



**André Tiago  
Marques Reis**

**Sistema de classificação de sinais de  
electroencefalograma**





**André Tiago  
Marques Reis**

**Sistema de classificação de sinais de  
electroencefalograma**

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Eng. Electrónica e Telecomunicações, realizada sob a orientação científica de Ana Maria Tomé, Professora do Departamento Electrónica Telecomunicações e Informática da Universidade de Aveiro



**o júri**

presidente

**Prof. Dr. João Paulo Trigueiros da Silva Cunha**

Professor Associado da Universidade de Aveiro

vogais

**Prof. Dra. Ana Maria Perfeito Tomé**

Professora Associada do Departamento de Electrónica, Telecomunicações e Informática de Aveiro

**Prof. Dr. Miguel Angel Guevara Lopez**

Professor Titular da Facultad de Informática da Universidade de Ciego de Ávila de Cuba



## **agradecimentos**

À minha orientada Professora Doutora Ana Maria Tomé, pela orientação científica, apoio e disponibilidade demonstrada ao longo destes meses de trabalho.

Dedico este trabalho aos meus pais, pelo incansável apoio e sacrifício prestado, em prol de uma educação mais alicerçada e credenciada.

À Mestre Ana Rita Teixeira, pela disponibilidade, colaboração e interesse em todo o projecto.

À futura psicóloga Daniela Caprichoso, pela colaboração prestada em todo o processo de escrita.

Aos meus colegas de curso, em especial ao Samuel Costa, pela ajuda e companheirismo demonstrados ao longo de todos estes anos.



## Resumo

Este trabalho apresenta um método para detectar actividade epiléptica em sinais de *Electroencefalograma* (EEG).

O sistema é constituído por dois blocos, o primeiro trata da extracção de características do sinal EEG e o segundo procede à classificação do sinal em 2 classes.

As características são medidas (como energia média, máximo do sinal, entre outras) extraídas a partir das saídas da decomposição discreta em wavelet do sinal.

A classificação será tratada como um problema binário, assumindo-se que os dois tipos de classes (Classe não epiléptica e Classe epiléptica), são linearmente separáveis. O classificador utilizado é do tipo *Support Vector Machine* (SVM).

O desempenho do sistema é também apresentado segundo um método experimental, onde diferentes estratégias para a organização do conjunto de treino são debatidas.

Para avaliar a performance deste sistema, comparativamente a outros, utilizou-se uma base de dados publicada [1].

Os resultados obtidos usando um classificador linear apresentaram-se prometedores, com uma precisão situada entre 89,74% - 99,87%.



## **Abstract**

This work presents a method for detecting epileptic activity in Electroencephalogram (EEG) signals.

The system is divided into two blocks, the first dealing with feature extraction from the EEG and the second with the classification problem.

The features are measures (like energy, maximum and so on) taken into the outputs of the discrete wavelet decomposition of the signal.

To perform the classification the support vector machine was chosen. This binary classifier was designed assuming that the two classes (epileptic and non-epileptic activity) are linearly separable. The performance of the system is also presented using an experimental study where different strategies to organize the training data sets are also discussed.

The accuracy of the system is in range of 89,74% - 99,87% in a publicly available data set used by other works.



# Lista de Figuras

1.1	Formação do conjunto de treino . . . . .	2
1.2	Abordagem à detecção . . . . .	2
2.1	Representação da disposição de eléctrodos segundo o sistema 10-20 . . . . .	6
2.2	Representação temporal de um sinal EEG . . . . .	6
2.3	Exemplos de sinais EEG . . . . .	7
2.4	Exemplos de sinais EEG com spikes . . . . .	8
2.5	Exemplos de sinais EEG com ondas abruptas . . . . .	9
2.6	Exemplos de sinais EEG com artefactos musculares . . . . .	10
2.7	Exemplos de sinais EEG com artefactos oculares . . . . .	11
4.1	Diagrama de blocos do problema de classificação de um sinal EEG . . . . .	17
4.2	Diagrama de blocos do bloco de extracção de características . . . . .	18
4.3	Estrutura dos filtros com sub amostragem . . . . .	18
4.4	Decomposição em árvore . . . . .	19
4.5	Estrutura do resultado da função $wt$ . . . . .	20
4.6	Diagrama de blocos do bloco do processo de Classificação . . . . .	22
4.7	Divisão dos dados segundo diversos hiperplanos . . . . .	23
4.8	Definição dos hiperplanos . . . . .	24
4.9	Dados linearmente não linearmente separáveis . . . . .	27
4.10	Hiperplano de separação de dados não linearmente separáveis . . . . .	28
4.11	Aprendizagem considerando $C=1$ . . . . .	29
4.12	Aprendizagem considerando $C=100$ . . . . .	30
5.1	Representação dos diversos tipos de segmentos EEG (A a E) . . . . .	36
5.2	Dados organizados à saída da extracção de características . . . . .	37
5.3	Relação entre características do nível 3 (a) Máximo, (b) Mínimo, (c) Desvio, (d) Energia Média . . . . .	39
5.4	Relação entre características do nível 6 (a) Máximo, (b) Mínimo, (c) Desvio, (d) Energia Média . . . . .	40
5.5	Boxplot do erro associado ao Teste no conjunto A e ao Teste no conjunto E utilizando as 4 características (Experiência A caso 1) . . . . .	41
5.6	Boxplot do erro associado ao Teste no conjunto B, Teste no conjunto C e Teste no conjunto D, utilizando as 4 características (Experiência A caso 1) . . . . .	41
5.7	Boxplot do erro associado ao Teste no conjunto A e ao Teste no conjunto E utilizando a Energia média como característica (Experiência A caso 2) . . . . .	42

5.8	Boxplot do erro associado ao Teste no conjunto B, Teste no conjunto C e Teste no conjunto D utilizando a Energia média como característica (Experiência A caso 2) . . . . .	42
5.9	Boxplot do erro associado ao Teste no conjunto A, Teste no conjunto B, Teste no conjunto C, Teste no conjunto D, Teste no conjunto E utilizando as 4 características (Experiência B caso 1) . . . . .	44
5.10	Boxplot do erro associado ao Teste no conjunto A, Teste no conjunto B, Teste no conjunto C, Teste no conjunto D e Teste no conjunto E utilizando a Energia média como caracteística(Experiência B caso 2) . . . . .	45
5.11	Análise a um segmento do tipo A . . . . .	48
5.12	Análise a um segmento do tipo E . . . . .	48
5.13	Análise a um segmento do tipo B . . . . .	49
5.14	Análise a um segmento do tipo C . . . . .	50
5.15	Análise a um segmento do tipo D . . . . .	50

# Lista de Tabelas

2.1	Gamas de Frequência de actividade EEG [2] . . . . .	7
5.1	Caracterização dos segmentos EEG . . . . .	35
5.2	Frequências correspondentes aos níveis de decomposição . . . . .	36
5.3	Dimensão dos diversos treinos na experiência A . . . . .	40
5.4	Dimensão dos diversos treinos na experiência B . . . . .	43
5.5	Matriz Confusão . . . . .	46
5.6	Accuracy associada ao Caso 1 e 2 da experiência A . . . . .	47
5.7	Accuracy associada ao Caso 1 e 2 da experiência B . . . . .	47



