas previsões de 0,29/h [67]. Com a utilização da montagem bipolar, foi possível obter um desempenho um pouco melhor e com a montagem *time-differential*, uma pequena diminuição da taxa de falsos alarmes, enquanto que a sensibilidade se manteve quase inalterada em relação às obtidas com as montagens anteriores. Apesar da diminuição da taxa de falsos positivos, esta continua acima da máxima proposta.

O desenvolvimento de um algoritmo baseado na combinação de características não-lineares univariadas e bivariadas com poder preditivo revelado em estudos anteriores demonstrou uma sensibilidade média de 79,9% e 90,2% e uma taxa média de alarmes falsos de 0,17/h e 0,11/h, respetivamente, e uma especificidade média de 97%, para períodos de ocorrência de crises de 30 e 50 minutos [68].

Williamson e os seus colaboradores avaliaram um método que combina *machine learning* específica de cada paciente e características multivariadas num estudo onde foram avaliados 3 limiares de previsão diferentes, para os quais os valores da

sensibilidade média e da taxa de ocorrência de falsas previsões variaram entre 86% e 95% e entre 0,07/h e 0,03/h, respetivamente. Foi também analisado o impacto da variação de diferentes parâmetros, como a duração do período, o número de escalas de atraso, o número de componentes principais e variações da covariância, tendo sido reportado que os melhores resultados foram obtidos com a utilização dos parâmetros padrão [69].

A segurança e eficácia de um sistema de previsão de crises implantado em 15 pacientes adultos com epilepsia focal foram estudadas. Apenas os pacientes em que foi possível obter uma sensibilidade superior a 65% na primeira fase passaram para a fase de teste posterior, o que aconteceu com 11 dos doentes. O algoritmo desenvolvido baseou-se na energia média, na energia de Teager-Kaiser e *line-length*, que foram extraídos através da aplicação de uma janela de análise de 5 segundos aos sinais filtrados. No estudo, os valores de sensibilidade e da FPR/h não foram referidos, mas foi reportado que para 9 dos 11 pacientes, nos quais foi realizado o estudo, foi possível obter resultados estatisticamente significativos. Este estudo foi o primeiro a usar dados de EEG intracraniano ambulatorio e a demonstrar a viabilidade de prever crises em ambulatorio [70].

Um estudo baseado em medidas de semelhança entre o estado pré-ictal e inter-ictal derivadas na entropia e energia de *wavelets* reportou uma sensibilidade média superior a 85%, uma taxa de ocorrência de previsões falsas inferior a 0,1/h e um tempo médio de previsão de cerca de 36 minutos [71].

A taxa de pontas epiléticas, no EEG intracraniano, foi analisada como medida para prever crises, apresentando diferenças significativas para os diferentes estados epiléticos, quando calculada em períodos de 5 segundos selecionados aleatoriamente em cada segmento. Esta medida foi usada como característica na previsão de crises após o seu cálculo com uma janela deslizante de 5 segundos deslizante ao longo do sinal e da sua suavização através de um filtro de média móvel de 30 segundos. Durante o período de treino foi definido um limiar específico otimizado para cada paciente, que quando ultrapassado pela taxa de pontas epiléticas suavizada, gerava um alarme. O algoritmo, quando testado com um tempo de intervenção de 10 segundos, atingiu uma sensibilidade de 56% e de 72,7% e uma taxa de alarmes falsos de 0,15/h e de 0,11/h para um horizonte de previsão de 30 minutos e de 50 minutos, respetivamente. Os autores reportaram uma sensibilidade média de 75,8% e uma taxa de previsões falsas de 0,09/h sem qualquer informação de horizonte de previsão usado na obtenção deste resultado e um tempo médio

de previsão de 49,7 minutos [72].

Rasekhi e os seus colaboradores combinaram 22 características univariadas extraídas de seis canais de EEG em 10 pacientes com epilepsia refratária num espaço de características, com o objetivo de estudar o seu desempenho na previsão de crises. Entre as características extraídas destacaram-se os parâmetros de Hjorth (mobilidade e complexidade), a potência relativa da banda gama, os coeficientes *wavelet* das bandas 1 e 2 e a variância [73].

Em 2014, foi proposto um método baseado na sincronização de fase da atividade neuronal elétrica de diferentes áreas do cérebro [74]. Neste estudo, foi possível observar tanto um aumento como um decréscimo da sincronização de fase, anterior às crises. Foi reportado que o poder preditivo da característica base do método não era melhor que um classificador aleatório. Para calcular a coerência de fase média, foi utilizado um modo de decomposição empírico bivariado combinado com a transformada de Hilbert. Para uma taxa de falsas previsões máxima de 0,15/h, foi obtida uma sensibilidade média entre 25% e 70%, para um horizonte de previsão entre 2 minutos e 40 minutos, respetivamente, com um tempo de intervenção de 10 minutos.

Foi estudado um algoritmo baseado num modelo computacional com parâmetros calculados através do ajuste do modelo à densidade de potência espectral de sinais de EEG intracraniano e integrados através da informação obtida investigando as alterações anteriores às crises. O seu desempenho foi avaliado para um horizonte de previsão de 30 minutos e um de 50 minutos, obtendo uma sensibilidade média de 87,07% e 92,6% e FPR de 0,2/h e 0,15/h, respetivamente. Com condições que permitiram maximizar a especificidade, a sensibilidade diminuiu para 82,9% e 90,05% e a taxa de falsas previsões para 0,16/h e 0,12/h para os horizontes de previsão considerados anteriormente [75].

Mais tarde, foi apresentado um método de deteção e previsão de crises que usava uma análise simbólica dos dados de EEG baseada na modelação N-gram [76], que é uma técnica probabilística de reconhecimento de padrões que identifica e prevê a ocorrência de sequências de dados simbólicos dentro dos dados. Depois da aplicação de um filtro *notch* de 50Hz e um filtro de suavização para remover artefactos, os dados foram quantificados de 16 a 8 bits para serem adaptados ao algoritmo N-gram. Em seguida, foi usada uma janela deslizante de 1 minuto com o algoritmo para procurar padrões repetidos avaliando iterativamente a semelhança entre padrões de diferentes tamanhos através da

distância inversa de Hamming. Foi definido um limiar fixo baseado no desvio padrão do fundo e um dinâmico baseado na média móvel da contagem dos padrões para gerar avisos quando ultrapassados. Estes valores, o número de padrões de repetição e o comprimento do padrão foram otimizados na fase de treino usando uma crise e 1 hora de dados do período inter-ictal de cada paciente. O método foi testado para 2 horizontes de previsão, apresentando uma sensibilidade de 67% e uma taxa de falsas previsões igual a 0,04/h em pacientes temporais e para um horizonte de previsão de 20 minutos. Nos casos do lobo frontal, a sensibilidade obtida foi 72% e a FPR 0,62/h. Aumentando o número de crises na fase de treino foi possível obter uma sensibilidade até 100% e uma taxa de falsos alarmes de 0,06/h, para os casos do lobo temporal. O tempo de previsão não foi referido [77].

Em [78] foram analisadas 34 características baseadas na energia e em dinâmicas não lineares (expoentes máximos de Lyapunov e dimensão de correlação). Catorze dessas características foram usadas para comparar o algoritmo com algoritmos que as usavam anteriormente. As restantes 20 basearam-se em descritores estatísticos do EEG, na potência espectral em diversas bandas e na frequência espectral de borda. As características foram extraídas através de janelas de 5 segundos sem sobreposição, para cada um dos 6 canais selecionados para cada paciente e reduzidas para 14 atributos por canal, usando um procedimento de seleção de características. O algoritmo proposto era caracterizado pela utilização de um parâmetro de tempo em avanço t , uma vez que o modelo era treinado e usado para prever o início de um episódio epilético entre t e $t+5$ minutos. Fizeram variar este parâmetro entre 0 e 20. Os autores reportaram a sensibilidade, especificidade e a média geométrica destas (*SI score*). A sensibilidade média obtida variou entre 89% e 93,5%. O melhor desempenho, usando o *SI score*, correspondeu a um tempo de intervenção de 1 minutos ($t=1$). Apesar do tempo de previsão não ter sido reportado é possível inferir que deve ser apenas na ordem de alguns minutos [79].

A capacidade de previsão de crises epiléticas de métodos de inteligência computacional em EEG de longa duração foi avaliada em 278 pacientes com epilepsia refratária. Para 32% dos pacientes, foi possível encontrar um preditor capaz de antecipar crises com uma sensibilidade igual ou superior a 50% e com uma taxa de falsos alarmes não superior a 0,15/h. Foi também reportado que a localização do foco epilético, a frequência de amostragem dos dados, a duração do teste, o número de crises nos testes, o

tipo de técnica de *machine learning* e o período pré-ictal têm grande influência no desempenho da previsão. Em média, em todos os pacientes, foi possível obter uma sensibilidade de 73,55% e a FPR igual a 0,28/h para os canais de EEG de escalpe, para os registros de EEG intracranianos a sensibilidade obtida foi 67,66% e a taxa de falsos alarmes 0,39/h, diferenças que não se mostraram estatisticamente significativas quando avaliadas com o teste de significância de Kruskal-Wallis. Foi ainda reportado que nenhuma das três abordagens de seleção de elétrodos testadas levou a uma melhoria significativa de desempenho, o que sugeriu que diferentes áreas cerebrais estão envolvidas nos fenômenos pré-ictais e que essas áreas podem ser específicas para cada paciente [80].

Um algoritmo baseado em rácios de potência espectral foi treinado e testado com dados contínuos de longa duração de 24 pacientes. Após a aplicação de um filtro *notch* a 50Hz, foi calculada a potência espectral normalizada das bandas espectrais de EEG (Delta, Teta, Alfa, Beta e Gama) através de uma janela deslizante de 5 segundos e a razão entre estas como forma de rastrear as mudanças pré-ictais. Esta razão foi calculada usando todas as combinações de canais e das bandas espectrais de EEG. Após a redução do número de características obtidas, através da aplicação de um método de seleção destas baseado na diferença máxima nos histogramas de distribuição de amplitude do período pré-ictal e do período não pré-ictal, foi treinado um classificador SVM usando subconjuntos de dados depois de realizado o balanceamento do número de amostras correspondentes ao pré-ictal e ao inter-ictal, para uma classificação imparcial. Foram usadas 3 crises para treinar o modelo e as restantes para testar e o horizonte de previsão foi otimizado na fase de treino. O desempenho do modelo foi reportado separadamente para os pacientes com EEG de escalpe e com EEG intracraniano. Em média, nos últimos foi registada uma sensibilidade de 78,36% e uma taxa de falsas previsões igual a 0,15/h, para um horizonte de previsão de 33,7 minutos. Foi comentado que o EEG intracraniano contém informação epileptógena mais clara e localizada do que o de escalpe, mas para obter resultados melhores podem ser necessárias características menos discriminantes [81].

Em 2017, foi apresentado um algoritmo, para previsão de crises epiléticas, específico para cada paciente, baseado em SVM multiclasse, usando conjuntos de características multicanal originando um espaço de classificação de alta dimensão. O estudo foi realizado com 216 pacientes, 185 com registros de EEG de escalpe e 31 com registros de EEG

intracraniano, alcançando uma sensibilidade geral de 38,47% e uma taxa de falsas previsões igual a 0,20/h. Para 11% dos pacientes, o método alcançou um desempenho com significância estatística, o que corrobora a hipótese de existência de um período pré-ictal identificável para um pequeno subgrupo de pacientes [82].

Quando analisados estudos relativos à identificação de crises epiléticas é possível verificar que estes se baseiam em características extraídas de sinais de EEG e classificadores semelhantes aos analisados na previsão de crises, como transformadas de *wavelet* [83]–[85], análises tempo-frequência [86] e classificadores SVM [87], [88]. Foram também quantificadas, medidas e detetadas semelhanças de potenciais elétricos nas proximidades do foco epilético, como forma de obter um mecanismo de detecção de crises [89].

No que diz respeito à análise do estado ictal em ratos, em 2008, foi sugerido um método de análise quantitativo. Este método foi usado, especialmente, para analisar o comportamento da banda gama de sinais de EEG e assim quantificar a energia da atividade representada no estado epilético. Como foi demonstrado o aumento da atividade na banda gama, o método foi desenvolvido para estimar a eficácia do tratamento medicamentoso em mudanças da potência dos sinais de EEG na banda gama [90].

No melhor do nosso conhecimento parece não haver estudos de previsão/detecção de crises usando sinais relativos à pO_2 em humanos e modelos animais, o que torna os resultados apresentados nesta tese inovadores.

4. Materiais e Métodos

4.1. Dados

Os dados utilizados correspondem a registos *in vivo* em animais acordados e em comportamento como descrito detalhadamente em [6]. De forma resumida, a corrente amperométrica registada foi obtida a um potencial constante de -0,6V vs. Ag/AgCl/NaCl (0,15mM), utilizando um potencióstato para controlo e aquisição (FAST16mkIII, Quanteon LLC). Cada registo foi obtido utilizando um *array* de microelétrodos de platina com 4 sítios ativos, implementado no cérebro de ratos Wistar macho com 8 semanas de idade. A indução de crise epilética foi realizada por administração intra-hipocampal de uma solução de pilocarpina (1,2 mg em 1 μ L de salino 0,9%). A administração foi efetuada a um fluxo de 0,5 μ L min^{-1} através de um microinjetor ligado a uma bomba de infusão. Durante todo o procedimento, cada animal permaneceu no interior de uma caixa de comportamento (50x50x50 cm).

A cada conjunto de dados corresponde um ficheiro com informações relativas à frequência de amostragem e eventos assinalados ao longo da aquisição.

Uma vez que as crises epiléticas registadas no conjunto de dados analisados são crises induzidas, foram apenas identificados os instantes em que estas ocorreram, originando assim duas classes, ictal e não-ictal. Esta informação foi fornecida em conjunto com os dados utilizados. Esta informação tal como os detalhes da duração de cada registo e sua frequência de amostragem pode ser consultada na Tabela 4.1.

Tabela 4.1: Descrição dos conjuntos de dados utilizados no estudo

Nome	Duração total (min)	Frequência de amostragem (Hz)	Intervalo em que ocorreu a crise (segundos)	Observações
WO2_005	75	100	[2874; 3935]	Morte do animal no minuto 66
WO2_007	174	100	[1717; 7024]	Administração de Diazepam para controlo da crise no minuto 117
WO2_017	207	100	[8510; 10502]	Administração de Diazepam para controlo da crise no minuto 208, seguida da morte do animal
WO2_020	174	100	[4032; 9346]	Administração de Diazepam para controlo da crise no minuto 156
WO2_024	346	200	-	Sem identificação do período ictal

4.2. Filtros

Com o objetivo de selecionar apenas gamas de frequência de interesse foram aplicados filtros digitais aos sinais, cuja resposta em frequência se caracteriza pela existência de bandas de passagem e de rejeição, definindo a classificação dos mesmos como passa-alto, passa-baixo, passa-banda ou rejeita-banda.

A saída de filtros digitais pode ser obtida através da convolução do sinal de entrada com a resposta ao impulso do filtro. Quando a resposta a impulso é usada com este fim, é denominada *kernel*.

Em termos gerais podemos falar na existência de dois tipos de filtros: os filtros não recursivos ou de resposta finita (FIR – *Finite Impulse Response*) e os de resposta infinita (IIR – *Infinite Impulse Response*).

Na aplicação de filtros FIR, cada amostra de saída resulta da combinação linear das entradas do sinal. Considerando que $x(n)$ é o sinal de entrada, P a ordem do filtro e h_i os coeficientes do filtro, o sinal de saída $y(n)$ é dado por:

$$y(n) = \sum_{i=0}^P h_i x(n-i) \quad (4.1)$$

Por outro lado, no caso dos filtros recursivos, são usados valores calculados anteriormente na saída. A resposta de filtros IIR como o nome indica propaga-se por tempo infinito. Estes filtros são muito rápidos, menos flexíveis quando comparados com os FIR e podem tornar-se instáveis. No entanto têm a vantagem de com uma ordem muito mais reduzida se obter o mesmo efeito que com filtros FIR. Neste conjunto de filtros, foram usados quatro tipos – Butterworth, Chebyshev tipo 1 e tipo 2 e elíptico.

Quando não existe ondulação na resposta de frequência (figura 4.1), o filtro é denominado Butterworth. A estes correspondem respostas de frequência suaves, o mais planas possíveis na banda de passagem e na banda de corte. A alteração da sua ordem implica apenas a alteração da largura da banda de transição (*rol-off*).

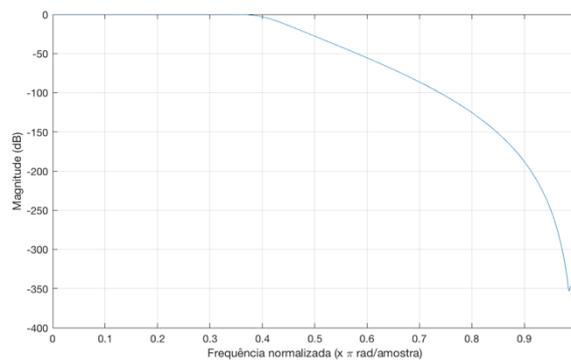


Figura 4.1: Resposta em frequência do filtro Butterworth passa-baixo de ordem 10.

O filtro Chebyshev apresenta uma estratégia matemática que permite alcançar uma banda de transição mais curta (*rol-off* mais rápido), permitindo ondulações na resposta em frequência. À medida que a ondulação aumenta (mau), a banda de transição torna-se mais estreita, assim a resposta deste filtro é um *trade-off* ótimo entre estes dois parâmetros. Os filtros Chebyshev tipo 1 apresentam ondulações apenas na banda de passagem (figura 4.2 A), e os de tipo 2 na banda de corte (figura 4.2 B). Estes filtros, em comparação com o filtro Butterworth apresentam uma banda de transição mais estreita, quando usados os mesmos coeficientes.

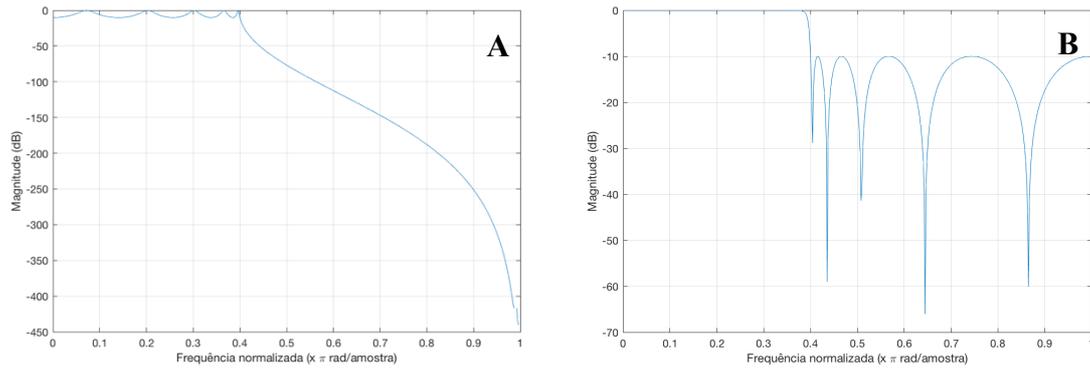


Figura 4.2: Resposta em frequência dos filtros Chebyshev passa-baixo tipo 1 e tipo 2, respectivamente, de ordem 10.

Quando existem ondulações na banda de passagem e de rejeição na resposta de frequência (figura 4.3), o filtro é denominado elíptico. Este, é o filtro que apresenta a banda de transição mais estreita, quando comparado com os anteriores, apresentando também mais ondulações em toda a largura de banda.

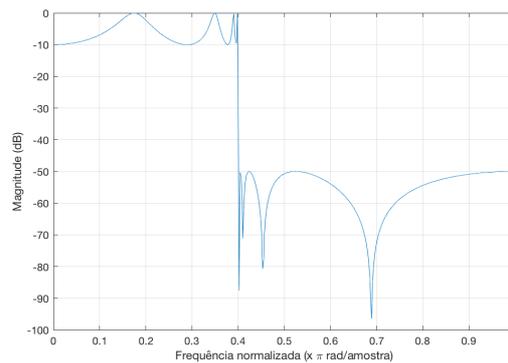


Figura 4.3: Resposta em frequência do filtro Elíptico passa-baixo de ordem 10.

4.3. **Análise de Componentes Independentes (ICA)**

A ICA representa uma via de separação cega de fontes, ou seja, faz uma separação de um sinal multivariado nas suas fontes latentes, assumindo que estas fontes latentes não seguem distribuições gaussianas e são estatisticamente independentes. Com o objetivo de realizar essa separação, foi utilizado o algoritmo *fastICA*, que é um algoritmo eficiente e popular nesta análise, desenvolvido por Aapo Hyvärinen [91].

Na implementação do ICA surge o processo de maximizar ou minimizar uma função objetivo, neste caso pretende-se maximizar a função de aproximação da neguentropia,

também conhecida por função de contraste, uma vez que mede a diferença entre a gaussianidade e a não-gaussianidade, dada por:

$$J(y_i) \propto [E\{G(y_i)\} - E\{G(v)\}]^2 \quad (4.2)$$

Em que $G(\cdot)$ pode assumir diferentes formulações, por exemplo:

$$G_1(u) = \frac{1}{a_1} \log \cosh a_1 u, \quad 1 \leq a_1 \leq 2 \quad (4.3)$$

ou

$$G_2(u) = -e^{-\frac{u^2}{2}} \quad (4.4)$$

u representa y_i ou v de acordo com o caso

v é uma variável gaussiana de média zero e variância unitária

$y_i = w^T x$ representa uma componente independente obtida através da projeção do vetor de mistura x no espaço definido por w .

Este algoritmo consiste num processo iterativo que progride sucessivamente para soluções até alcançar a solução ótima, começando com uma solução arbitrária. Assim, este método define uma direção (através do vetor unitário w) cuja projeção $w^T x$ maximiza a não-Gaussianidade através da aproximação da neguentropia [92].

4.4. Extração de características

Como forma de analisar os sinais recolhidos, foram extraídas características dos mesmos, através de janelas deslizantes com duração de 1 segundo, sem sobreposição.

4.4.1. Medidas estatísticas

Normalmente, a análise estatística é realizada através do cálculo da média, da mediana, da obliquidade e do *kurtosis*. Estes parâmetros são conhecidos, respetivamente, como momentos de primeira, segunda, terceira e quarta ordem, que permitem obter informação relativa à distribuição da amplitude de séries temporais.

A média (equação 4.5) permite determinar o ponto onde se concentram os dados.

$$\mu = \frac{1}{N} \sum_{i=1}^N x_i \quad (4.5)$$

x_i : representa cada ponto da amostra

N : número de pontos da amostra

A variância (equação 4.6) corresponde a uma medida de dispersão, indicando se os valores observados se encontram muito ou pouco dispersos em torno da média.

$$\sigma^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \mu)^2 \quad (4.6)$$

x_i : representa cada ponto da amostra

μ : média da amostra

A obliquidade, assimetria ou *skewness* (equação 4.7) é uma medida de assimetria dos dados em torno do valor médio da amostra. Se a assimetria for negativa, significa que os valores se encontram mais espalhados à esquerda da média, se for positiva, encontram-se espalhados mais à direita. A obliquidade de uma distribuição normal é zero.

$$s(x) = \frac{E(x-\mu)^3}{\sigma^3} \quad (4.7)$$

x : amostra

μ : média da amostra

σ : desvio padrão da amostra

O *kurtosis* ou curtose (equação 4.8) é uma medida que permite caracterizar o achatamento da curva da função de distribuição. Numa distribuição normal, o valor desta medida é igual a 3. Se superior a 3, significa que a curva da função de distribuição

apresenta um pico mais alto e é mais afunilada em relação à distribuição normal. Se for inferior a 3, significa que a curva da função de distribuição é mais achatada do que a correspondente à distribuição normal.

$$k(x) = \frac{E(x-\mu)^4}{\sigma^4} \quad (4.8)$$

x : amostra

μ : média da amostra

σ : desvio padrão da amostra

4.4.2. Coeficientes *wavelet*

As transformadas de *wavelet* permitem decompor outra função no domínio temporal, com o objetivo de a analisar em diferentes escalas de frequência e de tempo, ou seja, permitem descrever o sinal através dos componentes de frequência presentes e a sua localização na série temporal. Este tipo de análise tem tomado um papel cada vez mais importante no processamento de sinais, descrevendo os sinais através de coeficientes característicos de uma onda tipo e com propriedades matemáticas.

Com o objetivo de decompor os sinais em bandas de frequência semelhantes às das ondas Delta, Teta, Alfa, Beta e Gama, foi feita a decomposição do sinal através da *Discrete Wavelet Transform* (DWT), usando a função Daubiches-4 (db4) para descrever os sinais em seis níveis, para sinais com frequência de amostragem igual a 200Hz e em cinco níveis para sinais adquiridos a 100Hz, uma vez que a banda de sinal útil é metade da banda total.

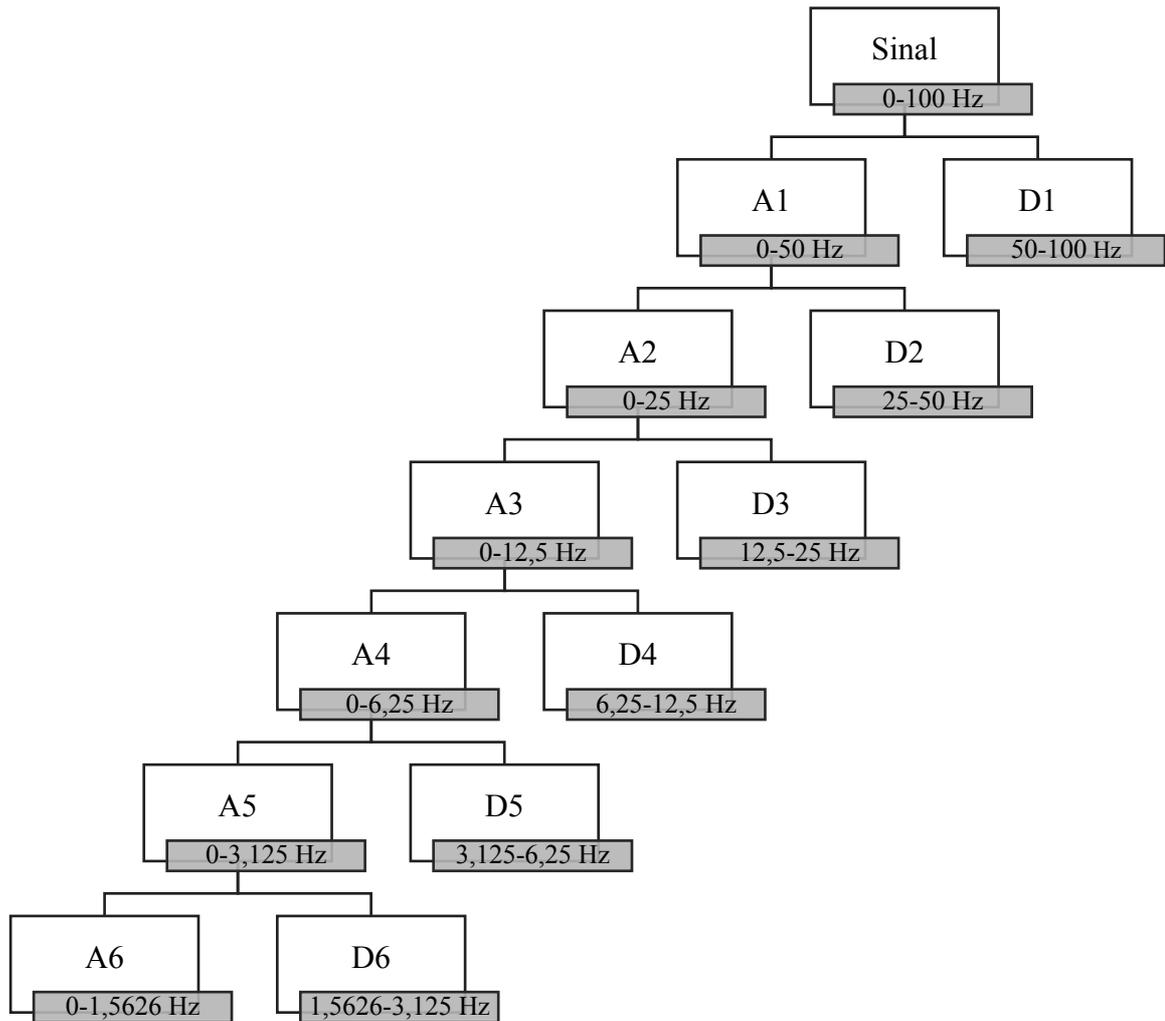


Figura 4.4: Esquema representativo do processo de decomposição implementado pela DWT para sinais com frequência de amostragem igual a 200Hz.

Desta forma, D6 irá corresponder a ondas Delta (0,5-4 Hz), D5 a ondas Teta (4-7 Hz), D4 a ondas Alfa (8-12 Hz), D3 a ondas Beta (13-30 Hz) e D2 e D1 a ondas Gama (>30 Hz). Quando são obtidos apenas cinco níveis de decomposição, D5 corresponderá a ondas Delta, D4 a ondas Teta, D3 a ondas Alfa, D2 a ondas Beta e D1 a ondas Gama. Os intervalos obtidos não são exatamente iguais aos desejados devido ao processo de decomposição das *wavelets*, possível de observar na figura 4.4, mas representam a aproximação possível.

Para cada janela são gerados cinco ou seis atributos, dependendo da frequência de amostragem, que correspondem à média quadrática dos coeficientes (equação 4.6), de cada um dos níveis.

$$x_q = \sqrt{\frac{x_1^2 + x_2^2 + \dots + x_n^2}{n}} \quad (4.9)$$

x_i : cada ponto da amostra

n : número de pontos da amostra

4.4.3. Mapa Tempo-Frequência

Com o objetivo de estudar o sinal no domínio temporal e de frequência, em simultâneo, tal como se fez com as *wavelets*, foram também calculados mapas de tempo-frequência através do modelo *short-time auto-regressive*, que permite obter estimativas da distribuição tempo-frequência de sinais não estacionários que são aplicadas a janelas consecutivas, apresentando como vantagem, o facto da sua resolução não ser afetada pela segmentação do sinal, ou seja, pelo tamanho e tipo de janela.

4.4.4. Potência Relativa

A potência relativa das bandas de frequência das ondas Delta, Teta, Alfa, Beta e Gama representa a razão entre a potência de cada uma das bandas e a potência total.

Esta medida é calculada, recorrendo ao mapa tempo-frequência, dividindo cada janela nas bandas pré-definidas da mesma forma como se encontra dividido o mapa tempo-frequência representado na figura 4.5. Após esta divisão é calculada a soma dos coeficientes de cada uma das bandas individualmente e a sua soma na totalidade e realizado o quociente entre estes valores obtidos para cada banda. Após a realização das operações anteriores são obtidos 5 parâmetros para cada janela analisada, cada valor corresponde a uma banda.

$$\text{Potência Relativa}_{\text{banda } i} = \frac{\sum \text{coeficientes da banda } i}{\sum_{k=1}^5 \sum \text{coeficientes da banda } k} \quad (4.10)$$

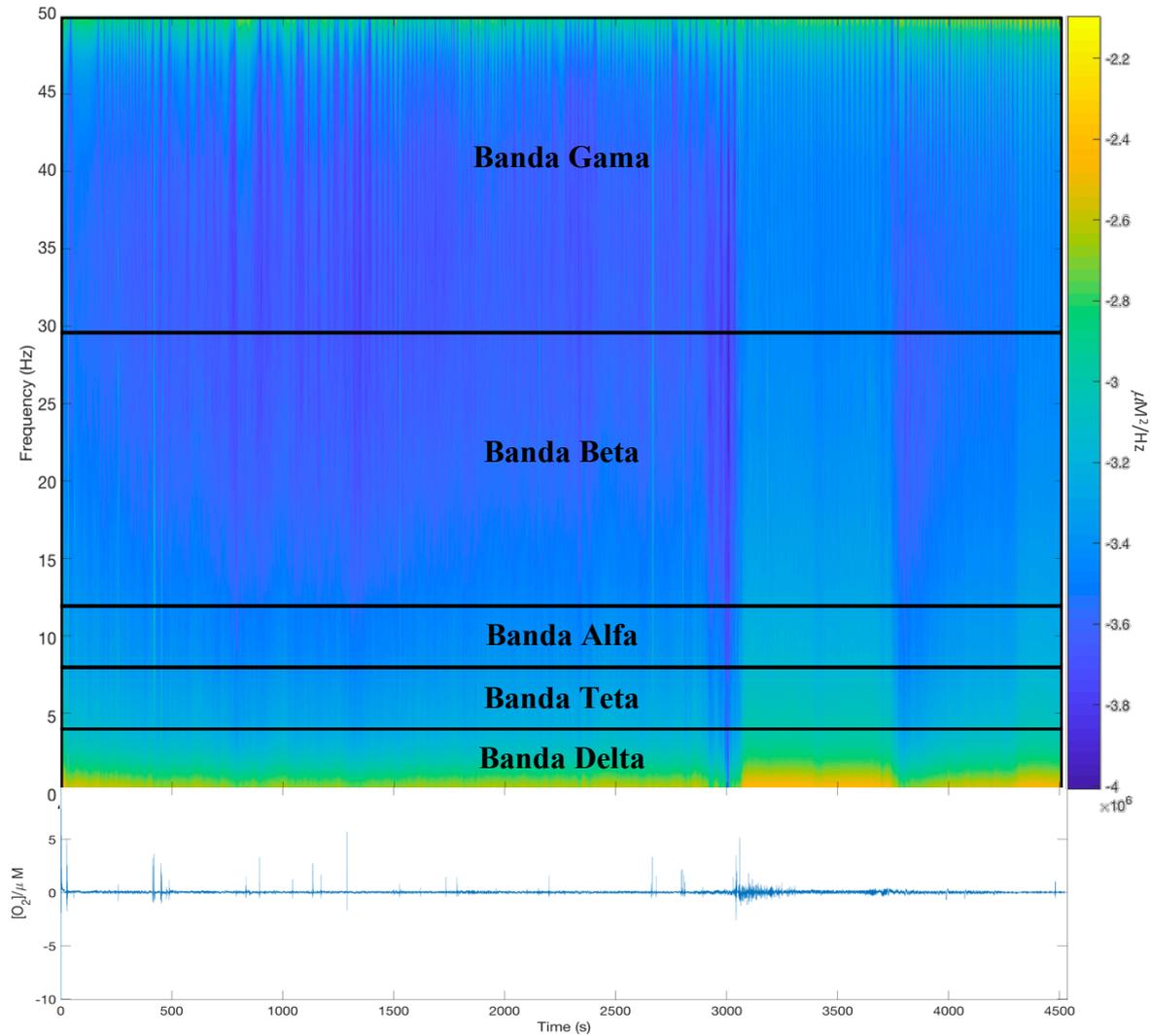


Figura 4.5: Exemplo de mapa tempo-frequência com a representação esquemática da divisão das bandas de frequência pré-definidas e a série temporal associada.

A potência relativa pode também ser calculada para bandas de frequência diferentes das pré-definidas.

Através desta medida é possível observar as variações em cada uma das bandas de frequência ao longo do tempo.

4.4.5. Parâmetros de Hjorth

A atividade, mobilidade e complexidade de Hjorth [93] são três parâmetros usados para descrever os sinais no domínio do tempo e muito usados na análise de EEG.

A atividade (equação 4.11) é quantificada através da variância da amplitude do sinal, desta forma, apenas foram adicionadas a mobilidade e a complexidade ao estudo, uma vez que o outro parâmetro faz também parte das medidas estatísticas utilizadas.

$$H_a = \sigma^2 \quad (4.11)$$

σ^2 : variância

A mobilidade (equação 4.12) corresponde à raiz quadrada do quociente entre a variância da primeira derivada e a variância da amplitude do sinal. Uma vez que estas medidas são igualmente dependentes da amplitude do sinal, a sua relação dependerá apenas da forma da curva, medindo assim, a inclinação média relativa da curva, para cada janela.

$$H_m = \sqrt{\frac{\sigma^2(x')}{\sigma^2(x)}} \quad (4.12)$$

σ^2 : variância

x : amostra

x' : primeira derivada da amostra

A complexidade (equação 4.13) corresponde ao erro quadrático médio da taxa de declive em relação a uma curva ideal possível. Este parâmetro fornece uma estimativa da largura da banda do sinal.

$$H_c = \sqrt{\frac{\sigma^2(x'')x\sigma^2(x)}{(\sigma^2(x'))^2}} \quad (4.13)$$

σ^2 : variância

x : amostra

x' : primeira derivada da amostra

x'' : segunda derivada da amostra

4.4.6. Potência acumulada

A potência acumulada, ou energia de longo termo, permite observar a ocorrência de variações de energia ao longo do tempo. Esta medida é obtida através do cálculo da potência do sinal em cada instante (equação 4.14). Para cada janela de tamanho N do sinal, é calculada a potência média e somada com a potência acumulada da janela anterior (equação 4.15).

$$E_a = \frac{1}{N} \sum_{i=1}^N x_i^2 \quad (4.14)$$

$$AE_m = \frac{i}{N} \sum E_k + AE_{m-1} \quad (4.15)$$

x_i : cada um dos pontos da amostra

E_k : potência instantânea de cada amostra

4.4.7. Modelo Quantitativo

Foi implementado um modelo quantitativo usado na análise de estados epiléticos em ratos [90]. Depois do sinal de O_2 filtrado, é determinada a potência calculando o quadrado do valor absoluto de cada ponto (equação 4.16). Após esta operação é feita uma subamostragem do sinal para uma amostra por segundo (equação 4.17). É calculada uma linha base que corresponde à média do sinal nos primeiros 10 minutos, isto é, 600 amostras (equação 4.18) e de seguida é calculada a potência do sinal em decibéis (dB) ajustada à linha base (equação 4.19). Após o cálculo da potência do sinal, foi ainda determinada uma curva correspondente à média das curvas suaves que permitem delinear os extremos do sinal.

$$e(t) = |f(t)|^2 \quad (4.16)$$

$$d(k) = e(k \times f_s) \quad (4.17)$$

$$b = \frac{1}{600} \sum_{k=1}^{600} d(k) \quad (4.18)$$

$$p(k) = 10 \log_{10} \left(\frac{d(k)}{b} \right) \quad (4.19)$$

$f(t)$: sinal

f_s : frequência de amostragem do sinal

4.5. Normalização dos dados

Quando se está perante um problema de reconhecimento de padrões é comum proceder à normalização dos dados como forma de fazer variar as diferentes características no mesmo intervalo de valores e assim evitar o surgimento de uma tendência no classificador.

Neste estudo foi aplicada a normalização estatística, também conhecida como *z-score normalization*, desta forma o valor médio de cada atributo é zero e a variância um. A transformação aplicada é dada por:

$$x'_i = \frac{(x_i - \mu_i)}{\sigma_i} \quad (4.20)$$

Onde:

x'_i : é a amostra normalizada do atributo i

x_i : é a amostra do atributo i

μ_i : é a média do atributo i

σ_i : é o desvio padrão do atributo i

4.6. Teste U de Mann-Whitney e Teste de Kruskal-Wallis

O teste U de Mann-Whitney ou teste de Wilcoxon-Mann-Whitney, é um teste não paramétrico usado para comparar duas amostras e determinar se foram obtidas a partir de populações com a mesma distribuição. Com este teste pretende-se aceitar ou rejeitar a hipótese nula (os dois grupos apresentam a mesma distribuição), para tal é calculado um valor de U para cada uma das amostras (U_1 e U_2), sendo considerado como U final o valor mínimo calculado. Um valor de U menor indica uma certeza maior de que as populações são diferentes. Da mesma forma quanto menor for o valor- p obtido na aplicação deste teste, maior é a probabilidade de rejeição da hipótese nula.

Quando o número de classes (grupos) é superior a dois, é usado o teste de Kruskal-Wallis. Este teste é uma extensão do anterior e é usado para testar a hipótese nula de que todos os grupos têm distribuições semelhantes contra a hipótese alternativa de que pelo menos dois apresentam distribuições diferentes. A estatística F utilizada na tabela ANOVA clássica dá lugar a uma estatística de qui-quadrado e o valor- p mede a significância desta.

4.7. Classificadores

Em estatística, reconhecimento de padrões e *machine learning* são usados métodos de análise discriminante como forma de encontrar uma combinação linear de atributos que caracterizem duas ou mais classes de objetos ou eventos. Esta combinação pode ser usada como um classificador linear. Neste trabalho foram considerados classificadores lineares de forma a potenciar a interpretabilidade dos resultados.

4.7.1. Classificador *Fisher LDA*

Como forma de identificar um hiperplano de separação entre duas classes foi implementado um classificador linear, o *Fisher LDA* fornecido na *toolbox* de reconhecimento de padrões *stprtool* desenvolvida para o *Matlab*, bastante usado em reconhecimento de padrões, uma vez que ambiciona obter uma separação máxima entre dois grupos e minimizar a variabilidade dentro das classes, com base numa combinação linear das características. A transformação aplicada procura fazer uma rotação dos eixos

com o objetivo de maximizar as diferenças entre os grupos, quando são projetados nos novos eixos.

A matriz de dispersão dentro da classe é dada por:

$$S_w = \sum_{k=1}^2 \sum_{x \in c_k} (x - m_k)(x - m_k)' \quad (4.21)$$

A matriz de dispersão entre classes é dada por:

$$S_b = (m_1 - m_2)(m_1 - m_2)' \quad (4.22)$$

Este método pretende maximizar o critério:

$$J(w) = \frac{w'S_b w}{w'S_w w} \quad (4.23)$$

Onde

w : direção de projeção a determinar

m_k : representa o valor médio de cada classe

c_k : representa cada classe

x' : matriz de dados a classificar (x) transposta

A direção de w que maximiza $J(w)$ é dada por:

$$w = S_w^{-1}(m_1 - m_2) \quad (4.24)$$

Após estas operações, é possível identificar um plano de separação e analisar a sua capacidade para classificar corretamente os dados através do cálculo da acurácia, precisão sensibilidade e especificidade.