# Glossário

**daub** *Daubechies.*

**ECG** *Electrocardiograma.*

**EEG** *Electroencefalograma.*

**EOG** *Electro-oculograma.*

**KKT** *Karush-Kuhn-Tucker.*

**RBF** *Radial Basis Function.*

**SVM** *Support Vector Machine.*

**TAC** *Tomografia Axial Computorizada.*

**TDW** *Transformada Discreta Wavelet.*



# Conteúdo

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Objectivos . . . . .	1
1.2	Abordagem . . . . .	2
<b>2</b>	<b>Epilepsia e EEG</b>	<b>5</b>
2.1	Aquisição de sinal EEG . . . . .	5
2.2	Bandas características do EEG . . . . .	6
2.3	Actividade anormal EEG . . . . .	7
2.4	Artefatos EEG . . . . .	9
<b>3</b>	<b>Sistemas Automáticos</b>	<b>13</b>
3.1	Extracção de características . . . . .	13
3.2	Classificação . . . . .	14
<b>4</b>	<b>Sistema baseado em classificador binário</b>	<b>17</b>
4.1	Extracção de características . . . . .	18
4.1.1	Transformada discreta de Wavelet . . . . .	18
4.1.2	Implementação da TDW . . . . .	19
4.1.3	Vector de Características . . . . .	21
4.2	Classificação . . . . .	22
4.2.1	Classificação com SVM . . . . .	22
4.2.2	Considerações finais sobre SVM . . . . .	27
4.2.3	Organização do Treino e do Teste . . . . .	28
4.2.4	Implementação do classificador . . . . .	29
4.2.5	Software desenvolvido . . . . .	31
<b>5</b>	<b>Resultados</b>	<b>35</b>
5.1	Caracterização dos sinais EEG . . . . .	35
5.2	Extracção de características . . . . .	36
5.2.1	Vector de características . . . . .	37
5.2.2	Organização do treino e do teste . . . . .	38
5.3	Validação das características . . . . .	38
5.4	Avaliação do classificador . . . . .	38
5.4.1	Experiência A . . . . .	40
5.4.2	Experiência B . . . . .	43
5.5	Análise global . . . . .	46
5.6	Análise Temporal . . . . .	47

<b>6 Conclusões</b>	<b>51</b>
6.1 Propostas futuras . . . . .	53
<b>Referências Bibliográficas</b>	<b>56</b>
<b>Apêndice A</b>	<b>57</b>

# Capítulo 1

## Introdução

Actualmente a epilepsia é considerada uma das desordens neurológicas mais comuns, com predominância de 0.6%-0.8% da população mundial. Esta caracteriza-se por um mau funcionamento repentino e periódico do cérebro, denominado por crises epiléticas, podendo ou não haver convulsões.

O diagnóstico de epilepsia é puramente clínico. Faz-se apenas pela descrição das crises, pelo que é muito importante ter uma descrição pormenorizada destas por quem as observa.

As crises epiléticas caracterizam-se por uma descarga eléctrica cerebral desorganizada que se propaga para todas as regiões do cérebro, levando a uma alteração de toda actividade cerebral.

A sua manifestação pode caracterizar-se pela alteração comportamental, na qual o indivíduo pode falar coisas sem sentido, por movimentos estereotipados de um membro, ou mesmo através de episódios nos quais o paciente fica com o olhar parado, fixo, podendo ocorrer esporadicamente ao longo de uma semana ou com frequência ao longo de um dia [3][4].

O EEG é um dos métodos actuais que regista a actividade eléctrica cerebral e permite identificar as áreas cerebrais onde as crises ocorrem.

Em alguns casos, apenas para se identificar a doença que pode estar a causar a epilepsia, pode ser necessário recorrer-se a exames de imagem - *Tomografia Axial Computorizada* (TAC) ou Ressonância Magnética Nuclear. No entanto, uma das desvantagens do TAC é o facto de utilizar radiação X. Para além do elevado custo que acarreta, a exposição do corpo humano a este tipo de radiações pode ter um efeito maléfico.

Embora o EEG, seja considerado o principal método de auxílio ao diagnóstico de epilepsia, a análise deste tipo de sinal é bastante árdua. Além da dificuldade na interpretação dos sinais devido sobretudo à presença de artefactos.

### 1.1 Objectivos

Este trabalho propõe-se a investigar, projectar e executar um método de análise e posteriormente de classificação de sinal EEG. A sua elaboração recairá, sobre aspectos fundamentais do ponto de vista da extracção de características relevantes para a detecção de anormalidade/normalidade de sinal EEG. A utilização da *Transformada Discreta Wavelet* (TDW), terá também um papel preponderante neste processo, pois permite uma descrição tempo-frequência do sinal em análise. Estes dois aspectos terão um papel fundamental na classificação.

Em suma, pretende-se obter uma classificação capaz de atingir patamares de desempenho satisfatório, quer na classificação de sinal epiléptico quer na classificação de sinal não epiléptico.

## 1.2 Abordagem

Para este trabalho foram escolhidos vários tipos de sinais; sinais provenientes de voluntários saudáveis, sinais que apresentam características de pacientes em plena crise epiléptica bem como sinais de pacientes com epilepsia mas sem crise epiléptica.

O sistema desenvolvido passa por tratar um determinado sinal EEG, que previamente tratado e caracterizado determina a qual de duas classes pertence o sinal (classe não epiléptica ou classe epiléptica).

Para cada tipo de sinal são extraídas características. Estas são organizadas num vector denominado por vector de características. O conjunto de vectores de características provenientes de diferentes tipos de sinais permite a elaboração do conjunto de treino (figura 1.1).

Com este conjunto formado, treinou-se classificador, tendo em vista uma posterior classificação de um determinado sinal EEG.

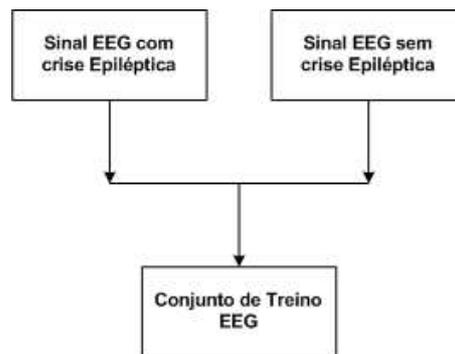


Figura 1.1: Formação do conjunto de treino

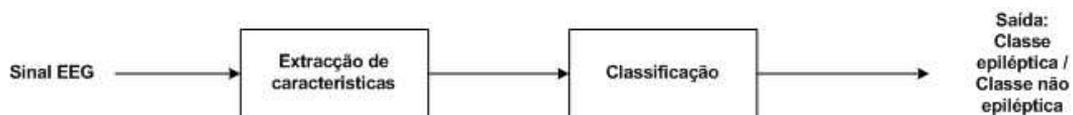


Figura 1.2: Abordagem à detecção

As figuras 1.1 e 1.2 ilustram o tipo de abordagem ao problema, ou seja, a primeira etapa passa pela elaboração de uma base de dados contendo diversas características (relacionadas com amplitude e frequência) de segmentos de sinal previamente etiquetados. A segunda etapa envolve a elaboração de um classificador do tipo SVM.

Numa primeira fase são analisados sinais do tipo normal (provenientes de voluntários saudáveis) e sinais com crise epiléptica (sinais que apresentam características de pacientes em plena crise epiléptica).

A segunda fase é mais abrangente, onde serão distinguidos casos de sinais não epiléticos (sinais provenientes de voluntários saudáveis e sinais de pacientes com epilepsia mas sem crise epilética), e sinais de crise epilética (sinais que apresentam características de pacientes em plena crise epilética).



## Capítulo 2

# Epilepsia e EEG

A epilepsia é conhecida hoje em dia, como sendo uma desordem neurológica do sistema nervoso central que predispõem indivíduos a apresentarem crises periódicas.

Os mecanismos genéticos e moleculares que causam epilepsia, permanecem ainda desconhecidos, no entanto, esta desordem é mais comum em indivíduos cujo cérebro foi comprometido por algum tipo de distúrbio (doenças genéticas, anomalias congénitas, trombozes, tumores, etc).

Uma forma de analisar a epilepsia ou sintomas epilépticos é através do electroencefalograma (EEG); trata-se de um sinal eléctrico produzido no cérebro cuja tensão pode variar entre uma amplitude de 2 a várias centenas de microvolts. Este sinal é gerado pelos milhões de neurónios corticais.

Na prática, os eléctrodos são distribuídos simetricamente em torno do escalpe, segundo as indicações da figura 2.1, permitindo desta forma a obtenção de uma relação temporal e espacial da actividade eléctrica no cérebro. Este processo é simples e indolor[5].

Um indivíduo sujeito a um EEG, pode ser exposto a vários estímulos, como luzes brilhantes ou cintilantes, com o objectivo de provocar uma crise convulsiva. Durante esta, a actividade eléctrica do cérebro acelera-se, produzindo um padrão desordenado em forma de ondas. Estes registos das ondas cerebrais ajudam a identificar a epilepsia.

### 2.1 Aquisição de sinal EEG

O EEG utiliza convencionalmente um sistema de 21 eléctrodos dispostos segundo a norma internacional 10-20 (norma para a colocação dos eléctrodos no escalpe). Este sistema é baseado na relação entre a posição de um eléctrodo e a área subjacente ao córtex cerebral.

Os pontos apresentados na figura 2.1 representam eléctrodos. Cada posição está associada uma letra de modo a identificar a posição no hemisfério.

As letras usadas são as F, T, C, P e O para Frontais, Temporais, Centrais, Parietais e Occipitais, respectivamente. As letras A funcionam como referência aquando a aquisição de potencial.

Os números pares (2,4,6 e 8) estão associados ao hemisfério direito, os números ímpares (1,3,5 e 7) estão associados ao hemisfério esquerdo e Z associado aos eléctrodos colocados na zona central [6].

Um exemplo de aquisição de sinal pode ser observado na figura 2.2 onde a visualização é relativa ao Cz.

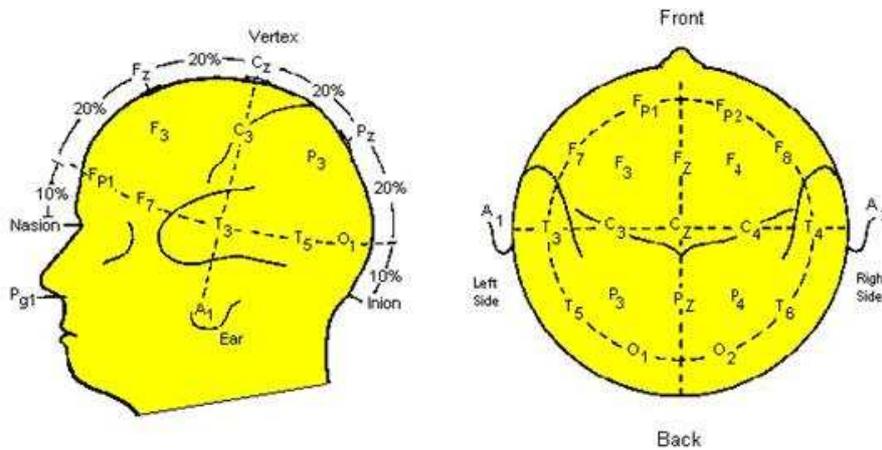


Figura 2.1: Representação da disposição de eléctrodos segundo o sistema 10-20

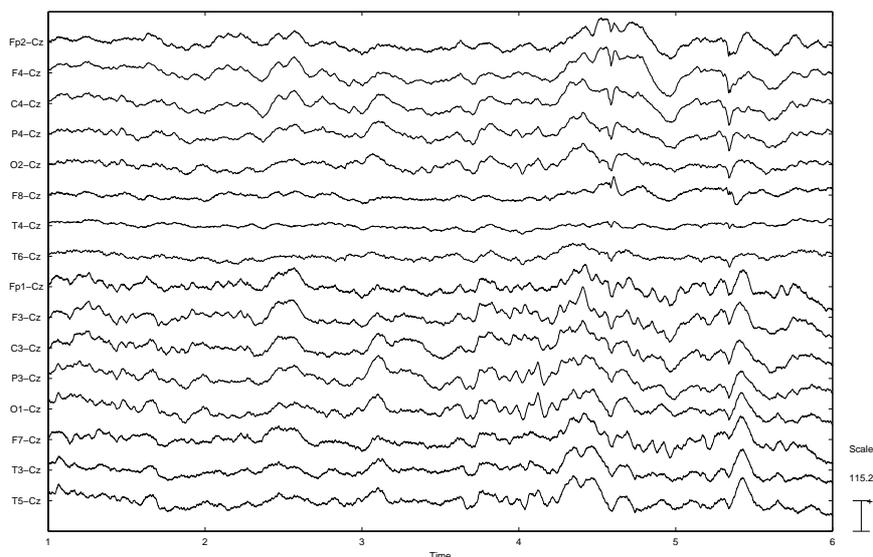


Figura 2.2: Representação temporal de um sinal EEG

## 2.2 Bandas características do EEG

A gama de frequências do sinal EEG com relevância clínica é limitada ao intervalo  $[0.5Hz, 50Hz]$ . As actividades de crise epiléptica são mais limitadas, correspondendo ao intervalo de frequências de  $[0.5Hz, 25Hz]$ [7], havendo contudo nesse intervalo actividade normal.

No sinal EEG podem ser observados padrões diferentes que estão relacionados com conteúdos de frequência diferentes.

Na maioria dos adultos, o padrão EEG de vigília consiste em ondas sinusoidais na frequência de  $8-12Hz$  que são predominantes nas áreas occipitais, sendo conhecida como banda alfa.

Esta banda pode ser atenuada pela abertura ocular, actividade mental ou pelo estado sonolência,

ou seja, esta actividade pode ser aumentada pelo fechar de olhos ou diminuída quando pensamos ou calculamos.

As frequências superiores a  $12Hz$  são chamadas bandas beta e são normalmente encontradas sobre as regiões frontais e centrais. A preocupação, a vigilância, o estado de escuta, de pensamento promovem o início desta actividade.

A banda teta é a actividade de EEG com frequência situada entre  $4-8Hz$ . Esta actividade é anormal em adultos acordados, mas pode ser observada frequentemente no sono ou em crianças ainda em idade jovem (até 13 anos). Pode-se também definir esta banda como o repositório para as memórias, emoções e sensações, concluindo-se que esta actividade esta associada ao subconsciente.

A banda delta está situada abaixo dos  $4Hz$ ; isto poderá ocorrer no sono profundo, onde as ondas delta estão envolvidas com a “mente inconsciente”. Esta banda também pode ser também associada a crianças até um ano de idade.

Assim, a figura 2.3 representa as bandas características, enquanto a tabela 2.1 apresenta o intervalo de gamas de frequências a pertencem [8].