Para o caso em que são identificadas mais do que duas classes, são realizadas classificações binárias, através da oposição de cada classe ao conjunto das restantes (*One vs All*), que são depois combinadas e permitem identificar a que classe pertence cada padrão.

4.7.2. Classificador baseado na distância euclidiana mínima

Os classificadores baseados na distância mínima permitem determinar o grupo a que pertence cada amostra através da distância a que esta se encontra dos valores centrais, isto é da média, de cada um dos grupos, ou seja, a classe atribuída a cada amostra corresponde à classe do grupo mais próximo desta. Estes classificadores podem ser definidos baseando-se em diferentes métricas para calcular as distâncias. O método de classificação testado neste estudo baseia-se no cálculo da distância euclidiana, uma das métricas mais utilizadas neste tipo de análises.

Considerando os pontos $P = (p_1, p_2, \dots, p_n)$ e $Q = (q_1, q_2, \dots, q_n)$, a distância euclidiana entre eles é dada por:

$$d = \sqrt{\sum_{i=1}^n (p_i - q_i)^2} \quad (4.25)$$

Na verdade este tipo de classificadores calculam a distância entre os padrões e a média de forma implícita, ou seja, são baseados na definição de funções discriminantes para cada classe dadas por [94]:

$$g_i(\mathbf{x}) = \mathbf{m}_i^T \mathbf{x} - \frac{1}{2} \mathbf{m}_i \quad (4.26)$$

onde m_i corresponde ao vetor constituído pelos valores médios das características da classe i . Finalmente, a atribuição de uma determinada classe a um padrão x é dada de acordo com:

$$x \in \omega_i \text{ if } \max_i \{g_i(\mathbf{x})\} \quad (4.27)$$

A função de decisão entre cada par de classes é dada por:

$$g_{ij}(\mathbf{x}) = g_i(\mathbf{x}) - g_j(\mathbf{x}) \quad (4.28)$$

5. Aplicação computacional para processamento de sinais de O₂

Foi desenvolvida uma aplicação para analisar e processar os sinais recolhidos através dos *arrays* de microelétrodos de platina dividida em quatro painéis principais. O primeiro corresponde ao painel onde são importados os dados recolhidos (figura 5.1).

Neste painel, é possível carregar os dados em ficheiros *.xls* ou *.xlsx* e representá-los graficamente de forma individual ou em conjunto. É também possível adicionar mais dados aos já carregados.

Na janela, existe também a opção para importar ficheiros *.xml* ou *.csv* com informações relativas aos eventos sinalizados durante a aquisição dos dados, bem como a frequência de amostragem do sinal. Após a importação destes, é possível sinalizar no gráfico o instante correspondente a cada evento.

No segundo painel (figura 5.2) é possível proceder ao pré-processamento dos dados, isto é, aplicar filtros IIR (Butterworth, Elíptico, Chebyshev tipo 1 e tipo 2) de diferentes tipos (passa-baixo, passa-alta, passa-banda e rejeita-banda). É também possível desenhar e aplicar um filtro de resposta finita ao impulso através da definição da amplitude de atenuação (entre 0 e 1, em que 0 representa rejeição total e 1 passagem total) para cada frequência definida. O painel contém também uma opção que permite fazer a análise dos componentes independentes através da aplicação do algoritmo *fastICA*. É possível seleccionar componentes, remove-los e reconstruir o sinal. No caso de as alterações obtidas não serem as desejadas, é possível repor o sinal original.

No terceiro painel (figura 5.3) é feita a extração de características do sinal. Neste painel, é possível extrair e representar cada uma das características individualmente ou extrair todas em conjunto. Para além das características predefinidas, é ainda possível visualizar a potência relativa para bandas diferentes das bandas padrão.

O conjunto de todas as características é composto pela média, mediana, obliquidade, *kurtosis*, coeficientes *Wavelet* (5 quando a frequência de amostragem é igual a 100Hz e 6 quando é igual a 200Hz), potência relativa das bandas Delta, Teta, Alfa, Beta e Gama, mobilidade, complexidade, potência acumulada, potência e curva correspondente obtidas

através do método quantitativo. Apesar do mapa tempo-frequência não integrar este conjunto e ser utilizado no cálculo da potência relativa, este pode ser representado na aplicação.

No quarto painel (figura 5.4), podem ser representadas duas ou três características em simultâneo, definir um *target* através da delimitação do intervalo (em segundos ou pontos de amostragem) onde ocorre o evento sinalizado, ordenar o conjunto de todas as características de acordo com a capacidade de discriminação dos grupos (através do teste de Mann-Whitney para dois grupos ou do teste de Kruskal-Wallis para três ou mais). Após a definição de cada grupo, é possível treinar um classificador linear, representar as linhas ou planos de separação obtidas quando usadas duas ou três características, uma vez que não é possível representar mais do que três dimensões, e avaliar a sua capacidade de separação dos grupos através da acurácia, sensibilidade, especificidade e precisão. No treino do classificador, é possível utilizar todas as características, as N primeiras do *ranking* obtido anteriormente ou usar as características selecionadas. As características utilizadas neste painel foram normalizadas, como forma de uniformizar os seus intervalos de variação.

Neste painel é ainda possível avaliar a capacidade de categorizar os dados através de um classificador baseado na distância euclidiana mínima. É também possível gravar um modelo de classificação e testá-lo posteriormente noutros conjuntos de dados.

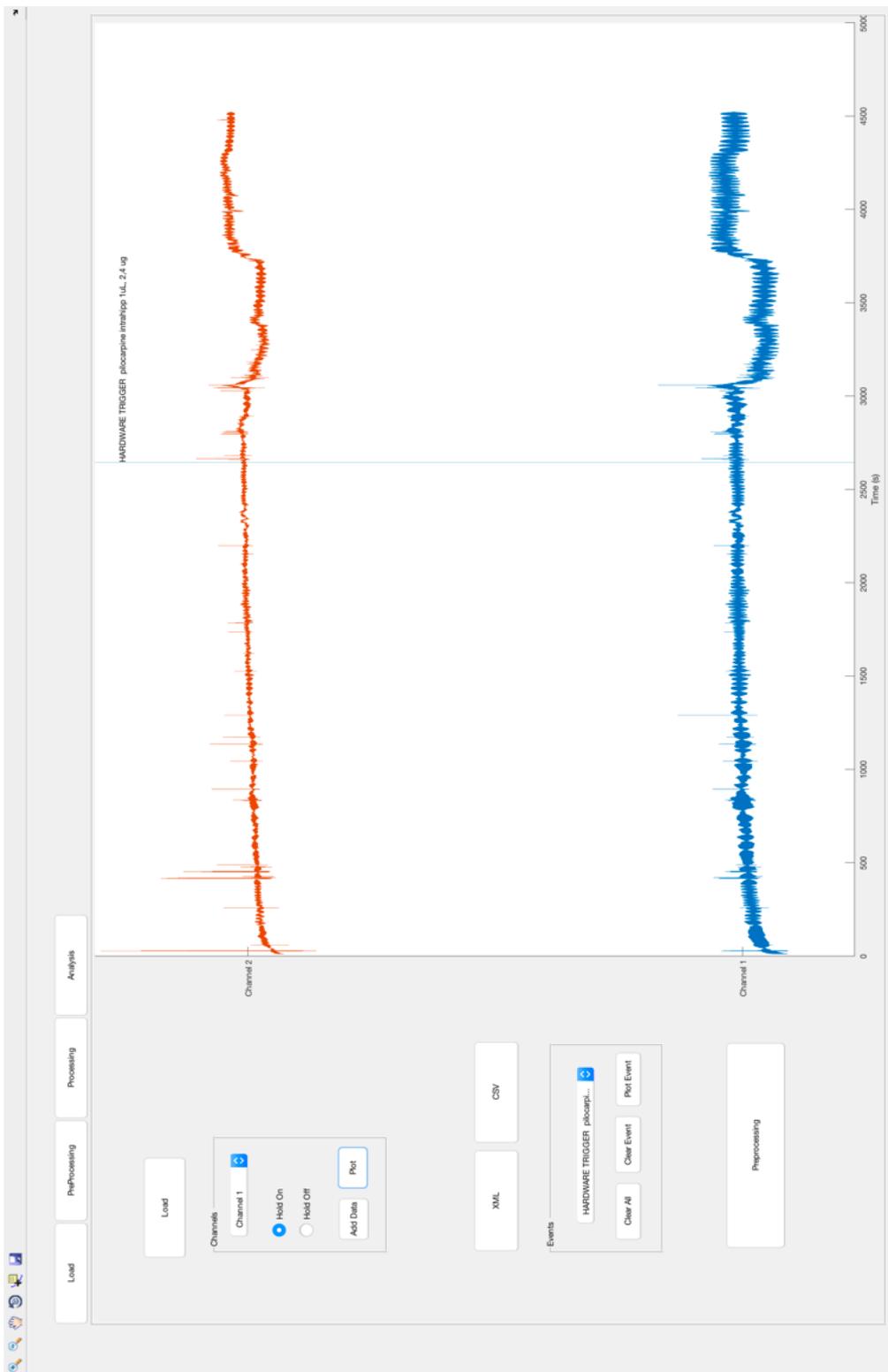


Figura 5.1: Interface de importação de dados da aplicação com o canal 1 e 2 do conjunto de dados WO2_005 representados e o instante correspondente à injeção de pilocarpina assinalado.

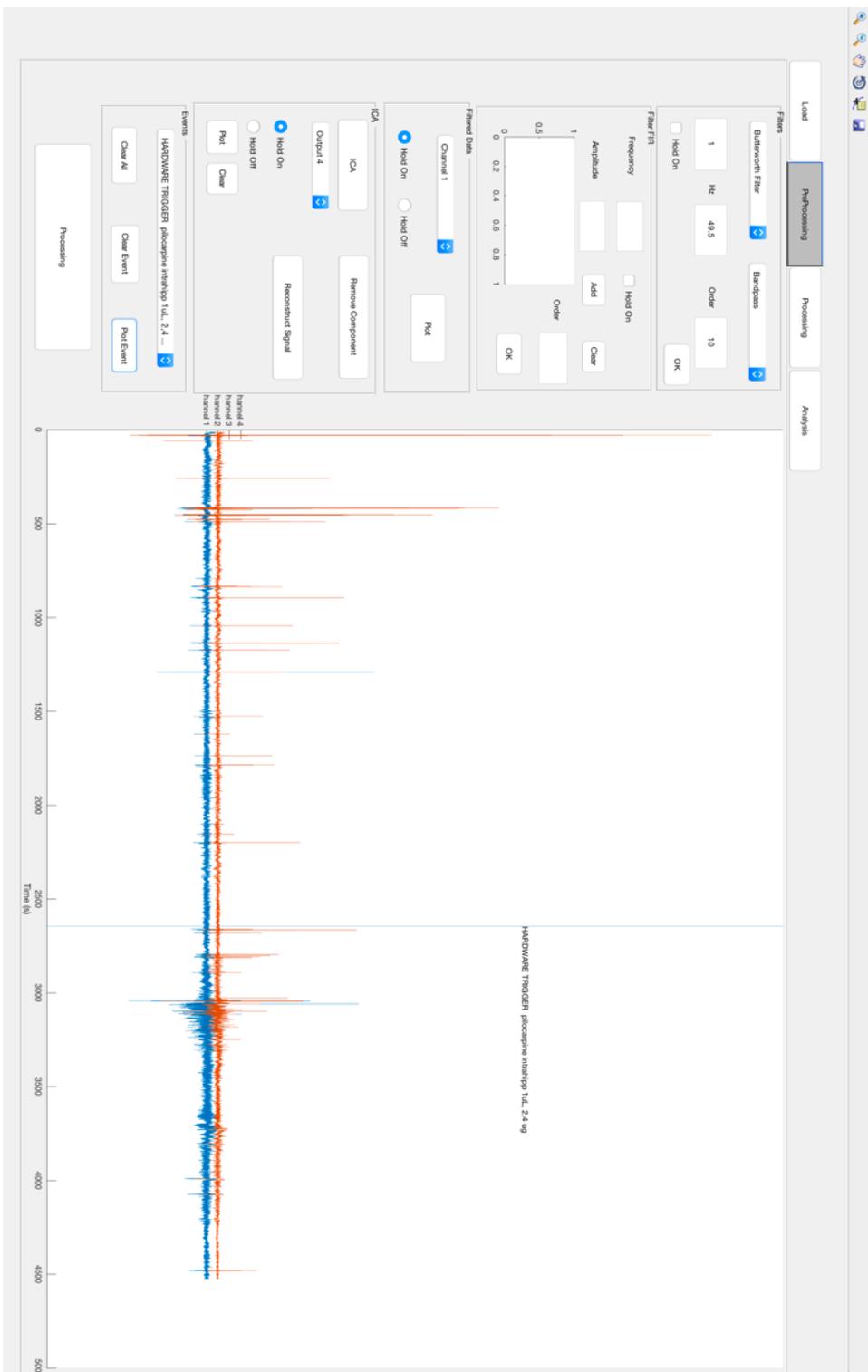


Figura 5.2: Interface da aplicação correspondente ao painel de pré-processamento, com a representação dos sinais correspondentes ao canal 1 e 2 do conjunto de dados WO2_055 após a aplicação de um filtro Butterworth passa-banda de ordem 10, com banda de passagem de 1 a 49,5Hz, e o instante em que ocorreu a injeção de pilocarpina assinalado.

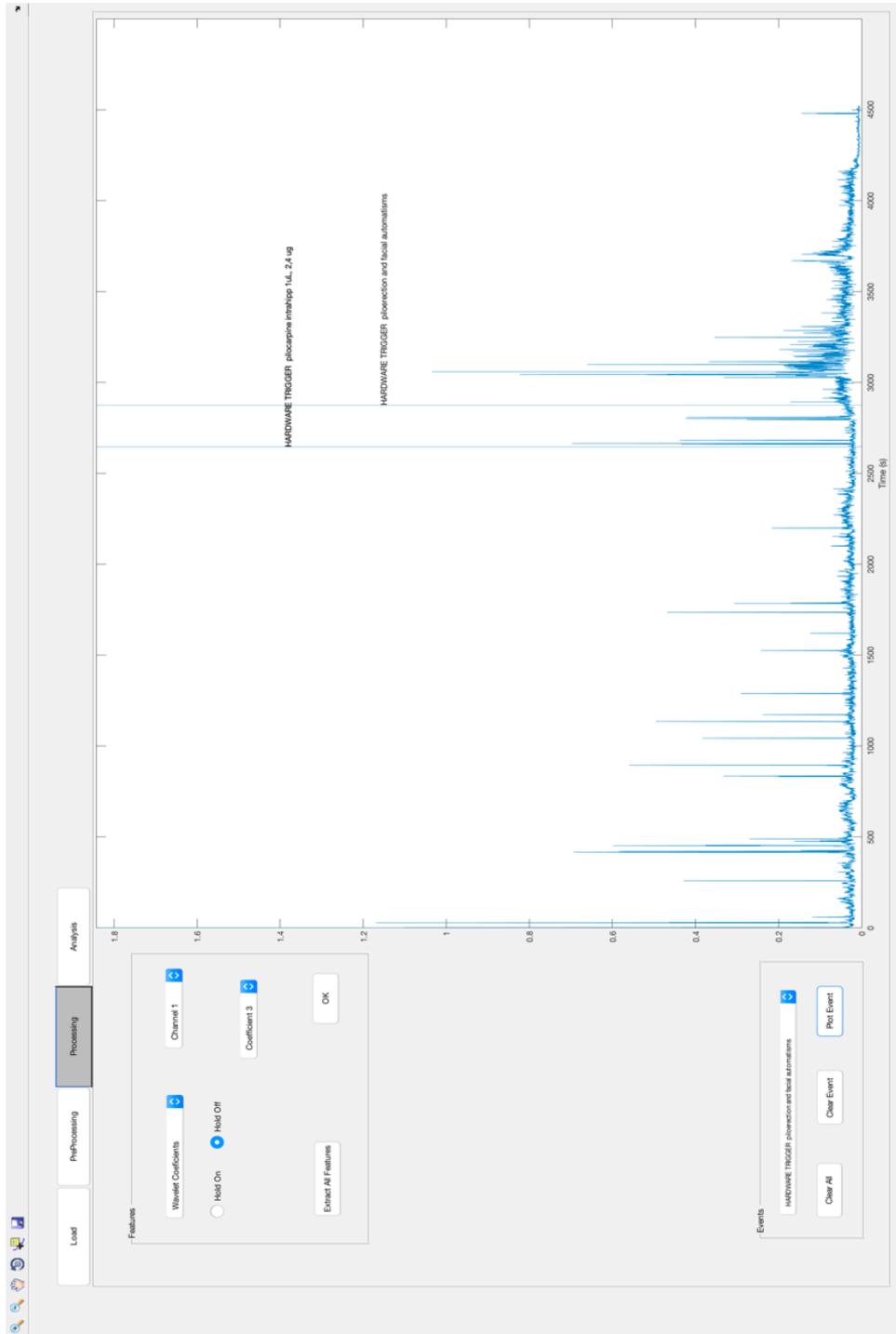


Figura 5.3: Interface da aplicação correspondente ao painel de extração de características, com a representação do *kurtosis* ao longo do tempo e os instantes da injeção de pilocarpina e da ocorrência de automatismos faciais assinalados.

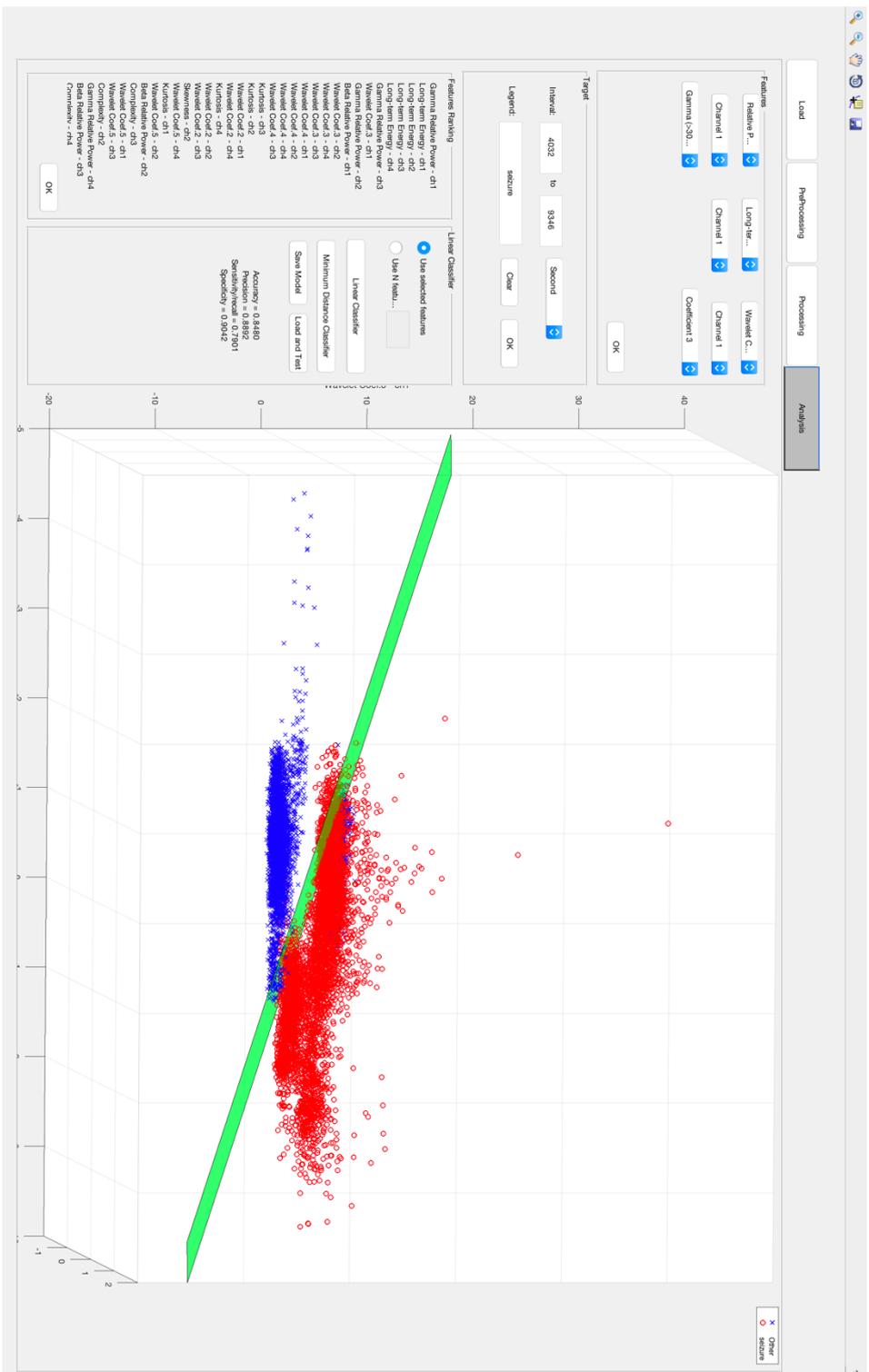


Figura 5.4: Interface da aplicação correspondente ao painel de análise de várias características em simultâneo, com a representação da potência relativa da banda gama correspondentes ao canal 2 em função da potência acumulada dos canais 1 e 2 e do plano de separação dos dois grupos (*seizure* e *other*), definidos através da opção de definição do *target*, obtida através do treino de um classificador linear.

6. Resultados e discussão

O objetivo principal deste estudo foi desenvolver uma aplicação para analisar os sinais recolhidos e encontrar padrões que permitissem identificar o período que precede as crises epiléticas e assim ser capaz de as prever. Contudo, as crises presentes nos dados existentes são crises induzidas, o que não permite avaliar as alterações que antecedem crises espontâneas. A não existência de dados relativos a crises espontâneas deve-se ao facto de os animais acabarem por morrer durante, ou pouco tempo após a crise induzida e, nos que sobrevivem, os conjuntos de microelétrodos não aguentam implantados tempo suficiente até à ocorrência de uma crise espontânea. Assim, foi desenvolvida uma aplicação que permite realizar, futuramente, a análise das alterações ocorridas nos períodos que antecedem crises espontâneas, no entanto, a análise feita neste estudo teve como foco a identificação das crises induzidas por pilocarpina.

Como forma de avaliar o funcionamento da aplicação foram realizados vários testes com os diferentes conjuntos de dados. Inicialmente, os sinais de cada conjunto de dados, isto é, dados de um único animal, foram analisados sem a realização de qualquer tipo de pré-processamento antes da extração de características; depois foram analisados os sinais filtrados através da aplicação um filtro Butterworth passa-banda, com banda de passagem 1 a 49,5 Hz com o objetivo de avaliar altas frequências (> 1 Hz) e desprezar ruído causado pela corrente elétrica (50 Hz). Após a filtragem, foram realizadas duas abordagens diferentes; uma em que foi feita a extração de características após a filtragem e outra, em que foi aplicado o algoritmo *fastICA* entre estes dois passos, obtendo estimativas de sinais independentes.

Foram determinados hiperplanos de separação entre as amostras correspondentes a instantes em que ocorreram as crises e as restantes, através de um método de classificação linear, *Fisher LDA*, e avaliada a capacidade de discriminação destes, quando utilizadas todas as características extraídas.

Como forma de reduzir o número de características utilizadas na classificação e de possibilitar a sua representação e visualização, foi realizada uma análise das mesmas, recorrendo ao teste de Mann-Whitney, com o objetivo de as ordenar de acordo com a sua capacidade de discriminação dos grupos, isto é, quanto maior a diferença na distribuição

das amostras da mesma característica dos diferentes grupos, maior será a sua capacidade de discriminação. Após esta análise foi calculado um hiperplano de separação dos dois grupos definidos (crise e não crise), utilizando as primeiras duas características do *ranking* obtido anteriormente e depois as três primeiras, como forma de observar graficamente a sua distribuição; de seguida usaram-se todas as características e as primeiras N que permitiram obter o desempenho ótimo em cada um dos casos.

Procedendo aos testes de análise de características e de classificação, sem realizar nenhum tipo de pré-processamento, para o conjunto de dados WO2_005, foi possível verificar que os atributos que melhor permitem discriminar os dois grupos definidos são os que correspondem à potência acumulada dos quatro canais e o coeficiente *wavelet* da banda teta dos canais 2, 3 e 4. O número de características que permitiu obter o melhor desempenho para cada um dos casos, foi obtido através do cálculo e análise da acurácia do classificador em função do número de atributos utilizados (figura 6.1). Neste conjunto de dados, foi necessário recorrer a 69 atributos para atingir o desempenho ótimo.

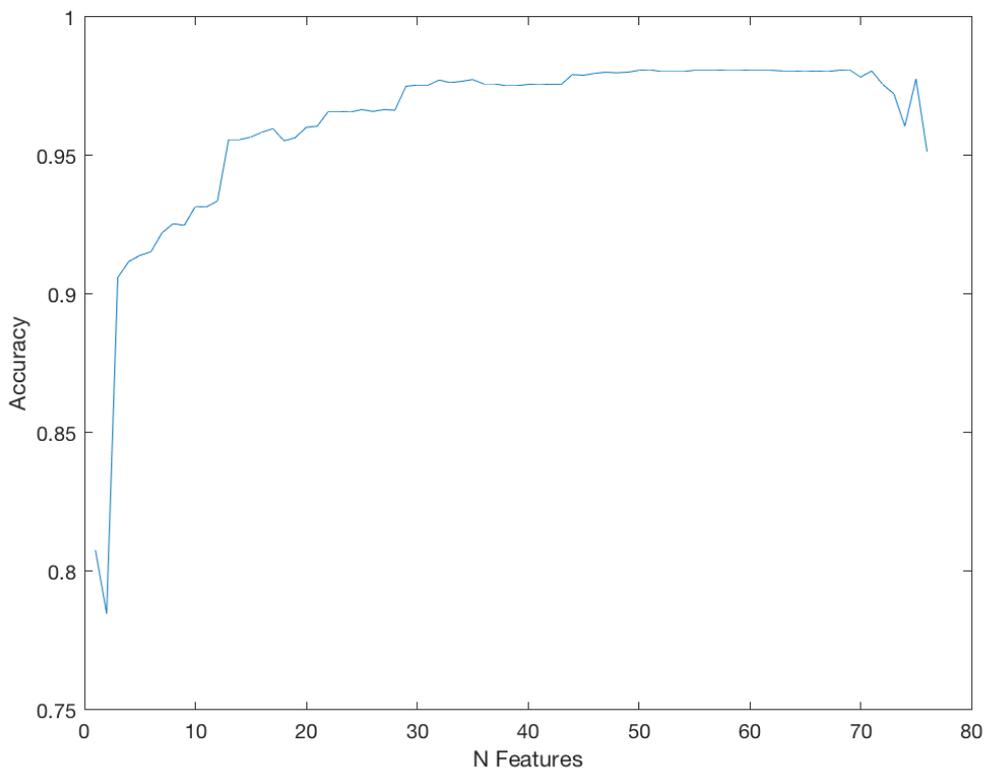


Figura 6.1: Representação da acurácia em função do número de características utilizadas na separação linear através do método *Fisher LDA*, no conjunto de dados WO2_005, sem pré-processamento.

No conjunto de dados WO2_007, os primeiros atributos da lista de características ordenada obtida foram o *kurtosis* e os coeficientes *wavelet* da banda alfa correspondentes aos canais 1, 2 e 3. A classificação ótima foi obtida através da utilização de todas as características.

No conjunto de dados WO2_017, as características com melhor capacidade discriminativa foram a potência relativa das bandas beta e gama dos canais 1 e 2. Utilizando 67 atributos, foi possível atingir uma solução ótima.

No conjunto de dados WO2_020, os atributos com maior capacidade discriminativa foram a potência relativa da banda gama do canal 1 e a potência acumulada correspondente aos quatro canais. O melhor desempenho foi conseguido utilizando 66 características.

Após esta primeira análise, foi aplicado um filtro Butterworth com banda de passagem de 1 a 49,5 Hz antes da extração de características e realizados os testes anteriores. Numa fase posterior, foi aplicado o algoritmo *fastICA* aos dados filtrados e foram realizados os mesmos testes.

No conjunto de dados WO2_005, as características que demonstraram maior capacidade de discriminação foram os coeficientes *wavelet* das bandas beta e alfa correspondentes a todos os canais. Utilizando 67 características atingiu-se a classificação ótima.

Quando introduzida a aplicação do algoritmo *fastICA* na fase de pré-processamento, as características que apresentaram melhor capacidade de discriminação foram as correspondentes à potência acumulada das quatro estimativas de componentes independentes obtidas. A classificação ótima, neste caso, foi alcançada utilizando 73 atributos.

Quando realizados testes anteriores para o conjunto de dados WO2_007, as características com maior capacidade de discriminação foram a variância, o *kurtosis* correspondente ao canal 2 e os coeficientes *wavelet* das bandas beta e alfa correspondentes aos canais 1, 2 e 3. Com a utilização de 71 atributos na classificação, foi registado o melhor desempenho.

Realizando a análise de componentes independentes e posteriormente a análise das características e classificação, as características que demonstraram melhor capacidade de

discriminação dos grupos foram os coeficientes *wavelet* das bandas alfa, beta e teta e a variância correspondentes à primeira estimativa. Para alcançar o desempenho ótimo foram usadas 75 características.

No conjunto de dados WO2_017, as características com maior capacidade discriminativa foram os coeficientes *wavelet* da banda gama de todos os canais. Ao selecionar 74 atributos foi atingida a classificação ótima.

Quando aplicado o algoritmo *fastICA* na fase de pré-processamento, os primeiros atributos na lista ordenada foram os coeficientes *wavelet* da banda gama correspondentes às estimativas 2, 3 e 4 e a potência das estimativas 1, 2 e 3. A classificação ótima implicou a utilização de 70 características.

No conjunto de dados WO2_020, as características com melhor capacidade de discriminação foram as curvas de potência obtidas através do método quantitativo e os coeficientes *wavelet* da banda gama de todos os canais. Para obter um desempenho ótimo foram selecionados 63 atributos.

Aplicando o algoritmo *fastICA* ao conjunto de dados já filtrado, as características que ocuparam os primeiros lugares na lista ordenada foram as curvas de potência obtidas através do método quantitativo e os coeficientes *wavelet* da banda gama correspondentes à primeira estimativa.

Após a realização dos testes com cada um dos conjuntos de dados, foram determinadas as medidas médias de desempenho da separação, de acordo com a utilização de diferentes opções de pré-processamento e apresentadas na tabela 6.1.

6. Resultados e discussão

Tabela 6.1: Medidas médias de desempenho da separação de cada um dos conjuntos de dados.

Dados	Acurácia(%)	Precisão(%)	Sensibilidade(%)	Especificidade(%)
WO2_005				
Sem pré-processamento	90,28	97,29	90,05	91,03
Filtro Butterworth	88,21	92,80	91,83	76,39
Filtro Butterworth + ICA	93,56	96,58	94,94	89,06
WO2_007				
Sem pré-processamento	82,41	78,69	91,35	73,70
Filtro Butterworth	85,15	82,00	91,61	78,86
Filtro Butterworth + ICA	84,28	81,77	89,18	79,49
WO2_017				
Sem pré-processamento	72,46	85,27	73,31	67,96
Filtro Butterworth	94,10	97,00	95,96	84,30
Filtro Butterworth + ICA	93,51	96,92	95,32	83,96
WO2_020				
Sem pré-processamento	88,28	87,30	90,67	85,95
Filtro Butterworth	85,04	80,57	93,11	77,19
Filtro Butterworth + ICA	85,57	80,10	95,02	76,37

Com base na análise dos resultados anteriores, é possível verificar que os sinais recolhidos apresentam diferenças nos momentos em que ocorrem crises e nos que não ocorrem, uma vez que a sua separação nestes dois grupos é possível de realizar com um desempenho médio superior ao de uma classificação aleatória, isto é, com uma acurácia superior a 50%.

Verificou-se que as diferentes opções de pré-processamento introduzem alterações nos valores das medidas de desempenho, apesar de não ser possível identificar um padrão, devido ao número reduzido de dados e ao facto de as alterações ocorridas não serem as mesmas em todos os conjuntos de dados, isto é, em alguns casos a filtragem dos dados favorece o desempenho na separação dos grupos, enquanto que em outros prejudica esta tarefa, o mesmo acontece com a aplicação do algoritmo *fastICA*.

Após esta análise inicial foi ainda testado um classificador baseado na distância euclidiana mínima nas estimativas obtidas através da análise de componentes independentes dos dados previamente filtrados. Nesta análise, à semelhança das anteriores, foram também usadas as N primeiras características que permitem atingir o desempenho ótimo da separação. No conjunto de dados WO2_005, o N obtido foi 3, no conjunto WO2_007 foi 37, no WO2_017 foi 1 e no WO2_020 foi 18. O desempenho médio obtido na realização dos testes com este tipo de classificador foi apresentado na tabela 6.2.

Tabela 6.2: Desempenho médio obtido na classificação baseada na distância euclidiana mínima de cada conjunto de dados após a aplicação do algoritmo *fastICA* na fase de pré-processamento.

Dados	Acurácia(%)	Precisão(%)	Sensibilidade(%)	Especificidade(%)
WO2_005	83,37	94,59	83,02	84,53
WO2_007	81,56	81,20	83,08	80,07
WO2_017	88,36	94,23	91,77	70,40
WO2_020	80,47	73,97	93,26	68,04

Como a utilização de um classificador baseado na distância euclidiana mínima não demonstrou melhorias significativas na separação das classes, nos testes seguintes apenas foi utilizado o método *Fisher LDA*, contudo é possível verificar que a sua utilização é possível na aplicação. O classificador de distância mínima Euclidiano não tem em conta a covariância dos dados e daí a pior performance.

Os testes realizados anteriormente serviram para avaliar a capacidade das características extraídas de conseguirem ou não separar a classe ictal da classe não-ictal. Desta forma usaram-se os dados do mesmo animal para definir o classificador e para avaliar a sua performance. No entanto, do ponto de vista de uma aplicação real queremos

avaliar o comportamento de um classificador treinado em dados completamente novos. Desta forma decidiu-se construir quatro conjuntos de dados, seguindo uma estratégia *live-one-out*, isto é dados de três animais foram considerados para treino e um deles para teste, como se mostra na tabela 6.3. Os resultados médios obtidos nestes testes foram apresentados na tabela 6.4.

Tabela 6.3: Descrição dos conjuntos usados para treinar e testar classificadores lineares *Fisher LDA*.

	Treino	Teste
Conjunto 1	WO2_005, WO2_007, WO2_017	WO2_020
Conjunto 2	WO2_005, WO2_007, WO2_020	WO2_017
Conjunto 3	WO2_005, WO2_017, WO2_020	WO2_007
Conjunto 4	WO2_007, WO2_017, WO2_020	WO2_005

Para o conjunto 1, quando realizada a extração de características sem pré-processamento, os atributos que apresentaram melhor capacidade de discriminação foram os coeficientes *wavelet* da banda beta do canal 2 e da banda alfa do canal 3 e os *kurtosis* dos canais 1 e 3. A separação ótima foi encontrada através da utilização de 71 características.

No conjunto 2, as características que se destacaram foram os coeficientes *wavelet* das bandas alfa e teta dos canais 1, 2 e 3, tendo sido atingido o desempenho ótimo quando foram selecionados 72 atributos.

As características que apresentaram melhor capacidade de separação no conjunto 3 foram a e a potência acumulada dos quatro canais e os coeficientes *wavelet* da banda alfa do canal 3. O melhor desempenho foi obtido usando as 76 características.

No conjunto 4, as características que mais se destacaram foram os coeficientes *wavelet* das bandas alfa e beta e o *kurtosis* correspondentes ao canal 3. O desempenho ótimo foi atingido usando 71 atributos.

Quando realizada a filtragem dos sinais, as características que mostraram melhor capacidade discriminativa no conjunto 1 foram os coeficientes *wavelet* da banda beta dos canais 2 e 3 e o *kurtosis* correspondente ao canal 2. O desempenho ótimo na fase de treino foi alcançado utilizando os atributos todos.

No conjunto 2, as características com melhor capacidade discriminativa foram as curvas relativas à potência obtida através do modelo quantitativo correspondentes aos canais 1,2 e 3 e o coeficiente *wavelet* da banda alfa do canal 1. O desempenho ótimo foi obtido através da utilização de 74 atributos.

No conjunto 3, os atributos que permitiram separar melhor as duas classes foram a potência e sua curva correspondente obtidas através do método quantitativo e os coeficientes *wavelet* da banda gama correspondentes ao canal 3. A separação ótima foi obtida usando 74 características.

As características que se destacaram no conjunto 4 foram a potência e a curva associada calculadas através do modelo quantitativo e os coeficientes *wavelet* das bandas gama e beta correspondentes ao canal 3. O desempenho ótimo foi atingido utilizando 73 atributos.

Após a filtragem dos sinais e a aplicação do algoritmo *fastICA*, no conjunto 1, as características que apresentaram melhor capacidade de separação das classes foram os coeficientes *wavelet* da banda beta correspondentes a todas as estimativas e o *kurtosis* correspondente à estimativa 2.

No conjunto 2, as características com melhor capacidade discriminativa foram a curva relativa à potência, os coeficientes *wavelet* da banda alfa e a variância correspondentes ao canal 3. O desempenho ótimo foi obtido usando 74 atributos.

Os atributos com melhor capacidade discriminativa, no conjunto 3, foram a potência acumulada de todos os canais. Para obter um desempenho ótimo, foi necessário utilizar as características todas.

No conjunto 4, os atributos que mais se destacaram foram a potência e a curva associada calculadas através do modelo quantitativo e os coeficientes *wavelet* das bandas gama, beta e alfa correspondentes ao canal 2. Utilizando 71 características foi possível obter uma separação ótima.

Através da análise da capacidade discriminativa das características para os dados de forma individual e para os conjuntos 1, 2, 3 e 4 foi possível observar que, na maior parte dos casos, os coeficientes *wavelet* se encontram no topo do *ranking* das características que melhor permitem separar as classes selecionadas.

6. Resultados e discussão

Tabela 6.4: Medidas médias de desempenho obtidas nas fases de treino e teste de classificadores lineares *Fisher LDA*, para cada um dos conjuntos de dados utilizando diferentes opções de pré-processamento.

			Acurácia (%)	Precisão (%)	Sensibilidade (%)	Especificidade (%)
Conjunto 1	Sem pré-processamento	Treino	77,70	86,02	81,39	69,27
		Teste	46,49	42,97	57,23	36,04
	Filtro Butterworth	Treino	65,50	74,12	62,25	72,95
		Teste	70,02	64,14	91,28	49,34
	Filtro Butterworth + ICA	Treino	82,13	84,67	91,19	61,40
		Teste	61,68	57,82	83,08	40,88
Conjunto 2	Sem pré-processamento	Treino	80,44	78,67	90,11	69,03
		Teste	43,78	75,59	46,70	28,98
	Filtro Butterworth	Treino	80,05	82,06	81,49	78,34
		Teste	77,65	95,35	77,15	80,29
	Filtro Butterworth + ICA	Treino	82,20	82,36	85,77	77,98
		Teste	75,51	96,07	73,90	83,99
Conjunto 3	Sem pré-processamento	Treino	80,00	86,90	83,58	71,81
		Teste	36,66	27,69	29,84	23,98
	Filtro Butterworth	Treino	82,29	88,89	85,40	75,14
		Teste	53,81	52,00	56,44	51,25
	Filtro Butterworth + ICA	Treino	80,15	86,71	84,61	69,94
		Teste	36,32	40,10	61,11	11,82
Conjunto 4	Sem pré-processamento	Treino	78,57	82,28	84,35	69,02
		Teste	75,25	83,24	85,40	42,14
	Filtro Butterworth	Treino	73,66	79,78	76,95	68,22
		Teste	66,22	87,52	64,82	70,76
	Filtro Butterworth + ICA	Treino	77,93	83,62	81,18	72,57
		Teste	63,89	81,61	68,28	49,57

Nos testes realizados com os conjuntos 1, 2, 3 e 4 definidos na tabela 6.3, foi possível verificar uma diminuição do desempenho de separação na fase de teste em relação à de treino na maior parte dos casos, como era de esperar. No entanto, nos testes realizados com o conjunto 2 foi possível verificar um aumento do desempenho da classificação quando utilizadas apenas 2 ou 3 características obtidas após a filtragem dos dados, o que pode ser verificado nas tabelas A.22 e A.23.

O melhor desempenho foi obtido com os dados filtrados do conjunto 2, utilizando 3 características no treino e teste do classificador, alcançando, na fase de teste, uma sensibilidade e especificidade de 80,77% e 81,54%, respetivamente.

Analisando o desempenho médio dos classificadores nestes testes apresentado na tabela 6.5, é possível verificar que, apesar de o desempenho estar muito distante do pretendido, que todas as medidas médias são superiores às obtidas teoricamente com uma classificação aleatória, isto é, superiores a 50%.

Tabela 6.5: Desempenho médio obtido no treino e teste de classificadores lineares Fisher LDA utilizando os conjuntos 1, 2, 3 e 4

	Acurácia (%)	Precisão	Sensibilidade	Especificidade
Treino	78,51	83,00	82,53	71,52
Teste	59,65	69,33	66,41	50,04

Os testes realizados com os conjuntos 1, 2, 3 e 4 demonstraram a possibilidade de treinar classificadores, guardá-los e testá-los posteriormente noutros conjuntos de dados, contudo os resultados obtidos revelaram que os classificadores testados não apresentam um bom desempenho, o que se deve à dimensão reduzida do conjunto de treino, pois torna o classificador muito ajustado aos casos com que é treinado, apresentando assim uma capacidade de generalização reduzida.