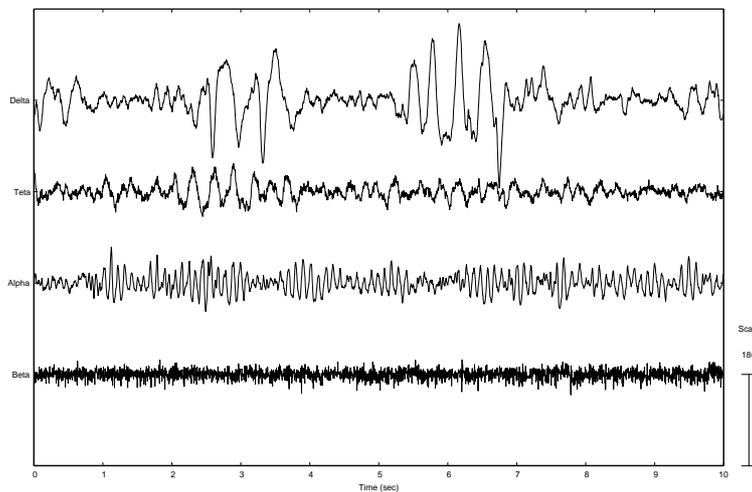


Figura 2.3: Exemplos de sinais EEG

Delta	0.5-4 Hz
Teta	4-8 Hz
Alpha	8-12 Hz
Beta	12-30 Hz

Tabela 2.1: Gamas de Frequência de actividade EEG [2]

## 2.3 Actividade anormal EEG

A actividade anormal de EEG é toda a actividade que é predominante no EEG dos grupos de pessoas que apresentam queixas neurológicas, ou doenças do foro neurológico.

Um EEG anormal pode ser uma onda invulgar, que pode ser caracterizada essencialmente pela amplitude e frequência.

A partir da análise de sinal EEG de um indivíduo com apresentação de crise é possível observar as constantes variações de amplitude e frequência nas formas de onda do sinal.

As variações abruptas de sinal EEG, mais comuns, são denominadas na literatura por “spikes” (figura 2.4) e “sharp waves” (figura 2.5). Estas apresentam picos de elevada amplitude, exibindo uma duração média de 20-70 milissegundos (spikes) e 70-200 milissegundos (sharp waves).

Para doentes afectados por epilepsia estas anomalias são frequentes, significando episódios de crise.

Uma crise EEG é uma actividade anormal, que resulta em descargas neuronais contínuas anormais, na qual o indivíduo apresenta alterações no comportamento involuntário, movimento, sensação ou consciência. Nas crises epilépticas podem ser distinguidas as crises parciais. Numa crise parcial a actividade epiléptica começa sobre uma região localizada no cérebro. As crises parciais que não afectam a consciência, estão classificadas como crises parciais simples, enquanto aquelas que afectam a consciência são classificadas como crises parciais complexas. As crises parciais simples não alteram a consciência, mas danificam temporariamente os sistemas sensoriais ou motor. Uma crise parcial simples que tenha como origem uma área localizada no cérebro pode ser chamada de crise sensorial parcial simples ou de crise parcial simples do sistema motor. Os sintomas sensoriais incluem alucinação que afecta a visão, a audição, ou o olfacto. As crises parciais simples motoras têm umas manifestações clínicas mais desobstruídas que incluem empurrões musculares rápidos e movimentos repentinos.

As crises parciais complexas conduzem a uma perda de consciência. São precedidas frequentemente por uma ilusão sensorial e acompanhadas de movimentos musculares *desorientados*.

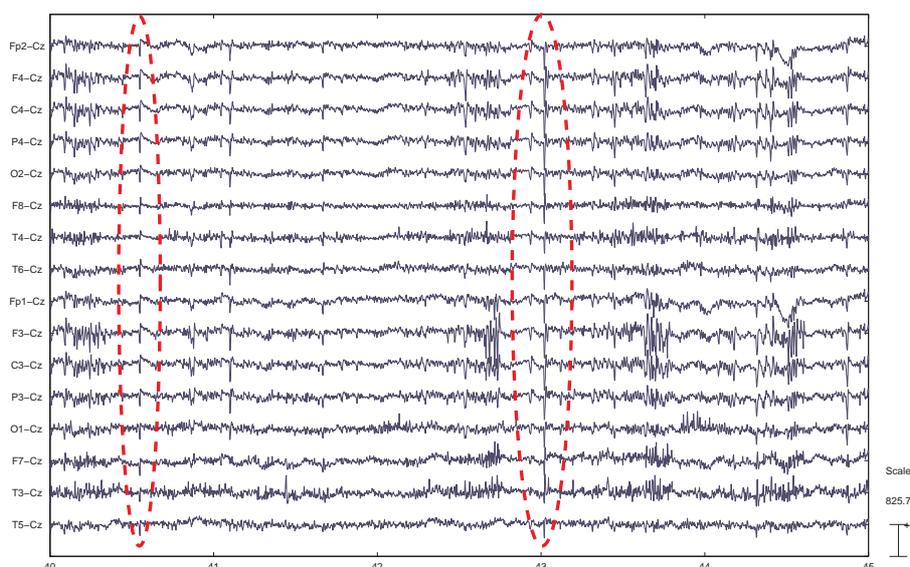


Figura 2.4: Exemplos de sinais EEG com spikes

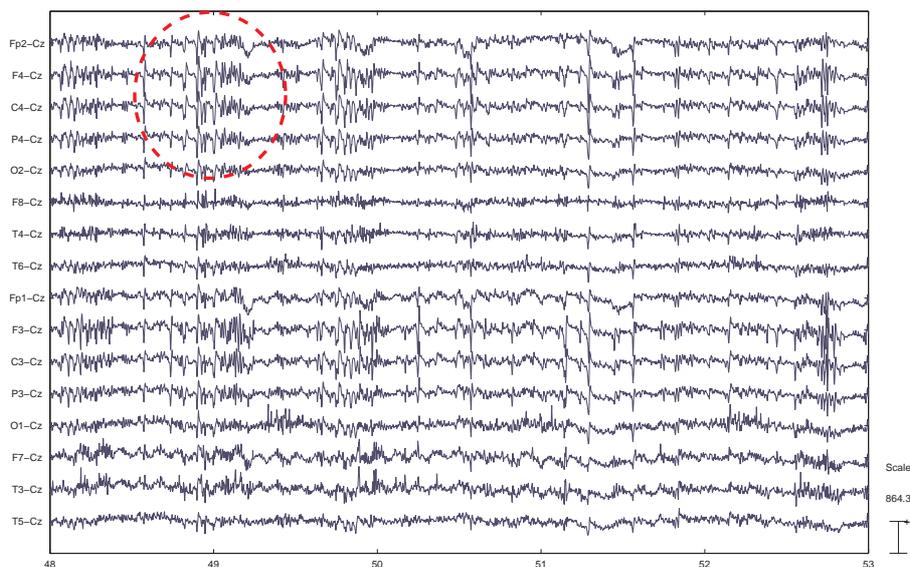


Figura 2.5: Exemplos de sinais EEG com ondas abruptas

## 2.4 Artefatos EEG

Toda a actividade eléctrica proveniente de um sinal EEG que não é de origem cerebral é considerada artefacto. Os artefactos podem ser de dois tipos, os de origem fisiológica e os de origem não fisiológica. Os artefactos de origem fisiológica podem ser resultado de actividades musculares (figura 2.6), de movimentos oculares (figura 2.7), EEG, *Electro-oculograma* (EOG) etc.

Os artefactos causados por actividades musculares são frequentes em gravações EEG, aparecem com elevado potencial, sendo frequentes nos eléctrodos frontais e temporais quer de uma aquisição bipolar, quer de uma aquisição segundo um referencial [9]. Embora os artefactos musculares nunca possam ser totalmente eliminados, eles podem ser fortemente atenuados com o uso de um filtro que limita a largura de banda no entanto corre-se o risco de poderem ser confundidos com a actividade Beta [3].

Contudo os potenciais gerados no cérebro são de maior duração do que os gerados nos músculos, com isto é possível que uma distinção entre os artefactos resultantes de actividade muscular em relação actividade cerebral, para além da duração o tempo, a morfologia e a frequência são factores igualmente distinguíveis.