Apesar dos resultados obtidos nesta última fase não serem satisfatórios, é possível verificar que a aplicação desenvolvida permite desempenhar todas as funções pretendidas corretamente. É importante referir que o desempenho dos classificadores poderia ser melhorado com o aumento da quantidade de dados disponíveis.

É de frisar que o conjunto de dados WO2_024 não foi utilizado na fase de testes, devido ao facto de não apresentar identificação do período em que ocorreu a crise induzida. No entanto, contribuiu para o desenvolvimento da aplicação, visto que esta foi

desenhada de forma a permitir a análise e processamento de dados com as mesmas condições de aquisição do conjunto em questão (frequência de amostragem igual a 200 Hz e informação relativa aos acontecimentos assinalados durante a aquisição gravados em formato *.csv*), uma vez que os restantes conjuntos têm frequência de amostragem igual a 100 Hz e as anotações gravadas em formato *.xml*.

7. Conclusão e Trabalho Futuro

Neste estudo, é possível verificar a capacidade da aplicação desenvolvida para processar e analisar sinais recolhidos através dos *arrays* de microelétrodos de platina utilizados na medição de oxigénio *in vivo*. A aplicação foi desenvolvida para colmatar uma necessidade de um grupo de investigação do CNC. Com as funcionalidades implementadas e os resultados obtidos é de prever que esta aplicação se torne numa ferramenta útil no estudo da epilepsia envolvendo análise de sinais de pO_2 , uma vez que permite utilizar ferramentas de análise e processamento de sinal de forma simples e intuitiva, mesmo não tendo muito conhecimento na área.

Apesar do estudo se basear na deteção de crises, é importante perceber a capacidade que a aplicação tem para se estender à previsão de crises e a problemas envolvendo várias classes, ou seja, à identificação dos períodos pré-crise ou de outros estados epiléticos, uma vez que para isso só é necessário alterar o *target* definido no separador correspondente à análise das características extraídas.

Como trabalho futuro, sugere-se o aumento do conjunto de dados e posterior análise destes, como forma de identificar padrões genéricos que permitam fazer uma boa separação dos grupos pretendidos. Propõe-se ainda que sejam estudadas e integradas na aplicação outros tipos de extração de características e classificação, uma vez que existe uma vasta gama de características e de classificadores que podem ser estudados.

Bibliografia

- [1] WHO, “WHO | Epilepsy,” 2018. [Online]. Available: <http://www.who.int/en/news-room/fact-sheets/detail/epilepsy>. [Accessed: 28-Jun-2018].
- [2] P. Kwan and M. J. Brodie, *Early Identification of Refractory Epilepsy*, vol. 342. 2000.
- [3] J. F. Tellez-Zenteno, L. Hernandez-Ronquillo, S. Buckley, R. Zahagun, and S. Rizvi, “A validation of the new definition of drug-resistant epilepsy by the International League Against Epilepsy.,” *Epilepsia*, vol. 55, no. 6, pp. 829–834, Jun. 2014.
- [4] J. A. French *et al.*, “Characteristics of medial temporal lobe epilepsy: I. Results of history and physical examination.,” *Ann. Neurol.*, vol. 34, no. 6, pp. 774–780, Dec. 1993.
- [5] A. Gustavsson *et al.*, “Cost of disorders of the brain in Europe 2010.,” *Eur. Neuropsychopharmacol.*, vol. 21, no. 10, pp. 718–779, Oct. 2011.
- [6] F. Panzica, G. Varotto, F. Rotondi, R. Spreafico, and S. Franceschetti, “Identification of the Epileptogenic Zone from Stereo-EEG Signals: A Connectivity-Graph Theory Approach.,” *Front. Neurol.*, vol. 4, p. 175, Nov. 2013.
- [7] M. Suh, H. Ma, M. Zhao, S. Sharif, and T. H. Schwartz, “Neurovascular coupling and oximetry during epileptic events.,” *Mol. Neurobiol.*, vol. 33, no. 3, pp. 181–197, Jun. 2006.
- [8] M. Zhao, M. Suh, H. Ma, C. Perry, A. Geneslaw, and T. H. Schwartz, “Focal increases in perfusion and decreases in hemoglobin oxygenation precede seizure onset in spontaneous human epilepsy.,” *Epilepsia*, vol. 48, no. 11, pp. 2059–2067, Nov. 2007.
- [9] S. Bahar, M. Suh, M. Zhao, and T. H. Schwartz, “Intrinsic optical signal imaging of neocortical seizures: the ‘epileptic dip’ .,” *Neuroreport*, vol. 17, no. 5, pp. 499–503, Apr. 2006.
- [10] I. E. Scheffer *et al.*, “ILAE classification of the epilepsies: Position paper of the

- ILAE Commission for Classification and Terminology.,” *Epilepsia*, vol. 58, no. 4, pp. 512–521, Apr. 2017.
- [11] J. Pimentel and C. Bentes, *Epilepsia: conceitos, diagnóstico e tratamento*. Lidel, 2016.
- [12] P. Kwan *et al.*, “Definition of drug resistant epilepsy: consensus proposal by the ad hoc Task Force of the ILAE Commission on Therapeutic Strategies.,” *Epilepsia*, vol. 51, no. 6, pp. 1069–1077, Jun. 2010.
- [13] A. Pitkanen and K. Lukasiuk, “Mechanisms of epileptogenesis and potential treatment targets.,” *Lancet. Neurol.*, vol. 10, no. 2, pp. 173–186, Feb. 2011.
- [14] S. Lodish, H. Berk, and A. Zipursky, “Molecular Cell Biology. 4th edition.,” New York: W. H. Freeman, 2000.
- [15] R. S. Sloviter, “The neurobiology of temporal lobe epilepsy: too much information, not enough knowledge.,” *C. R. Biol.*, vol. 328, no. 2, pp. 143–153, Feb. 2005.
- [16] G. Curia, D. Longo, G. Biagini, R. S. G. Jones, and M. Avoli, “The pilocarpine model of temporal lobe epilepsy,” *J. Neurosci. Methods*, vol. 172, no. 2–4, pp. 143–157, Jul. 2008.
- [17] J. P. Leite, N. Garcia-Cairasco, and E. A. Cavalheiro, “New insights from the use of pilocarpine and kainate models.,” *Epilepsy Res.*, vol. 50, no. 1–2, pp. 93–103, Jun. 2002.
- [18] M. Zhao, H. Ma, M. Suh, and T. H. Schwartz, “Spatiotemporal dynamics of perfusion and oximetry during ictal discharges in the rat neocortex.,” *J. Neurosci.*, vol. 29, no. 9, pp. 2814–2823, Mar. 2009.
- [19] R. S. Fisher and J. J. Engel, “Definition of the postictal state: When does it start and end?,” *Epilepsy Behav.*, vol. 19, no. 2, pp. 100–104, 2010.
- [20] J. J. Burmeister *et al.*, “Ceramic-based multisite microelectrode arrays for simultaneous measures of choline and acetylcholine in CNS.,” *Biosens. Bioelectron.*, vol. 23, no. 9, pp. 1382–1389, Apr. 2008.
- [21] V. Parikh, F. Pomerleau, P. Huettl, G. A. Gerhardt, M. Sarter, and J. P. Bruno, “Rapid assessment of in vivo cholinergic transmission by amperometric detection

- of changes in extracellular choline levels,” *Eur. J. Neurosci.*, vol. 20, no. 6, pp. 1545–1554, Sep. 2004.
- [22] A. Ledo, C. F. Lourenço, J. Laranjinha, G. A. Gerhardt, and R. M. Barbosa, “Combined in vivo amperometric oximetry and electrophysiology in a single sensor: A tool for epilepsy research,” *Anal. Chem.*, vol. 89, no. 21, pp. 12383–12390, 2017.
- [23] S. S. Viglione and G. O. Walsh, “Proceedings: Epileptic seizure prediction,” *Electroencephalogr. Clin. Neurophysiol.*, vol. 39, no. 4, pp. 435–436, Oct. 1975.
- [24] Z. Rogowski, I. Gath, and E. Bental, “On the prediction of epileptic seizures,” *Biol. Cybern.*, vol. 42, no. 1, pp. 9–15, 1981.
- [25] Y. Salant, I. Gath, and O. Henriksen, “Prediction of epileptic seizures from two-channel EEG,” *Med. Biol. Eng. Comput.*, vol. 36, no. 5, pp. 549–556, 1998.
- [26] A. Siegel, C. L. Grady, and A. F. Mirsky, “Prediction of Spike-Wave Bursts in Absence Epilepsy by EEG Power-Spectrum Signals,” *Epilepsia*, vol. 23, no. 1, pp. 47–60, 1982.
- [27] J. Gotman and M. G. Marciani, “Electroencephalographic spiking activity, drug levels, and seizure occurrence in epileptic patients,” *Ann. Neurol.*, vol. 17, no. 6, pp. 597–603, 1985.
- [28] J. Gotman and D. J. Koffler, “Interictal spiking increases after seizures but does not after decrease in medication,” *Electroencephalogr. Clin. Neurophysiol.*, vol. 72, no. 1, pp. 7–15, 1989.
- [29] A. Katz, D. A. Marks, G. McCarthy, and S. S. Spencer, “Does interictal spiking change prior to seizures?,” *Electroencephalogr. Clin. Neurophysiol.*, vol. 79, no. 2, pp. 153–156, 1991.
- [30] L. D. Iasemidis, J. Chris Sackellares, H. P. Zaveri, and W. J. Williams, “Phase space topography and the Lyapunov exponent of electrocorticograms in partial seizures,” *Brain Topogr.*, vol. 2, no. 3, pp. 187–201, Mar. 1990.
- [31] M. Le Van Quyen, C. Adam, J. Martinerie, M. Baulac, S. Clémenceau, and F. Varela, “Spatio-temporal characterizations of non-linear changes in intracranial activities prior to human temporal lobe seizures,” *Eur. J. Neurosci.*, vol. 12, no. 6,

- pp. 2124–2134, 2000.
- [32] M. Le Van Quyen *et al.*, “Anticipation of epileptic seizures from standard EEG recordings,” *Lancet*, vol. 357, no. 9251, pp. 183–188, 2001.
- [33] M. E. Weinand, L. P. Carter, W. F. El-Saadany, P. J. Sioutos, D. M. Labiner, and K. J. Oommen, “Cerebral blood flow and temporal lobe epileptogenicity,” *J. Neurosurg.*, vol. 86, no. 2, pp. 226–232, Feb. 1997.
- [34] C. Baumgartner *et al.*, “Preictal SPECT in temporal lobe epilepsy: regional cerebral blood flow is increased prior to electroencephalography-seizure onset.,” *J. Nucl. Med.*, vol. 39, no. 6, pp. 978–982, Jun. 1998.
- [35] P. D. Adelson, E. Nemoto, M. Scheuer, M. Painter, J. Morgan, and H. Yonas, “Noninvasive continuous monitoring of cerebral oxygenation periictally using near-infrared spectroscopy: a preliminary report.,” *Epilepsia*, vol. 40, no. 11, pp. 1484–1489, Nov. 1999.
- [36] P. Federico, D. F. Abbott, R. S. Briellmann, A. S. Harvey, and G. D. Jackson, “Functional MRI of the pre-ictal state.,” *Brain*, vol. 128, no. Pt 8, pp. 1811–1817, Aug. 2005.
- [37] R. S. Delamont, P. O. Julu, and G. A. Jamal, “Changes in a measure of cardiac vagal activity before and after epileptic seizures.,” *Epilepsy Res.*, vol. 35, no. 2, pp. 87–94, Jun. 1999.
- [38] V. Novak, A. L. Reeves, P. Novak, P. A. Low, and F. W. Sharbrough, “Time-frequency mapping of R-R interval during complex partial seizures of temporal lobe origin.,” *J. Auton. Nerv. Syst.*, vol. 77, no. 2–3, pp. 195–202, Sep. 1999.
- [39] D. H. Kerem and A. B. Geva, “Forecasting epilepsy from the heart rate signal.,” *Med. Biol. Eng. Comput.*, vol. 43, no. 2, pp. 230–239, Mar. 2005.
- [40] C. E. Elger and K. Lehnertz, “Seizure prediction by non-linear time series analysis of brain electrical activity,” *Eur. J. Neurosci.*, vol. 10, no. 2, pp. 786–789, 1998.
- [41] K. Lehnertz and C. E. Elger, “Can Epileptic Seizures be Predicted? Evidence from Nonlinear Time Series Analysis of Brain Electrical Activity,” *Phys. Rev. Lett.*, vol. 80, no. 22, pp. 5019–5022, Jun. 1998.

- [42] R. Aschenbrenner-Scheibe, T. Maiwald, M. Winterhalder, H. U. Voss, J. Timmer, and A. Schulze-Bonhage, “How well can epileptic seizures be predicted? An evaluation of a nonlinear method,” *Brain*, vol. 126, no. 12, pp. 2616–2626, 2003.
- [43] M. A. F. Harrison, I. Osorio, M. G. Frei, S. Asuri, and Y. C. Lai, “Correlation dimension and integral do not predict epileptic seizures,” *Chaos*, vol. 15, no. 3, pp. 1–15, 2005.
- [44] F. Mormann, K. Lehnertz, P. David, and C. E. Elger, “Mean phase coherence as a measure for phase synchronization and its application to the EEG of epilepsy patients,” *Phys. D Nonlinear Phenom.*, vol. 144, no. 3, pp. 358–369, 2000.
- [45] F. Mormann *et al.*, “Automated detection of a pre seizure state based on a decrease in synchronization in intracranial electroencephalogram recordings from epilepsy patients,” *Phys. Rev. E*, vol. 67, no. 2, p. 21912, 2003.
- [46] F. Mormann, T. Kreuz, R. G. Andrzejak, P. David, K. Lehnertz, and C. E. Elger, “Epileptic seizures are preceded by a decrease in synchronization,” *Epilepsy Res.*, vol. 53, no. 3, pp. 173–185, 2003.
- [47] R. Cerf and el H. el Ouasdad, “Spectral analysis of stereo-electroencephalograms: preictal slowing in partial epilepsies,” *Biol. Cybern.*, vol. 83, no. 5, pp. 399–405, 2000.
- [48] L. D. Iasemidis, P. Pardalos, J. C. Sackellares, and D.-S. Shiau, “Quadratic Binary Programming and Dynamical System Approach to Determine the Predictability of Epileptic Seizures,” *J. Comb. Optim.*, vol. 5, no. 1, pp. 9–26, Mar. 2001.
- [49] B. Litt *et al.*, “Epileptic Seizures May Begin Hours in Advance of Clinical Onset: A Report of Five Patients,” *Neuron*, vol. 30, no. 1, pp. 51–64, 2001.
- [50] S. Gigola, F. Ortiz, C. E. D’Attellis, W. Silva, and S. Kochen, “Prediction of epileptic seizures using accumulated energy in a multiresolution framework,” *J. Neurosci. Methods*, vol. 138, no. 1, pp. 107–111, 2004.
- [51] K. Schindler, R. Wiest, M. Kollar, and F. Donati, “EEG analysis with simulated neuronal cell models helps to detect pre-seizure changes,” *Clin. Neurophysiol.*, vol. 113, no. 4, pp. 604–614, 2002.
- [52] F. Mormann, R. G. Andrzejak, C. E. Elger, and K. Lehnertz, “Seizure prediction:

- The long and winding road,” *Brain*, vol. 130, no. 2, pp. 314–333, 2007.
- [53] K. Gadhoumi, J. M. Lina, F. Mormann, and J. Gotman, “Seizure prediction for therapeutic devices: A review,” *J. Neurosci. Methods*, vol. 260, no. 29, pp. 270–282, 2016.
- [54] K. Lehnertz and B. Litt, “The First International Collaborative Workshop on Seizure Prediction: summary and data description,” *Clin. Neurophysiol.*, vol. 116, no. 3, pp. 493–505, 2005.
- [55] M. D’Alessandro *et al.*, “A multi-feature and multi-channel univariate selection process for seizure prediction,” *Clin. Neurophysiol.*, vol. 116, no. 3, pp. 506–516, Mar. 2005.
- [56] R. Esteller *et al.*, “Continuous energy variation during the seizure cycle: towards an on-line accumulated energy,” *Clin. Neurophysiol.*, vol. 116, no. 3, pp. 517–526, 2005.
- [57] M. A. F. Harrison, M. G. Frei, and I. Osorio, “Accumulated energy revisited,” *Clin. Neurophysiol.*, vol. 116, no. 3, pp. 527–531, 2005.
- [58] C. C. Jouny, P. J. Franaszczuk, and G. K. Bergey, “Signal complexity and synchrony of epileptic seizures: is there an identifiable preictal period?,” *Clin. Neurophysiol.*, vol. 116, no. 3, pp. 552–558, 2005.
- [59] F. Mormann *et al.*, “On the predictability of epileptic seizures,” *Clin. Neurophysiol.*, vol. 116, no. 3, pp. 569–587, 2005.
- [60] L. D. Iasemidis *et al.*, “Long-term prospective on-line real-time seizure prediction,” *Clin. Neurophysiol.*, vol. 116, no. 3, pp. 532–544, 2005.
- [61] M. Le Van Quyen *et al.*, “Preictal state identification by synchronization changes in long-term intracranial EEG recordings,” *Clin. Neurophysiol.*, vol. 116, no. 3, pp. 559–568, 2005.
- [62] B. Schelter *et al.*, “Testing statistical significance of multivariate time series analysis techniques for epileptic seizure prediction,” *Chaos An Interdiscip. J. Nonlinear Sci.*, vol. 16, no. 1, p. 13108, Jan. 2006.
- [63] J. C. Sackellares *et al.*, “Predictability analysis for an automated seizure prediction

- algorithm.,” *J. Clin. Neurophysiol.*, vol. 23, no. 6, pp. 509–20, 2006.
- [64] P. Mirowski, D. Madhavan, Y. LeCun, and R. Kuzniecky, “Classification of patterns of EEG synchronization for seizure prediction,” *Clin. Neurophysiol.*, vol. 120, no. 11, pp. 1927–1940, 2009.
- [65] L. Kuhlmann *et al.*, “Patient-specific bivariate-synchrony-based seizure prediction for short prediction horizons,” *Epilepsy Res.*, vol. 91, no. 2, pp. 214–231, 2010.
- [66] C. A. Teixeira *et al.*, “EPILAB: A software package for studies on the prediction of epileptic seizures,” *J. Neurosci. Methods*, vol. 200, no. 2, pp. 257–271, 2011.
- [67] Y. Park, L. Luo, K. K. Parhi, and T. Netoff, “Seizure prediction with spectral power of EEG using cost-sensitive support vector machines,” *Epilepsia*, vol. 52, no. 10, pp. 1761–1770, 2011.
- [68] A. Aarabi and B. He, “A rule-based seizure prediction method for focal neocortical epilepsy,” *Clin. Neurophysiol.*, vol. 123, no. 6, pp. 1111–1122, 2012.
- [69] J. R. Williamson, D. W. Bliss, D. W. Browne, and J. T. Narayanan, “Seizure prediction using EEG spatiotemporal correlation structure,” *Epilepsy Behav.*, vol. 25, no. 2, pp. 230–238, 2012.
- [70] M. J. Cook *et al.*, “Prediction of seizure likelihood with a long-term, implanted seizure advisory system in patients with drug-resistant epilepsy: a first-in-man study,” *Lancet Neurol.*, vol. 12, no. 6, pp. 563–571, 2013.
- [71] K. Gadhomi, J.-M. Lina, and J. Gotman, “Seizure prediction in patients with mesial temporal lobe epilepsy using EEG measures of state similarity,” *Clin. Neurophysiol.*, vol. 124, no. 9, pp. 1745–1754, 2013.
- [72] S. Li, W. Zhou, Q. Yuan, and Y. Liu, “Seizure Prediction Using Spike Rate of Intracranial EEG,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 21, no. 6, pp. 880–886, 2013.
- [73] J. Rasekhi, M. Reza, K. Mollaei, M. Bandarabadi, C. A. Teixeira, and A. Dourado, “Preprocessing effects of 22 linear univariate features on the performance of seizure prediction methods,” *J. Neurosci. Methods*, vol. 217, no. 1–2, pp. 9–16, 2013.