Os artefactos ECG correspondem aos artefactos resultantes da actividade eléctrica do coração. Estes artefactos não podem ser facilmente removidos por filtragem, mas podem ser diferenciados de actividade EEG pois o período do sinal EEG corresponde perfeitamente ao período do sinal *Electrocardiograma* (ECG), verificando-se assim uma sincronização com cada complexo QRS do sinal ECG.

Os movimentos oculares dão origem frequentemente a ondas EEG lentas, estes artefactos aparecem mais predominantemente nos canais frontais.

No entanto não existem só estes tipos de artefactos, podemos encontrar artefactos resultantes da respiração, artefactos resultantes de movimentos linguísticos, entre outros.

Os artefactos de origem não fisiológica podem resultar de eléctrodos não correctamente colocados, de interferência electromagnética.

Aprender as características dos artefactos é fundamental para um médico como para um detector de crises, pois os artefactos podem ser confundidos com uma actividade de crise [3].

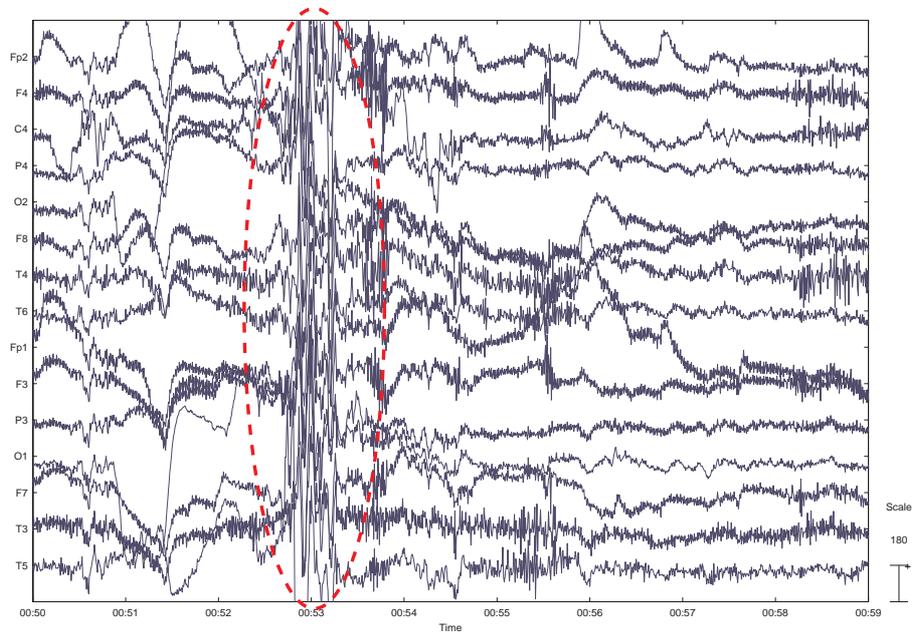


Figura 2.6: Exemplos de sinais EEG com artefactos musculares

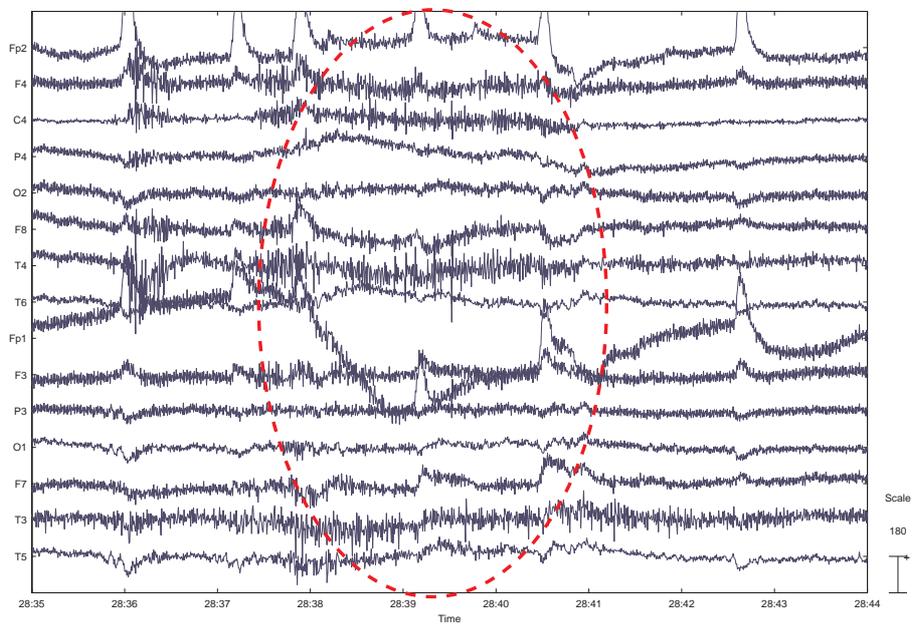


Figura 2.7: Exemplos de sinais EEG com artefactos oculares



## Capítulo 3

# Sistemas Automáticos

O diagnóstico e caracterização de síndromes epilépticas envolveram ao longo dos anos estudos clínicos que exigiam uma interacção prolongada com o paciente. As gravações EEG que podem durar dias ou até semanas são fundamentais para a captação de actividade epiléptica, daí o desenvolvimento de algoritmos que visam a detecção e caracterização de actividade anormal em sinais EEG.

Encontrar actividade epiléptica, em registos de sinais EEG que duram dias ou semanas é uma tarefa ingrata e demorada, já que esta actividade constitui frequentemente uma percentagem muito pequena da gravação. Esta dificuldade tem motivado o desenvolvimento de sistemas automáticos que percorrem o sinal, identificando se há, ou não ocorrência de eventos epilépticos.

Várias metodologias e algoritmos foram propostos desde 1976, com a publicação das primeiras abordagens de Gotman e Gloor (1976)[10], na tentativa de tornar automática a detecção de actividade epiléptica. Os autores projectaram e executaram um dos primeiros sistemas para a identificação de actividade epiléptica. São apresentados muitos exemplos demonstrando a detecção de eventos epilépticos e a rejeição de artefactos, mas não realizam uma avaliação sistemática do método.

O mesmo Gotman uns anos mais tarde desenvolveu um dos sistemas mais avançados para a detecção automática de crises de longa duração em sinais EEG [11].

O sistema foi testado em 16 sinais com um conjunto médio de 12.4 horas e exibiu uma variabilidade considerável no seu desempenho. Gotman relatou que 22% de todas as detecções tiveram origem em descargas epilépticas, 58% de descargas não epilépticas e 20% ocorreram devido a artefactos.

Os sistemas desenvolvidos, tal como qualquer sistema de classificação, podem ser divididos em dois blocos, um de extracção de características e outro de classificação. Para qualquer um dos blocos podem ser encontradas diferentes soluções.

### 3.1 Extracção de características

A extracção de características do sinal EEG, tem sofrido uma evolução acentuada ao longo dos anos. Na generalidade dos trabalhos analisados, a extracção de características é efectuada de forma sequencial. Primeiramente o sinal é dividido em subsegmentos de pequenas amostras, de onde posteriormente são extraídas as características.

Em 1997 Qu e Gotman [12], projectaram um classificador para detectar inícios de crise epiléptica. Os autores propuseram a extracção de características de sinal EEG baseando-se em técnicas no domínio do tempo e da frequência. Foram seleccionadas características como, amplitude, média, coeficiente de variação do sinal, entre outras.

Alguns dos primeiros estudos, centraram-se na detecção de ondas características como spikes, sharp waves, entre outras.

Em 1998 Tarassenko e outros autores [13], apresentaram uma metodologia para a classificação de spikes, baseada igualmente em técnicas no domínio do tempo e da frequência.

Os autores optam, no entanto pela utilização de características como mobilidade, complexidade, inclinação média (Average slope), Agudeza (Sharpness), entre outras.