- [74] Y. Zheng, G. Wang, K. Li, G. Bao, and J. Wang, “Epileptic seizure prediction using phase synchronization based on bivariate empirical mode decomposition,” *Clin. Neurophysiol.*, vol. 125, no. 6, pp. 1104–1111, 2014.
- [75] A. Aarabi and B. He, “Seizure prediction in hippocampal and neocortical epilepsy using a model-based approach,” *Clin. Neurophysiol.*, vol. 125, no. 5, pp. 930–940, 2014.
- [76] S. Banerjee and T. Pedersen, “The Design, Implementation, and Use of the Ngram Statistics Package BT - Computational Linguistics and Intelligent Text Processing,” 2003, pp. 370–381.
- [77] A. Eftekhar, W. Juffali, J. El-Imad, T. G. Constandinou, and C. Toumazou, “Ngram-Derived Pattern Recognition for the Detection and Prediction of Epileptic Seizures,” *PLoS One*, vol. 9, no. 6, pp. 1–15, 2014.
- [78] R. P. Costa, P. Oliveira, G. Rodrigues, B. Leitão, and A. Dourado, “Epileptic Seizure Classification Using Neural Networks with 14 Features BT - Knowledge-Based Intelligent Information and Engineering Systems,” 2008, pp. 281–288.
- [79] N. Moghim and D. W. Corne, “Predicting Epileptic Seizures in Advance,” *PLoS One*, vol. 9, no. 6, p. e99334, Jun. 2014.
- [80] C. Alexandre Teixeira *et al.*, “Epileptic seizure predictors based on computational intelligence techniques: A comparative study with 278 patients,” *Comput. Methods Programs Biomed.*, vol. 114, no. 3, pp. 324–336, 2014.
- [81] M. Bandarabadi, C. A. Teixeira, J. Rasekhi, and A. Dourado, “Epileptic seizure prediction using relative spectral power features,” *Clin. Neurophysiol.*, vol. 126, no. 2, pp. 237–248, 2015.
- [82] B. Direito, C. A. Teixeira, F. Sales, M. Castelo-Branco, and A. Dourado, “A Realistic Seizure Prediction Study Based on Multiclass SVM,” *Int. J. Neural Syst.*, vol. 27, no. 3, p. 1750006, 2017.
- [83] L. Guo, D. Rivero, and A. Pazos, “Epileptic seizure detection using multiwavelet transform based approximate entropy and artificial neural networks,” *J. Neurosci. Methods*, vol. 193, no. 1, pp. 156–163, 2010.
- [84] S. Ghosh-Dastidar, H. Adeli, and N. Dadmehr, “Mixed-Band Wavelet-Chaos-

- Neural Network Methodology for Epilepsy and Epileptic Seizure Detection,” *IEEE Trans. Biomed. Eng.*, vol. 54, no. 9, pp. 1545–1551, 2007.
- [85] H. Ocak, “Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy,” *Expert Syst. Appl.*, vol. 36, no. 2, Part 1, pp. 2027–2036, 2009.
- [86] A. T. Tzallas, M. G. Tsipouras, and D. I. Fotiadis, “Epileptic Seizure Detection in EEGs Using Time-Frequency Analysis,” *IEEE Trans. Inf. Technol. Biomed.*, vol. 13, no. 5, pp. 703–710, 2009.
- [87] Y. Liu, W. Zhou, Q. Yuan, and S. Chen, “Automatic Seizure Detection Using Wavelet Transform and SVM in Long-Term Intracranial EEG,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 20, no. 6, pp. 749–755, 2012.
- [88] Y. Kumar, M. L. Dewal, and R. S. Anand, “Epileptic seizure detection using DWT based fuzzy approximate entropy and support vector machine,” *Neurocomputing*, vol. 133, pp. 271–279, 2014.
- [89] M. Bandarabadi, C. A. Teixeira, and A. Dourado, “Early Seizure Detection Using Neuronal Potential Similarity: A Generalized Low-Complexity and Robust Measure,” vol. 25, no. 5, 2015.
- [90] M. J. Lehmkuhle *et al.*, “A Simple Quantitative Method for Analyzing Electrographic Status Epilepticus in Rats A Simple Quantitative Method for Analyzing Electrographic Status Epilepticus in Rats,” no. January 2009, pp. 1660–1670.
- [91] A. Hyvärinen and E. Oja, “Independent Component Analysis: Algorithms and Applications,” vol. 1, no. 1, 2000.
- [92] A. Hyvärinen, “Fast and Robust Fixed-Point Algorithms for Independent Component Analysis,” pp. 1–16, 1999.
- [93] B. Hjorth, “EEG analysis based on time domain properties.,” *Electroencephalogr. Clin. Neurophysiol.*, vol. 29, no. 3, pp. 306–310, 1970.
- [94] J. P. M. de Sa, “Pattern Recognition: Concepts, Methods, and Applications,” *Springer-Verlag Berlin Heidelb. GmbH*, 2001.

A. Medidas de desempenho obtidas nos diferentes testes

Tabela A.1: Desempenho obtido na classificação linear dos dados WO2_005 utilizando 2, 3, 69 e 76 características extraídas sem pré-processamento do sinal original.

Número de características	2	3	69	76
Acurácia (%)	78,49	90,60	98,10	93,94
Precisão (%)	100	92,59	98,42	98,13
Sensibilidade (%)	71,89	95,35	99,10	93,87
Especificidade (%)	100	75,14	94,82	94,16

Tabela A.2: Desempenho obtido na classificação linear dos dados WO2_007 utilizando 2, 3 e 76 características extraídas sem pré-processamento do sinal original.

Número de características	2	3	76
Acurácia (%)	73,70	74,47	99,07
Precisão (%)	67,80	69,68	98,58
Sensibilidade (%)	88,99	85,49	99,56
Especificidade (%)	58,78	63,72	98,61

Tabela A.3: Desempenho obtido na classificação linear dos dados WO2_017 utilizando 2, 3, 67 e 76 características extraídas sem pré-processamento do sinal original.

Número de características	2	3	67	76
Acurácia (%)	85,88	90,18	98,23	15,53
Precisão (%)	95,75	97,57	98,55	49,21
Sensibilidade (%)	87,06	90,58	99,35	16,25
Especificidade (%)	79,68	88,11	92,32	11,74

Tabela A.4: Desempenho obtido na classificação linear dos dados WO2_020 utilizando 2, 3, 66 e 76 características extraídas sem pré-processamento do sinal original.

Número de características	2	3	66	76
Acurácia (%)	83,77	92,05	98,53	78,77
Precisão (%)	89,16	90,87	97,11	72,04
Sensibilidade (%)	76,36	93,23	100	93,08
Especificidade (%)	90,97	90,89	97,10	64,85

Tabela A.5: Desempenho obtido na classificação linear dos dados WO2_005 filtrados, utilizando 2, 3, 67 e 76 características extraídas.

Número de características	2	3	67	76
Acurácia (%)	88,28	79,46	97,97	87,11
Precisão (%)	90,65	86,11	99,56	94,88
Sensibilidade (%)	94,42	87,23	97,78	87,88
Especificidade (%)	68,27	54,14	98,59	84,56

Tabela A.6: Desempenho obtido na classificação linear dos dados WO2_005 filtrados, utilizando 2, 3, 73 e 76 características extraídas, após a aplicação do algoritmo *fastICA* na fase de pré-processamento.

Número de características	2	3	73	76
Acurácia (%)	88,81	92,53	96,51	96,37
Precisão (%)	93,73	95,83	98,42	98,33
Sensibilidade (%)	91,51	94,34	97,00	96,91
Especificidade (%)	80,04	86,63	94,92	94,63

Tabela A.7: Desempenho obtido na classificação linear dos dados WO2_007 filtrados utilizando 2, 3, 71 e 76 características extraídas.

Número de características	2	3	71	76
Acurácia (%)	79,51	79,48	91,59	90,03
Precisão (%)	72,96	72,93	91,74	90,35
Sensibilidade (%)	92,97	92,95	91,17	89,36
Especificidade (%)	66,39	66,35	91,99	90,69

Tabela A.8: Desempenho obtido na classificação linear dos dados WO2_007 filtrados utilizando 2, 3, 74 e 76 características extraídas, após a aplicação do algoritmo *fastICA* na fase de pré-processamento.

Número de características	2	3	74	76
Acurácia (%)	78,15	76,42	96,25	86,28
Precisão (%)	73,57	71,60	96,81	85,08
Sensibilidade (%)	87,00	86,61	95,56	87,56
Especificidade (%)	69,52	66,48	96,93	85,02

Tabela A.9: Desempenho obtido na classificação linear dos dados WO2_017 filtrados utilizando 2, 3, 74 e 76 características extraídas.

Número de características	2	3	74	76
Acurácia (%)	90,14	90,38	98,06	97,80
Precisão (%)	94,36	94,34	99,72	99,59
Sensibilidade (%)	93,88	94,20	97,97	97,79
Especificidade (%)	70,50	70,25	98,54	97,89

Tabela A.10: Desempenho obtido na classificação linear dos dados WO2_017 filtrados utilizando 2, 3, 70 e 76 características extraídas, após a aplicação do algoritmo *fastICA* na fase de pré-processamento.

Número de características	2	3	70	76
Acurácia (%)	89,37	89,49	97,64	97,52
Precisão (%)	94,09	94,05	99,80	99,75
Sensibilidade (%)	93,20	93,40	97,39	97,29
Especificidade (%)	69,19	68,94	99,00	98,70

Tabela A.11: Desempenho obtido na classificação linear dos dados WO2_020 filtrados utilizando 2, 3, 63 e 76 características extraídas.

Número de características	2	3	63	76
Acurácia (%)	79,96	79,92	94,68	85,59
Precisão (%)	73,25	73,19	91,88	83,95
Sensibilidade (%)	93,52	93,56	97,85	87,50
Especificidade (%)	66,77	66,66	91,59	83,73

Tabela A.12: Desempenho obtido na classificação linear dos dados WO2_020 filtrados utilizando 2, 3, 75 e 76 características extraídas, após a aplicação do algoritmo *fastICA* na fase de pré-processamento.

Número de características	2	3	75	76
Acurácia (%)	79,06	78,87	92,82	91,51
Precisão (%)	72,78	72,48	88,46	86,68
Sensibilidade (%)	91,90	92,11	98,26	97,81
Especificidade (%)	66,57	65,98	87,53	85,38

Tabela A.13: Resultados obtidos na classificação através do método Fisher LDA utilizando as 2 características com melhor capacidade de discriminação para cada um dos casos.

Dados	Acurácia(%)	Precisão(%)	Sensibilidade(%)	Especificidade(%)
WO2_005				
Sem pré-processamento	78,49	100	71,89	100
Filtro Butterworth	88,28	90,65	94,42	68,27
Filtro Butterworth + ICA	88,77	93,72	91,45	80,04
WO2_007				
Sem pré-processamento	73,70	67,80	88,99	58,78
Filtro Butterworth	79,51	72,96	92,97	66,39
Filtro Butterworth + ICA	78,15	73,57	87,00	69,52
WO2_017				
Sem pré-processamento	85,88	95,75	87,06	79,68
Filtro Butterworth	90,14	94,36	93,88	70,50
Filtro Butterworth + ICA	89,37	94,09	93,20	69,19
WO2_020				
Sem pré-processamento	83,77	89,16	76,36	90,97
Filtro Butterworth	79,96	73,25	93,52	66,77
Filtro Butterworth + ICA	79,06	72,78	91,90	66,57

Tabela A.14: Resultados obtidos na classificação através do método Fisher LDA utilizando as 3 características com melhor capacidade de discriminação para cada um dos casos.

Dados	Acurácia(%)	Precisão(%)	Sensibilidade(%)	Especificidade(%)
WO2_005				
Sem pré-processamento	90,60	92,59	95,35	75,14
Filtro Butterworth	79,46	86,11	87,23	54,14
Filtro Butterworth + ICA	92,53	95,83	94,34	86,83
WO2_007				
Sem pré-processamento	74,47	69,68	85,49	63,72
Filtro Butterworth	79,48	72,93	92,95	66,35
Filtro Butterworth + ICA	76,42	71,60	86,61	66,48
WO2_017				
Sem pré-processamento	90,18	97,57	90,58	88,11
Filtro Butterworth	90,38	94,34	94,20	70,25
Filtro Butterworth + ICA	89,49	94,05	93,40	68,94
WO2_020				
Sem pré-processamento	92,05	90,87	93,23	90,89
Filtro Butterworth	79,92	73,19	93,56	66,66
Filtro Butterworth + ICA	78,87	72,48	92,11	65,98

Tabela A.15: Resultados obtidos na classificação através do método Fisher LDA utilizando todas as características cada um dos casos.

Dados	Acurácia(%)	Precisão(%)	Sensibilidade(%)	Especificidade(%)
WO2_005				
Sem pré-processamento	93,94	98,13	93,87	94,16
Filtro Butterworth	87,11	94,88	87,88	84,56
Filtro Butterworth + ICA	96,53	98,39	97,05	94,82
WO2_007				
Sem pré-processamento	99,07	98,58	99,56	98,61
Filtro Butterworth	90,03	90,35	89,36	90,69
Filtro Butterworth + ICA	86,28	85,08	87,56	85,02
WO2_017				
Sem pré-processamento	15,53	49,21	16,25	11,74
Filtro Butterworth	97,80	99,59	97,79	97,89
Filtro Butterworth + ICA	97,52	99,75	97,29	98,70
WO2_020				
Sem pré-processamento	78,77	72,04	93,08	64,85
Filtro Butterworth	85,59	83,95	87,50	83,73
Filtro Butterworth + ICA	91,51	86,68	97,81	85,38

Tabela A.16: Resultados obtidos na classificação através do método Fisher LDA utilizando as N características com melhor capacidade de discriminação, que permitem obter o desempenho ótimo para cada um dos casos.

Dados	N	Acurácia (%)	Precisão (%)	Sensibilidade (%)	Especificidade (%)
WO2_005					
Sem pré-processamento	69	98,10	98,42	99,10	94,82
Filtro Butterworth	67	97,97	99,56	97,78	98,59
Filtro Butterworth + ICA	71	96,42	98,36	96,94	94,73
WO2_007					
Sem pré-processamento	76	99,07	98,58	99,56	98,61
Filtro Butterworth	71	91,59	91,74	91,17	91,99
Filtro Butterworth + ICA	74	96,25	96,81	95,56	96,93
WO2_017					
Sem pré-processamento	67	98,23	98,55	99,35	92,32
Filtro Butterworth	74	98,06	99,72	97,97	98,54
Filtro Butterworth + ICA	70	97,64	99,80	97,39	99,00
WO2_020					
Sem pré-processamento	66	98,53	97,11	100	97,10
Filtro Butterworth	63	94,68	91,88	97,85	91,59
Filtro Butterworth + ICA	75	92,54	87,85	98,49	86,75

Tabela A.17: Desempenho obtido na classificação baseada na distância euclidiana mínima dos dados WO2_005 utilizando 2, 3 e 76 características extraídas, após a aplicação do algoritmo *fastICA* na fase de pré-processamento.

Número de características	2	3	76
Acurácia (%)	82,93	83,62	83,57
Precisão (%)	93,95	94,91	94,91
Sensibilidade (%)	83,04	83,04	82,98
Especificidade (%)	82,58	85,50	85,50

Tabela A.18: Desempenho obtido na classificação baseada na distância euclidiana mínima dos dados WO2_007 utilizando 2, 3, 37 e 76 características extraídas, após a aplicação do algoritmo *fastICA* na fase de pré-processamento.

Número de características	2	3	37	76
Acurácia (%)	78,38	78,79	85,66	83,40
Precisão (%)	73,86	74,40	88,24	88,30
Sensibilidade (%)	87,00	86,96	81,86	76,51
Especificidade (%)	69,97	70,82	89,36	90,11

Tabela A.19: Desempenho obtido na classificação baseada na distância euclidiana mínima dos dados WO2_017 utilizando 1, 2, 3 e 76 características extraídas, após a aplicação do algoritmo *fastICA* na fase de pré-processamento.

Número de características	1	2	3	76
Acurácia (%)	88,91	88,83	88,76	86,93
Precisão (%)	94,01	94,01	94,00	94,90
Sensibilidade (%)	92,70	92,61	92,52	89,24
Especificidade (%)	68,94	68,94	68,94	74,76

Tabela A.20: Desempenho obtido na classificação baseada na distância euclidiana mínima dos dados WO2_020 utilizando 2, 3, 18 e 76 características extraídas, após a aplicação do algoritmo *fastICA* na fase de pré-processamento.

Número de características	2	3	18	76
Acurácia (%)	78,76	78,78	82,89	81,46
Precisão (%)	72,46	72,38	76,62	74,41
Sensibilidade (%)	91,82	92,13	93,97	95,11
Especificidade (%)	66,06	65,79	72,12	68,18

Tabela A.21: Desempenho obtido no treino e teste dos diferentes conjuntos de dados sem pré-processamento.

		Número de características	Acurácia (%)	Precisão (%)	Sensibilidade (%)	Especificidade (%)
Conjunto 1	Treino	2	74,97	81,12	83,44	55,60
		3	80,14	84,61	87,34	63,69
		71	84,41	92,75	84,17	84,95
		76	71,27	85,59	70,59	72,82
	Teste	2	62,84	58,36	86,05	40,26
		3	60,46	56,32	88,30	33,38
		71	38,76	37,75	37,27	40,21
		76	23,89	19,46	17,31	30,29
Conjunto 2	Treino	2	69,44	67,08	85,60	50,35
		3	69,28	66,82	85,97	49,56
		72	93,31	92,37	95,55	90,67
		76	89,74	88,40	93,31	85,53
	Teste	2	58,59	86,44	60,67	49,92
		3	60,27	86,56	62,41	49,02
		72	27,60	64,26	31,18	8,78
		76	28,66	65,10	32,55	8,18
Conjunto 3	Treino	2	75,36	83,45	80,55	63,50
		3	70,86	82,25	74,08	63,50
		76	93,77	94,99	96,10	88,43
	Teste	2	37,80	35,93	33,17	42,31
		3	45,49	43,23	33,17	57,52
		76	26,68	27,69	29,84	23,98
Conjunto 4	Treino	2	67,63	72,24	78,00	50,51
		3	67,60	72,22	77,97	50,48
		71	90,44	93,29	91,21	89,16
		76	88,60	91,37	90,21	85,94
	Teste	2	70,86	77,67	86,91	18,55
		3	70,95	77,77	86,85	19,11
		71	82,25	89,55	86,94	66,95
		76	76,92	87,97	80,90	63,94

Tabela A.22: Desempenho obtido no treino e teste dos diferentes conjuntos de dados após a aplicação do filtro Butterworth.

		Número de características	Acurácia (%)	Precisão (%)	Sensibilidade (%)	Especificidade (%)
Conjunto 1	Treino	2	30,26	49,08	6,28	85,10
		3	79,36	81,00	91,90	50,71
		76	86,89	92,28	88,57	83,04
	Teste	2	63,15	59,40	79,83	46,92
		3	63,16	57,76	94,08	33,08
		76	83,76	75,25	99,92	68,03
Conjunto 2	Treino	2	78,70	77,75	85,00	71,26
		3	78,58	77,58	85,02	70,97
		74	81,66	86,45	78,43	85,48
		76	81,24	86,46	77,50	85,66
	Teste	2	80,85	95,83	80,72	81,54
		3	80,89	95,84	80,77	81,54
		74	74,42	94,78	73,61	78,68
		76	74,45	94,94	73,51	79,38
Conjunto 3	Treino	2	79,89	85,72	85,30	67,54
		3	80,04	84,90	86,72	64,78
		74	84,68	92,48	84,87	84,23
		76	84,55	92,45	84,70	84,19
	Teste	2	68,52	65,69	75,85	61,36
		3	71,23	68,65	76,80	65,79
		74	36,73	35,58	34,71	38,70
		76	38,76	38,08	38,38	39,13
Conjunto 4	Treino	2	77,44	80,35	84,43	65,89
		3	77,65	80,62	84,39	66,51
		73	83,92	91,30	81,99	87,10
		76	55,62	66,86	56,98	53,36
	Teste	2	70,46	89,54	69,52	73,54
		3	70,26	89,80	68,97	74,48
		73	71,08	88,71	71,28	70,43
		76	53,06	82,01	49,52	64,60

Tabela A.23: Desempenho obtido no treino e teste dos diferentes conjuntos de dados após a aplicação do filtro Butterworth e do algoritmo *fastICA*.

		Número de características	Acurácia (%)	Precisão (%)	Sensibilidade (%)	Especificidade (%)
Conjunto 1	Treino	2	79,83	81,42	92,01	51,99
		3	79,92	81,51	92,02	52,27
		76	86,63	91,08	89,55	79,95
	Teste	2	63,31	57,97	93,02	34,41
		3	61,97	57,09	92,17	32,61
		76	59,77	58,40	64,04	55,62
Conjunto 2	Treino	2	78,97	78,00	85,21	71,60
		3	79,05	78,00	85,40	71,54
		75	85,92	86,94	87,09	84,54
		76	84,86	86,48	85,39	84,23
	Teste	2	79,56	95,77	79,17	81,59
		3	79,60	95,74	79,25	81,44
		75	71,65	96,11	69,05	85,30
		76	71,22	96,66	68,11	87,61
Conjunto 3	Treino	2	75,88	83,57	81,29	63,50
		3	79,54	83,96	87,25	61,92
		76	85,02	92,59	85,29	84,41
	Teste	2	18,57	26,84	36,61	00,00
		3	42,51	45,64	86,09	00,00
		76	47,89	47,81	60,63	35,46
Conjunto 4	Treino	2	71,79	76,05	79,84	58,49
		3	73,21	76,65	81,96	58,77
		71	83,58	91,28	81,42	87,15
		76	83,14	90,49	81,50	85,85
	Teste	2	61,80	79,89	66,92	45,10
		3	62,46	78,69	69,86	38,32
		71	66,22	83,36	69,78	54,61
		76	65,07	84,51	66,54	60,26

B. Manual do Utilizador

B.1. Iniciar a aplicação

Antes de iniciar a aplicação, deverá ter instalado um compilador SDK compatível com o Matlab.