Com a evolução dos anos a detecção de anormalidade de sinal EEG tornou-se mais generalizada, optando por não centrar o estudo em ondas características, mas sim, em crises epilépticas.

Trabalhos mais recentes [3], sugerem a utilização da TDW. A decomposição em wavelet reflecte simultaneamente características tempo e frequência do sinal em análise. No resultado da decomposição de cada segmento de sinal são extraídas diversas medidas. Em 2004 [3], os autores sugerem a energia do sinal como característica relevante. Esta abordagem é também utilizada em 2005 [14], seleccionando como características, a média do sinal, o desvio padrão, entre outras.

Em 2006 [15], os autores optam pela amplitude máxima e mínima do sinal, desvio padrão e média como características a utilizar. Na realização deste trabalho optou-se igualmente pela utilização da TDW, seleccionando como características a energia média, a amplitude máxima e mínima do sinal, e o desvio padrão.

## 3.2 Classificação

No processo de classificação, são apresentados uma grande variedade de tipos de algoritmos. Qu e Gotman, em 1997, num estudo [12], sugerem a elaboração de um algoritmo baseado no classificador *nearest-neighbor* (vizinho mais próximo) no apoio a detecção de anormalidades no sinal EEG. No trabalho realizado, os autores apresentam uma taxa de 100% na detecção de inícios de crises epilépticas, com uma percentagem de falsos positivos igual a 0.02, resultante de uma série de dados elaborada com 12 pacientes e 47 crises. Os autores demonstram um desempenho menos favorável quando o algoritmo foi treinado com um sinal anormal de um paciente e testado com um sinal normal de outro paciente.

Tarassenko e outros autores em 1998, num estudo [13], defenderam que a utilização de redes neuronais são eficazes no apoio à detecção de spikes num sinal EEG.

Os autores apresentaram uma sensibilidade entre 83%-97% e uma especificidade entre 85%-95% quando o seu classificador foi treinado e testado em várias séries de dados de pacientes. O classificador apresentou uma sensibilidade semelhante quando treinado e experimentado em conjuntos de dados que combinavam muitos pacientes, embora seja apresentado um grande número de falsas detecções, muito devido à presença de diversos artefactos.

Os autores ainda apresentaram os seus pontos de vista sobre o processo de detecção, no entanto para eles um processamento mais sofisticado não reduziria os falsos positivos significativamente.

Já no século XXI Ali Shoeb [3] publica um artigo no qual apresenta um método para a de-

tecção de anormalidade de sinal EEG baseado no uso de um classificador do tipo SVM. No seu artigo apresenta resultados, em que chega a detectar 131 de 139 crises em 60 horas de sinal. Este método é aquele que mais se identifica com o trabalho realizado, pois parte dos mesmos princípios.

Em 2005, Inan Guler e Elif Derya Ubeyli [16], propuseram uma classificação baseada na aplicação de um sistema *neuro-fuzzy*. Este sistema foi treinado com o algoritmo *backpropagation*. No processo de classificação, foram utilizados 5 tipos de sinais EEG, que funcionaram como testes padrões à entrada do classificador. Para melhorar a exactidão, os autores introduziram um sexto classificador, treinado com as saídas dos outros cinco classificadores. Os resultados obtidos apresentaram uma precisão de 98,68%.

No mesmo ano, Subasi [14] publica um estudo baseando-se em técnicas recentes de redes neuronais, tendo em vista à detecção de anormalidades de sinal EEG, onde apresenta resultados de 93.1% de especificidade e 92.8% de sensibilidade.



## Capítulo 4

# Sistema baseado em classificador binário

O diagnóstico e caracterização de síndromes epilépticos envolvem estudos clínicos que exigem uma iteração prolongada com o paciente. Os sinais EEG são registados por largos períodos de tempo (dias ou mesmo semanas). Como forma de apoiar este tipo de abordagem exaustiva, procedeu-se a elaboração de um algoritmo de apoio ao diagnóstico.

Este algoritmo visa uma caracterização do sinal EEG classificando-o em epiléptico ou não epiléptico.

O processo de classificação passa inicialmente por uma extracção das características. Esta extracção tem como objectivo desenvolver procedimentos para se obter informações subjacentes aos sinais EEG.

Para a classificação automática é necessário extrair várias medidas de diferentes sinais EEG. Estas medidas são organizadas num vector que servirá de entrada para o classificador.

O diagrama da figura 4.1 elucida os passos necessários para a classificação de um determinado segmento de sinal EEG.

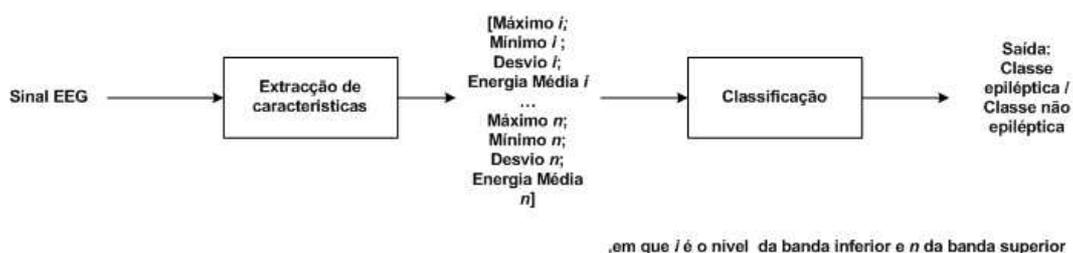


Figura 4.1: Diagrama de blocos do problema de classificação de um sinal EEG

O estudo passa assim por uma análise de um segmento de sinal EEG, extraindo características relevantes do ponto de vista epiléptico, seguindo-se uma classificação. A classificação é um processo de atribuir a classe a que pertence o segmento.

## 4.1 Extracção de características

O processo de extracção e selecção de características, tem uma grande importância no desempenho do classificador. Por isso uma selecção antecipada das melhores características que caracterizam o problema é fundamental para se obter uma classificação eficiente. Este processo pode ser dividido em dois blocos (figura 4.2). O primeiro bloco (Transformada Discreta Wavelet) trata da decomposição simultaneamente no tempo e na frequência. O segundo bloco (Vector de Características) é uma organização das características extraídas de cada subsegmento de sinal EEG.

Para a extracção de características, é fundamental a escolha de medidas com uma relação directa na variação de sinal, pois uma das características que distingue o sinal epiléptico de um sinal normal é uma oscilação brusca da amplitude.

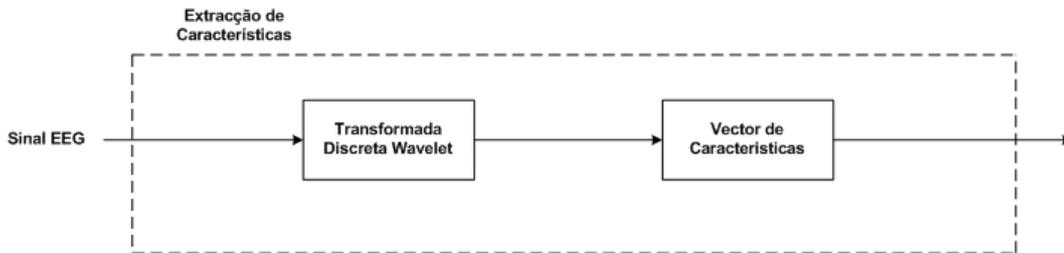


Figura 4.2: Diagrama de blocos do bloco de extracção de características

### 4.1.1 Transformada discreta de Wavelet

A TDW é uma ferramenta versátil de tratamento de sinal que podemos encontrar em muitas aplicações científicas [15].