Para que as funções de classificação funcionem corretamente deverá executar os *scripts* *stprpath.m* e *compilemex.m* presentes na *toolbox* *sptrtool*.

Para iniciar a aplicação deverá executar o *script* *APP.m*.

B.2. Painel 1 - *Load*

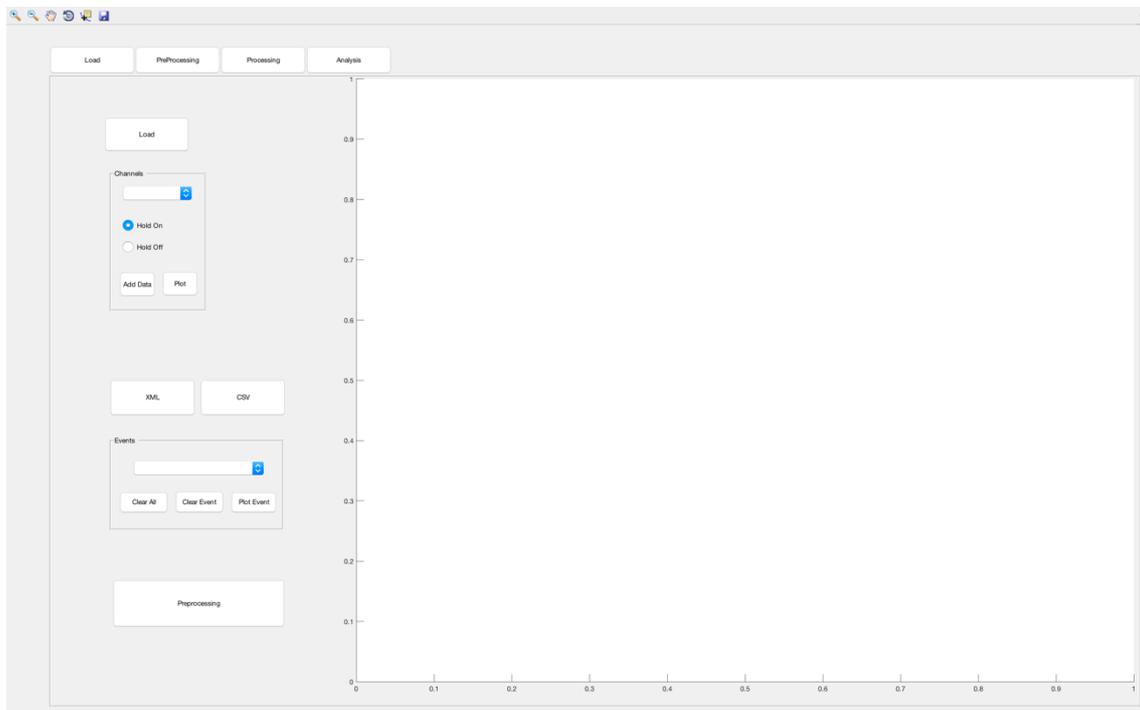


Figura B.1

Os botões existentes no primeiro painel, representado na figura B.1, permitem desempenhar as seguintes funções:

Load – importar dados no formato *.xls* e *.xlsx*

Add Data – adicionar mais dados aos já inseridos

Plot – representar a informação do canal selecionado

Hold On – a seleção deste botão permite representar a informação do canal selecionado, mantendo a representação presente no gráfico

Hold Off – a seleção deste botão implica a representação de apenas o canal selecionado

XML – adicionar ficheiros com informação relativa aos dados em formato *.xml*

CSV – adicionar ficheiros com informação relativa aos dados em formato *.csv*

Plot Event – desenhar uma linha vertical no instante de ocorrência do evento selecionado. É possível representar vários eventos em simultâneo.

Clear Event – apagar a linha de representação do evento selecionado do gráfico

Clear All – apagar as linhas de representação de todos os eventos assinalados no gráfico

PreProcessing – avançar para o painel seguinte

B.3. Painel 2 – *PreProcessing*

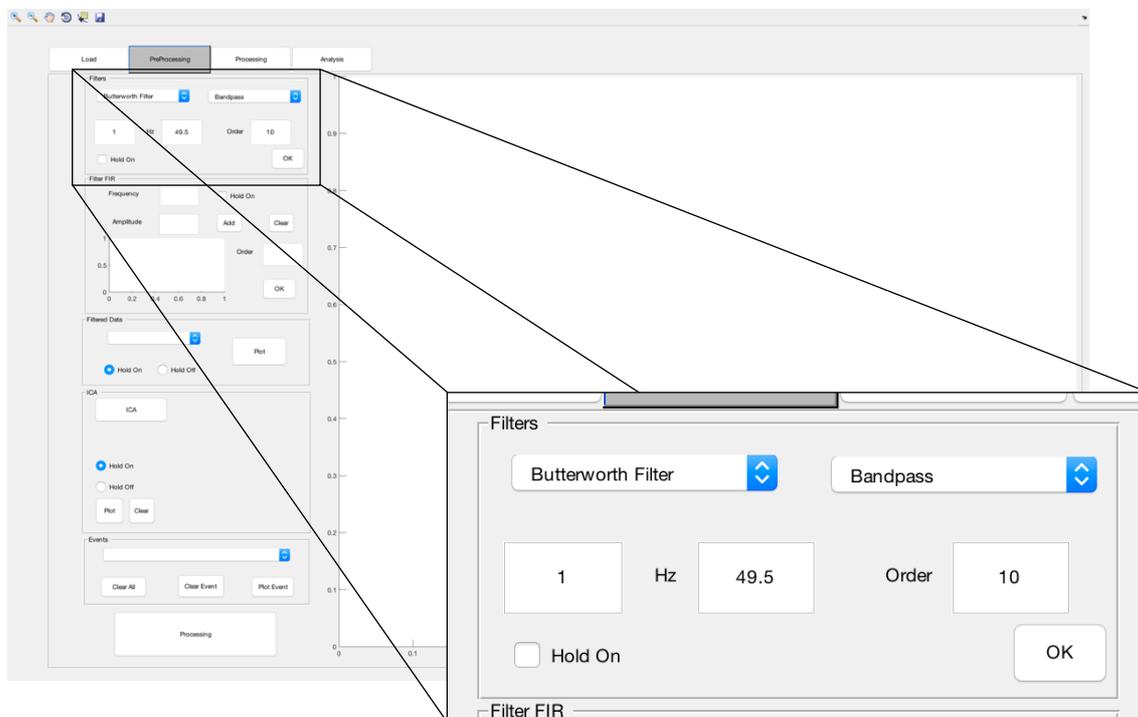


Figura B.2

Na área selecionada na figura B.2 é possível aplicar diferentes tipos de filtros de reposta a impulso infinita. As opções da primeira lista são:

- Butterworth
- Chebyshev Tipo 1
- Chebyshev Tipo 2
- Elíptico

Na segunda lista, é possível completar o tipo de filtro selecionado como:

- Passa-baixo
- Passa-alto
- Passa-banda
- Rejeita-banda

Na definição do filtro é ainda possível adicionar a frequência de corte e a ordem do filtro.

Através da seleção do botão *Hold On* é possível aplicar mais filtros ao sinal já filtrado.

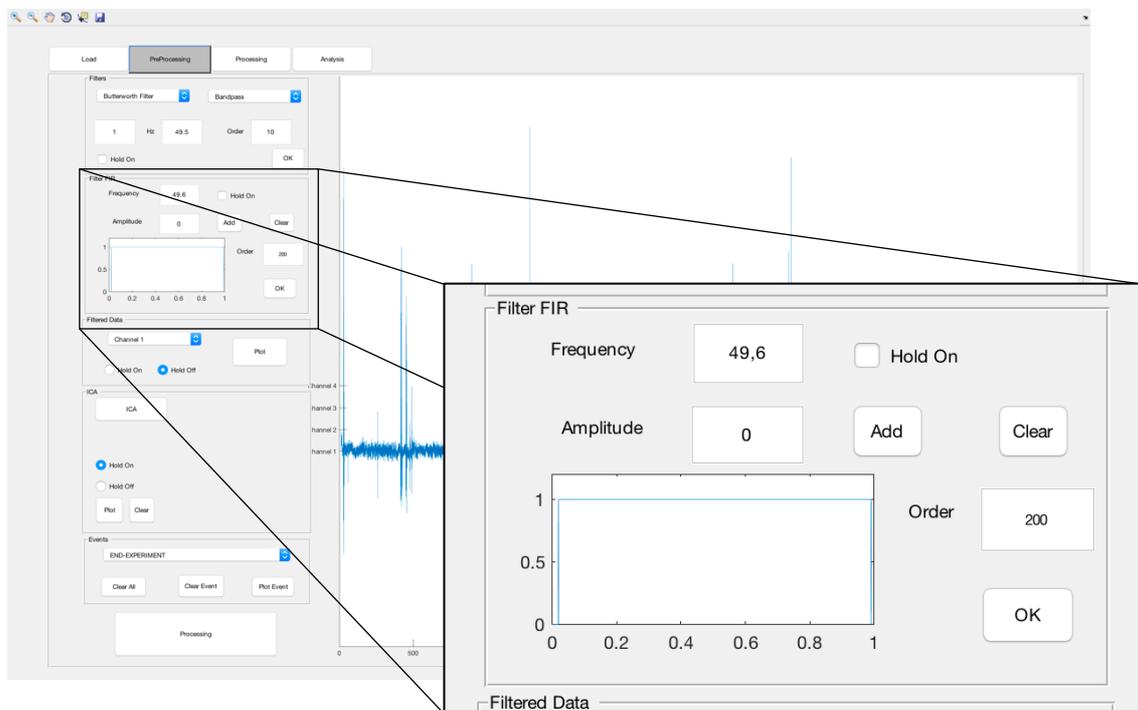


Figura B.3

Na área selecionada na figura B.3 é possível desenhar e aplicar filtros de resposta finita ao impulso através do preenchimento dos campos *Frequency* e *Amplitude*, definindo a atenuação desejada, entre 0 e 1 (0- rejeita totalmente e 1- passa totalmente), para cada amplitude e do clique no botão *Add* para adicionar os pontos ao filtro.

O filtro desejado é representado no gráfico presente na área selecionada à medida que os pontos são inseridos.

A ordem do filtro deve ser definida no campo *Order*.

Através da seleção do botão *Hold On* é possível aplicar mais filtros ao sinal já filtrado.

O botão *OK* permite aplicar o filtro ao sinal e visualizar o filtro aplicado numa janela nova.

O botão *Clear* apaga o filtro desenhado.

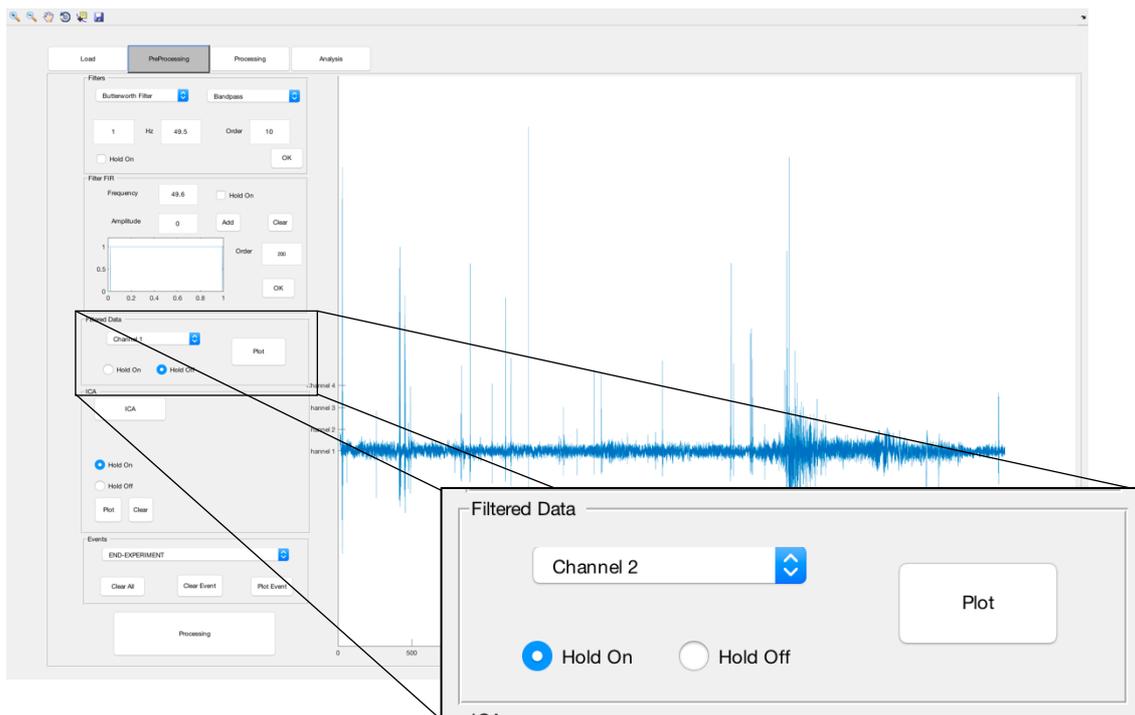


Figura B.4

A área selecionada na figura B.4 permite representar graficamente os sinais filtrados correspondentes a todos os canais de forma individual através da seleção do botão *Hold Off* antes do clicar no botão *Plot*.

É possível representar mais do que um sinal em simultâneo mantendo o botão *Hold On* selecionado enquanto é feita a representação de cada um dos sinais através do botão *Plot*.

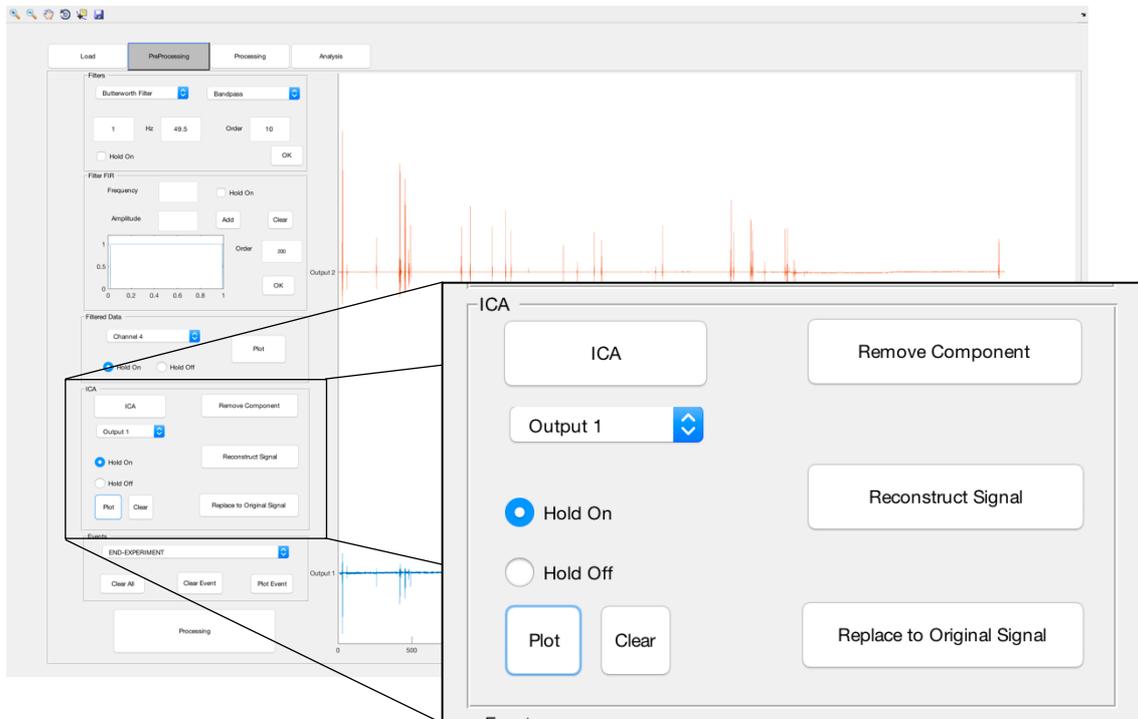


Figura B.5

Na área selecionada na figura B.5, é possível desempenhar as seguintes funções:

ICA – aplicar o algoritmo *fastICA*

Lista *Outputs* – selecionar a estimativa desejada

Plot – representar graficamente a estimativa selecionada

Hold On – manter a informação representada no gráfico quando representada outra estimativa.

Hold Off – representar apenas a estimativa pretendida

Clear – limpar a representação gráfica

Remove Component – remover a estimativa selecionada

Reconstruct Signal – reconstruir o sinal sem os componentes removidos

Replace to Original Signal – repor o sinal original

O botão *Processing* presente neste painel permite avançar para o painel seguinte.

B.4. Painel 3 – *Processing*

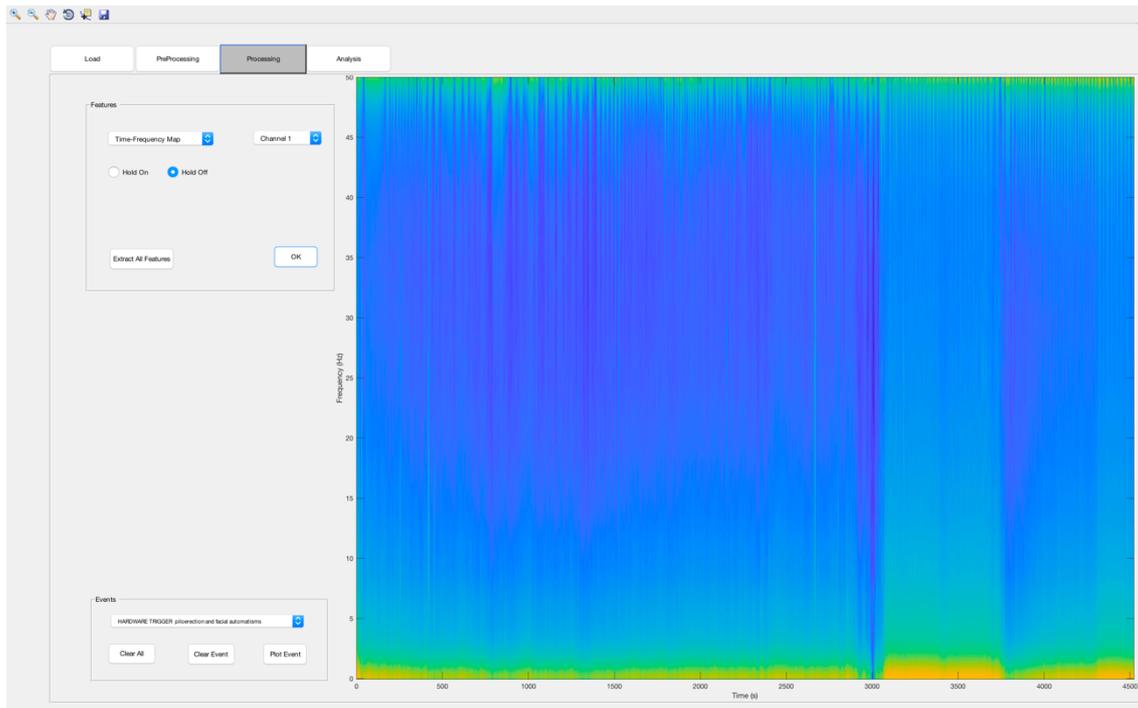


Figura B.6

Neste painel é possível representar as características extraídas dos sinais de forma individual, mantendo o botão *Hold Off* selecionado, ou em simultâneo, mantendo o botão *Hold On* selecionada quando pressionado o botão *OK*.

Através do botão *Extract All Features* é possível extrair todas as características pré-definidas dos sinais e analisá-las no painel seguinte (*Analysis*).

Após a seleção da opção *Wavelet Coefficients* na lista, aparece uma lista para seleção do coeficiente pretendido assinalada na figura B.7.

Após a seleção da opção *Relative Power* na lista, aparece uma lista para seleção da banda de frequência pretendida assinalada na figura B.8. Para além das bandas de frequência presentes na lista, é possível representar a potência relativa de bandas

customizadas, através da seleção do botão *Custom* e da definição dos limites da banda através do preenchimento dos campos assinalados na figura B.9.

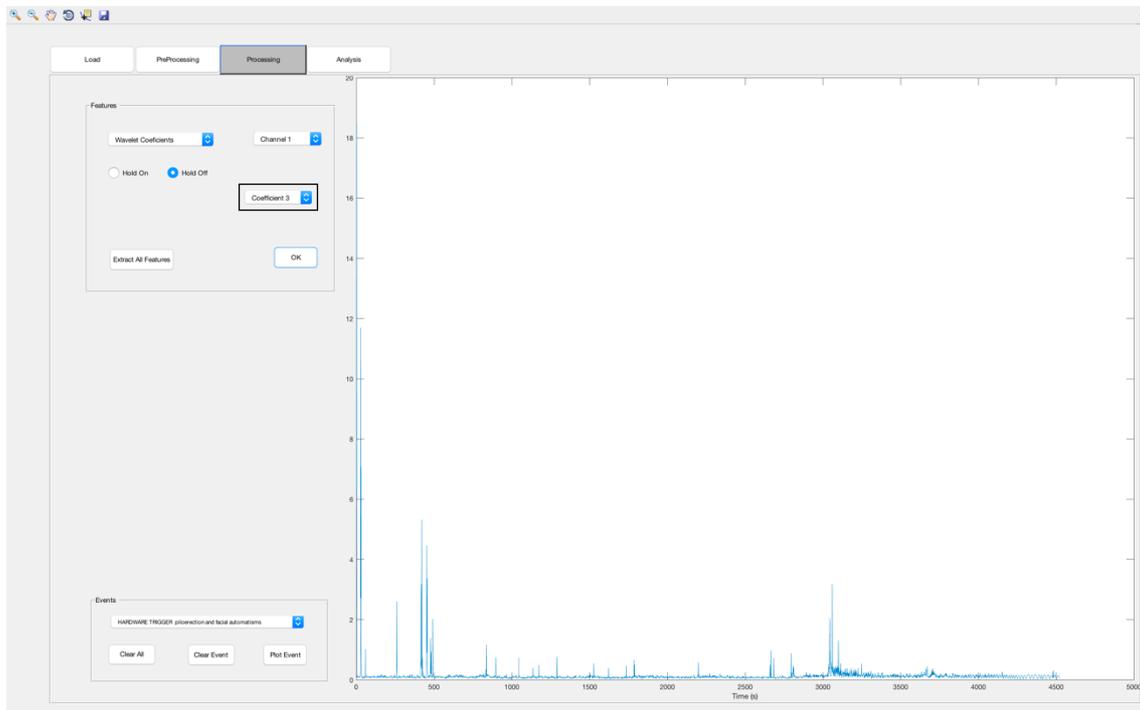


Figura B.7

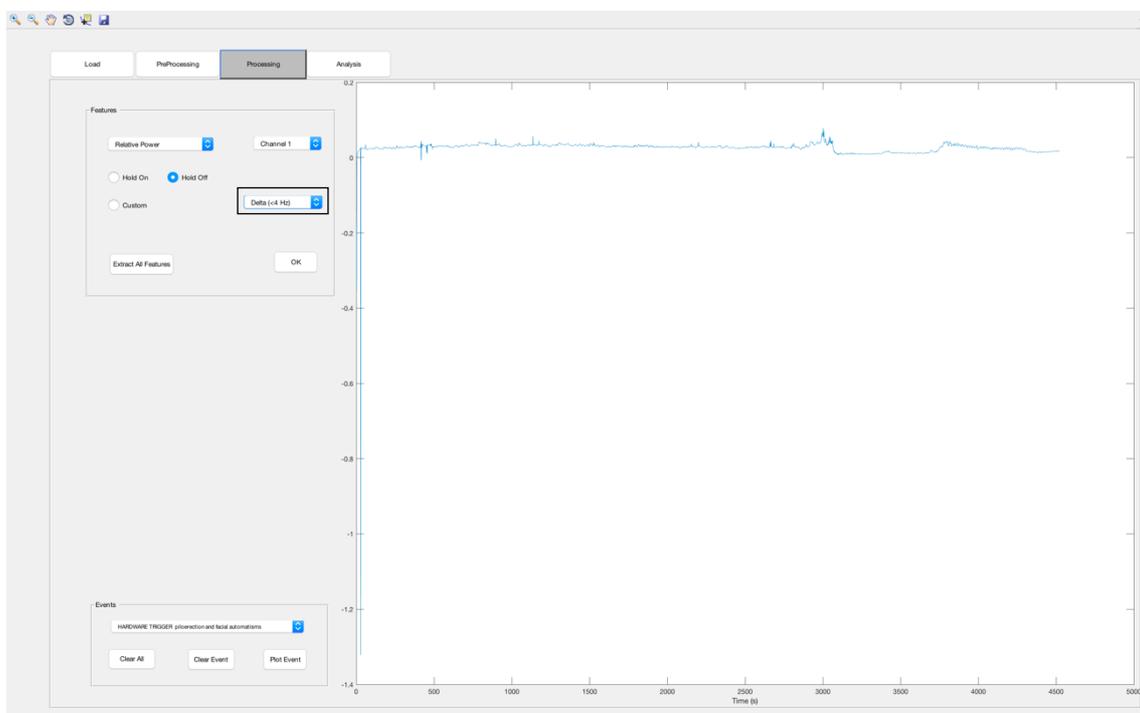


Figura B.8

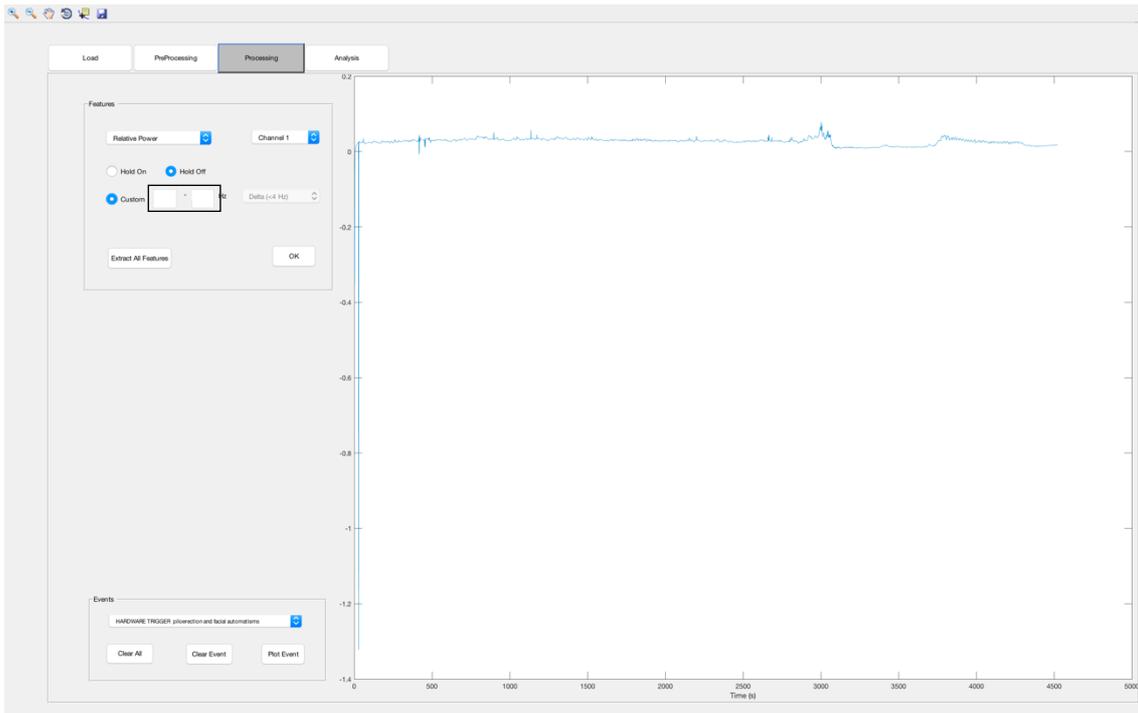


Figura B.9

B.5. Painel 4 – Analysis

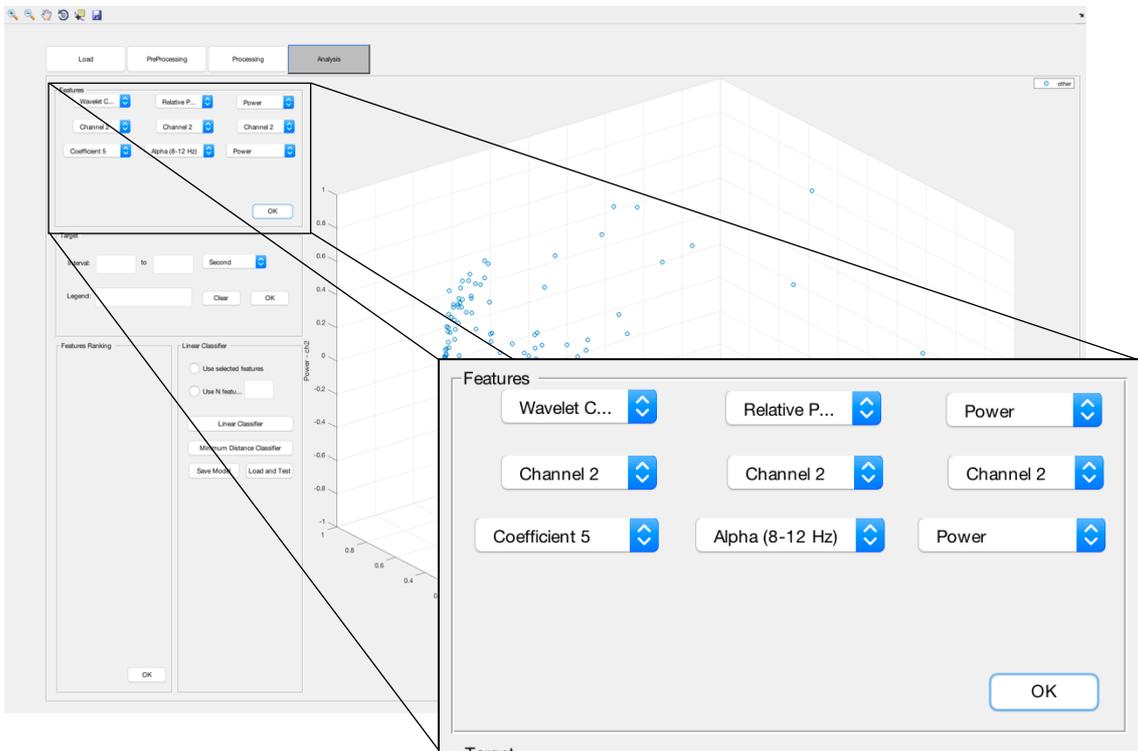


Figura B.10

Na área selecionada na figura B.10 é possível representar graficamente 2 ou 3 características selecionadas.

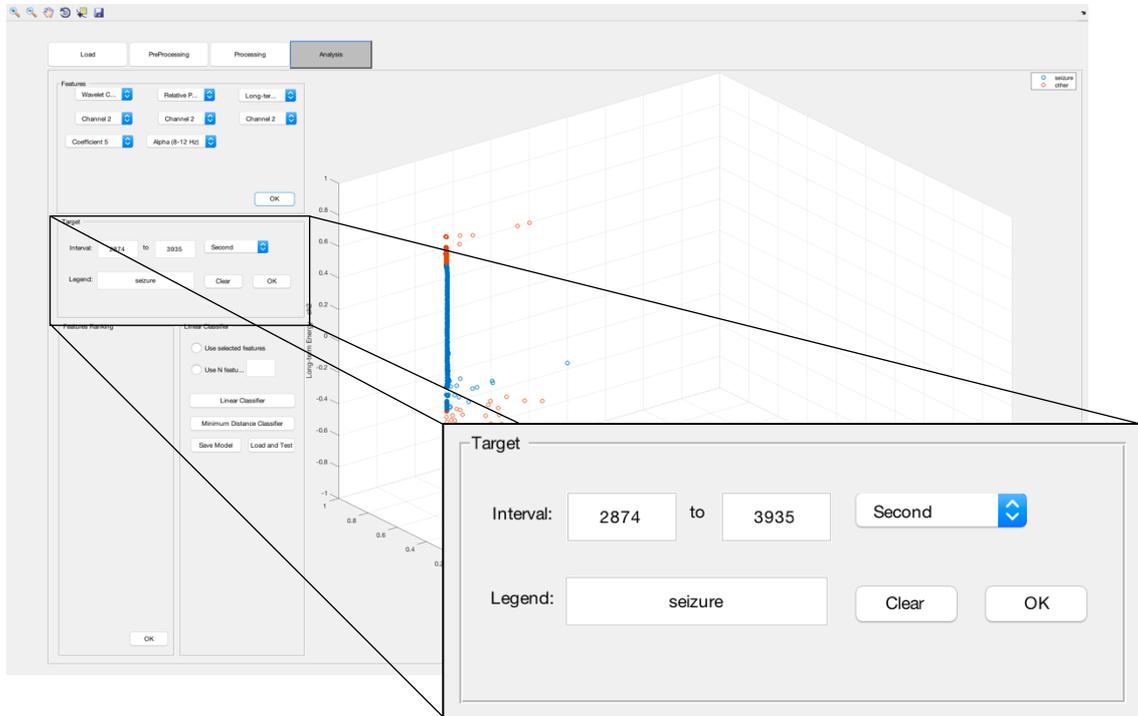


Figura B.11

Na área selecionada na figura B.11 é possível definir o *target*, identificando o intervalo de tempo (em segundos) ou de pontos em que se encontram os dados relativos ao grupo definido.

O botão *Clear* limpa o *target* definido anteriormente.

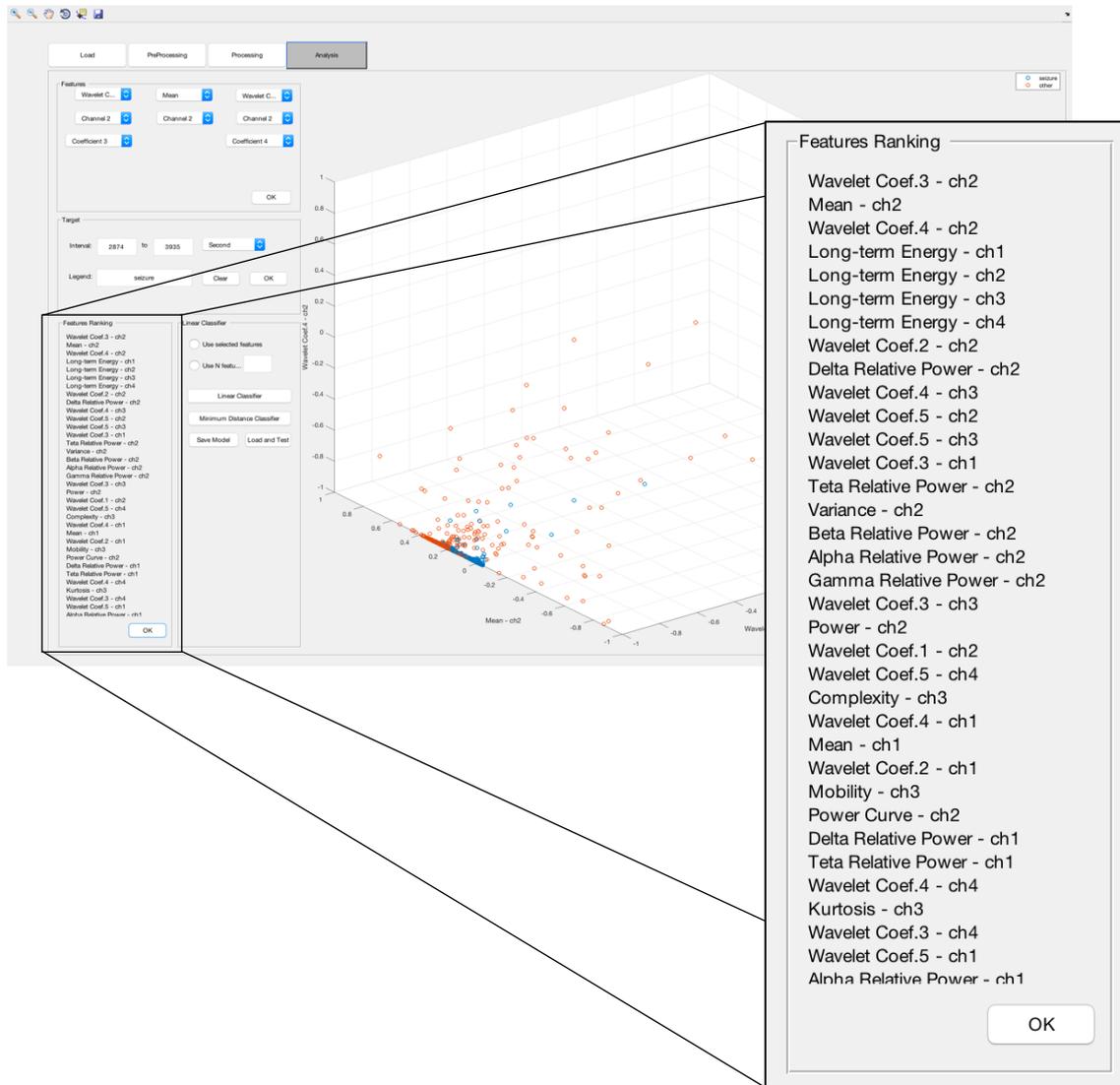


Figura B.12

Na área selecionada da figura B.12 é possível visualizar as características ordenadas de forma decrescente de acordo com a sua capacidade de discriminação dos grupos definidos anteriormente no *target*.

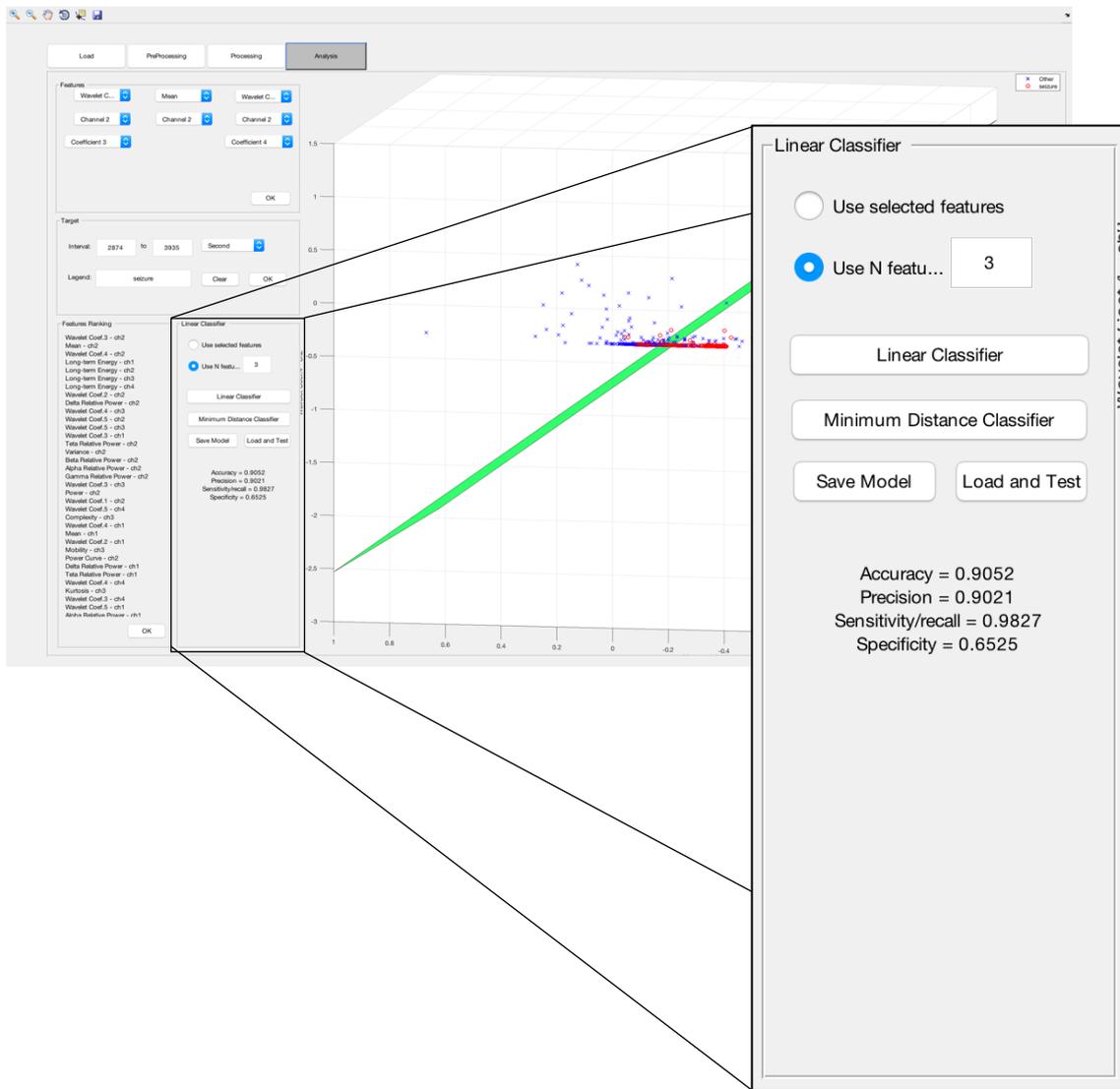


Figura B.13

Na área selecionada na figura B.13, é possível desempenhar as seguintes funções:

Use selected features – utilizar as características selecionadas na secção assinalada na figura B.10.

Use N features – usar as primeiras N características da lista ordenada assinalada na figura B.12

Para utilizar todas as características extraídas basta não selecionar nenhuma das duas opções anteriores

Linear Classifier – treinar um classificador *Fisher LDA*

Minimum Distance Classifier – treinar um classificador baseado na distância euclidiana mínima

Save Model – gravar o modelo treinado na diretoria onde se encontra o *script* da aplicação com o nome *model.mat*. Após esta operação deverá alterar o nome do ficheiro para evitar que seja substituído por outro quando realizar de novo esta operação.

Load and Test – importar e testar um modelo já treinado

Visualizar as medidas de desempenho da s