Uma das áreas em que a TDW é utilizada é no apoio à extracção de características num sinal EEG, pois permite uma decomposição do sinal no tempo e na frequência.

A figura 4.4 elucida o tipo de análise ao sinal que se pretende, ou seja, a TDW analisa o sinal em gamas de frequência diferentes.

A decomposição do sinal nos diferentes níveis de frequência é obtida pela filtragem consecutiva do sinal utilizando para isso um filtro passa-baixo  $h_0(z)$  e um filtro passa-alto  $h_1(z)$  (figura 4.3).

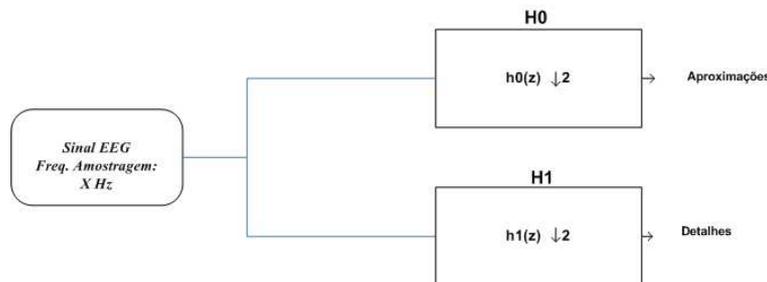


Figura 4.3: Estrutura dos filtros com sub amostragem

A utilização recursiva do bloco da figura 4.3 à saída do bloco passa-baixo da decomposição Wavelet, dá origem à estrutura representada na figura 4.4. Na figura estão referidos os conteúdos de frequência do sinal original que cada saída representa.

Considerando um sinal com uma frequência de amostragem igual a  $F_s$ , à saída  $d_1$  o sinal encontra-se situado entre o intervalo de frequência  $F_s/4$  a  $F_s/2$ . Para a saída  $d_2$  o sinal encontra-se situado entre o intervalo de frequência  $F_s/8$  e  $F_s/4$ , e assim sucessivamente para as outras saídas  $d$ .

A letra  $H_0$  representa um filtro passa-baixo aliado à sub amostragem de 2, a letra  $H_1$  representa um filtro passa-alto aliado à sub amostragem 2 (figura 4.3).

Neste tipo de decomposição de sinais, é usual utilizar filtros do tipo *Daubechies*. Na lit-

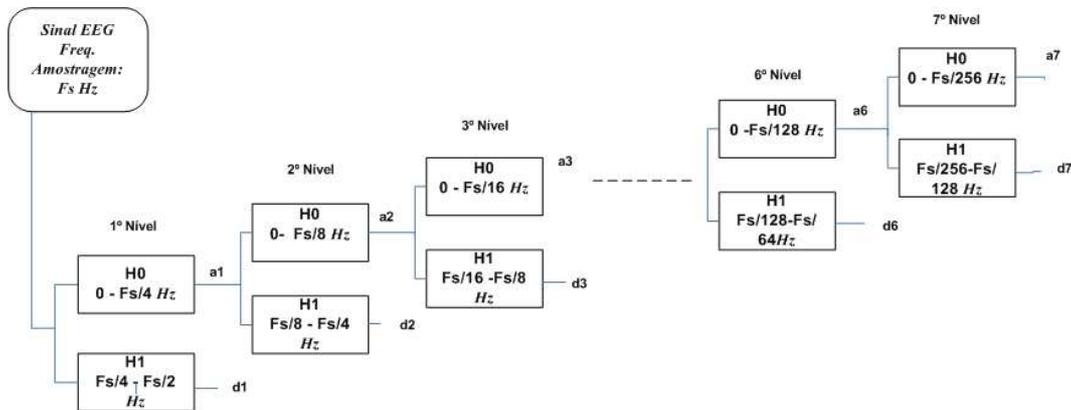


Figura 4.4: Decomposição em árvore

eratura, o filtro do tipo *Daubechies* (daub) de ordem 2 é o mais apropriado para detectar mudanças de sinais EEG [15] [16], no entanto há referências onde são utilizadas outras ordens.

#### 4.1.2 Implementação da TDW

Para apoiar na resolução da TDW, utilizaram-se algumas funções da toolbox Uviwave [17], desenvolvida na Universidade de Vigo.

Funções da toolbox Uviwave utilizadas:

- Função  $wt$   
Esta função tem como objectivo calcular a TDW.

$$wx = wt(X, h0, h1, K).$$

Parâmetros de entrada:

- $X$ : É o sinal de entrada
- $h_0$ : É o filtro passa-baixo
- $h_1$ : É o filtro passa-alto
- $K$ : Simboliza o número de níveis da decomposição

Parâmetros de saída:

- $wx$ : É o resultado da função  $wt$ , apresenta-se como um vector com estrutura idêntica à da figura 4.5. Note-se que para um sinal com  $N$  amostras, o vector  $wx$  tem também  $N$  amostras.  
Assim só é possível a distribuição das amostras pelas bandas indicada na figura 4.5 quando  $N$  é uma potência 2.

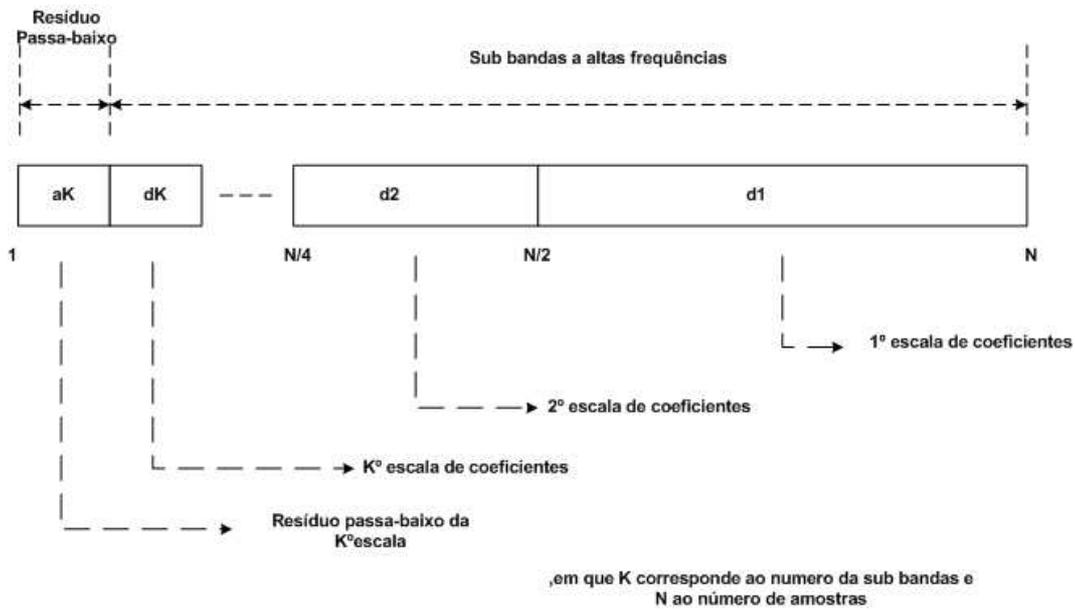


Figura 4.5: Estrutura do resultado da função  $wt$

- Função *bandsite*

Esta função permite a extração das sub-bandas pretendidas, através da posição dos índices.

$$[X1, Y1, X2, Y2] = \text{bandsite}(wx, K, Q, BT)$$

Parâmetros de entrada:

- $wx$ : É a saída da função  $wt$
- $K$ : Simboliza o número de níveis da decomposição efectuada
- $Q$ : Corresponde ao nível em análise
- $BT$  0: Baixas frequências 3: Altas frequências

Parâmetros de saída:

- $(X1, Y1)$ : Coordenada da parte direita da TDW, neste estudo como sinal é de uma dimensão o valor de  $X1=1$ , variando apenas o valor de  $Y1$

- $(X2, Y2)$ : Coordenada da parte esquerda da TDW, neste estudo como sinal é de uma dimensão o valor de  $X2=1$ , variando apenas o valor de  $Y2$

- Função *daub*

Esta função retorna os coeficientes de um filtro do tipo *Daubechies*.

$$[h, g, rh, rg] = daub(num\_coeffs)$$

Parâmetros de entrada:

- *num\_coeffs*: É o número de coeficientes desejados para o filtro

Parâmetros de saída:

- *h*: Filtro de análise passa-baixo
- *g*: Filtro de análise passa-alto
- *rh*: Filtro de síntese passa-baixo
- *rg*: Filtro de síntese passa-alto

### 4.1.3 Vector de Características

Os sinais provenientes das bandas de interesse não são usados directamente como entradas no classificador. Primeiramente extraem-se as características de cada banda de interesse, procedendo-se posteriormente à organização de um **Vector de características**:

$$\mathbf{X}_i = \begin{bmatrix} Maximo_i \\ Minimo_i \\ Desvio_i \\ EnergiaMedia_i \\ \vdots \\ Maximo_n \\ Minimo_n \\ Desvio_n \\ EnergiaMedia_n \end{bmatrix}$$

, em que  $i$  é o nível da banda inicial desejada e  $n$  é o nível da banda final desejada.

No entanto este **Vector de características** também foi elaborado a seguinte forma:

$$\mathbf{X}_i = \begin{bmatrix} EnergiaMedia_i \\ \vdots \\ EnergiaMedia_n \end{bmatrix}$$

, em que  $i$  corresponde igualmente ao nível da banda inicial desejada e  $n$  é o nível da banda final desejada.

## 4.2 Classificação

Classificação traduz-se num processo no qual se atribui uma etiqueta (classe) a um objecto caracterizado por um conjunto de parâmetros característicos (features).

Os parâmetros do classificador são determinados em função de um conjunto de treino, sendo portanto a elaboração deste determinante para o processo de classificação. A classificação em duas classes é neste contexto um problema binário, em que a etiqueta a atribuir a um objecto tem apenas valores binários (0/1 ou -1/1).

O processo de classificação foi então dividido em duas etapas, uma trata da organização do conjunto de treino e do teste, e a outra do treino e do teste do classificador (figura 4.6).

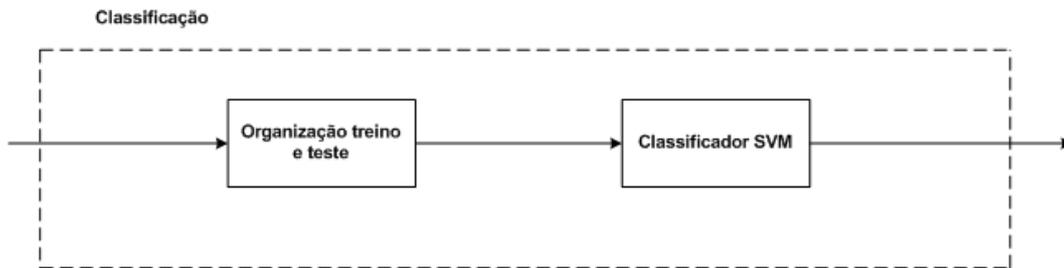


Figura 4.6: Diagrama de blocos do bloco do processo de Classificação

### 4.2.1 Classificação com SVM

Nesta secção apresentam-se os princípios de funcionamento dos classificadores designados por SVM.

Existem contudo muitos tipos de classificadores lineares que possibilitam a separação dos dados, no entanto a escolha recaiu sobre o SVM. Sendo que uma das vantagens relativamente a outros tipos de classificadores a maximização da margem, isto é, a distância entre o hiperplano de separação das classes e os dados das duas classes mais próximas [18].

#### Classificação em duas classes

Assumindo que temos um objecto  $\mathbf{k}$  que está descrito pelo vector das características  $\mathbf{x}_{\mathbf{k}}$ , no classificador caracterizado pelos parâmetros  $\mathbf{w}$  e  $b$  é efectuada a seguinte função:

$$f(\mathbf{x}_{\mathbf{k}}) = \mathbf{w} \cdot \mathbf{x}_{\mathbf{k}} + b \quad (4.1)$$

, onde  $\cdot$  é o produto interno entre os vectores  $\mathbf{w}$  e  $\mathbf{x}_{\mathbf{k}}$ . O valor da função anterior determina a classe a que  $\mathbf{x}_{\mathbf{k}}$  pretence, então:

$$\begin{cases} f(\mathbf{x}_{\mathbf{k}}) > 0 & \text{classe} + 1 \\ f(\mathbf{x}_{\mathbf{k}}) < 0 & \text{classe} - 1 \end{cases} \quad (4.2)$$

## Interpretação Geométrica

A figura 4.7 ilustra um conjunto de dados de dimensão 2. Uma classe corresponde aos círculos, enquanto que a outra corresponde às cruces.

Estes dados são linearmente separáveis, ou seja, é possível encontrar um hiperplano de separação entre dados de ambas as classes.

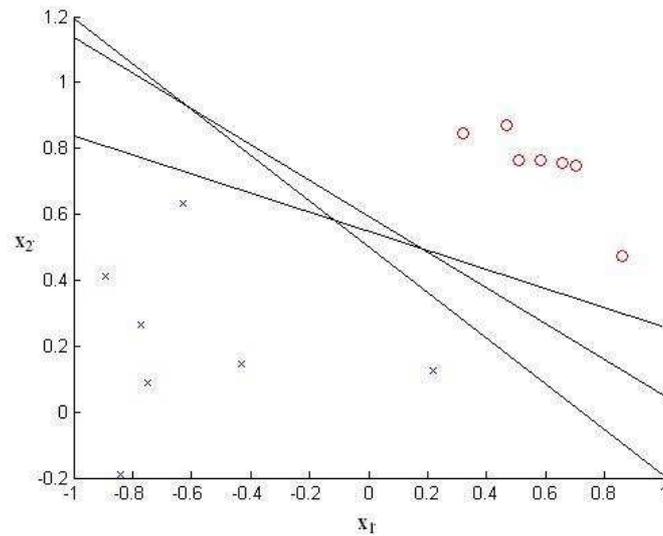


Figura 4.7: Divisão dos dados segundo diversos hiperplanos

No entanto há um número infinito de hiperplanos que seriam possíveis, tal como ilustra a figura 4.7.

Num classificador binário a posição do hiperplano é determinada pelos parâmetros  $\mathbf{w}$  e  $b$ .

Na figura 4.8 podemos observar o hiperplano (traço negro escuro), onde o vector  $\mathbf{w}$ , é perpendicular a este, enquanto que  $b$  determina a posição espacial do hiperplano.

Para qualquer valor  $\mathbf{x}$  no hiperplano  $H$  pode escrever-se a seguinte equação:

$$\mathbf{w} \cdot \mathbf{x} + b = 0 \quad (4.3)$$

O hiperplano de decisão  $H$  fica situado entre os hiperplanos  $H1$  e  $H2$ . Estes hiperplanos são paralelos e podem ser definidos pelas seguintes expressões:

$$\begin{aligned} H1 : \mathbf{w} \cdot \mathbf{x}_i + b &= +1 \\ H2 : \mathbf{w} \cdot \mathbf{x}_i + b &= -1 \end{aligned} \quad (4.4)$$

Considera-se um ponto ( $\mathbf{x}_0$ ) situado no hiperplano  $H$  e um ponto ( $\mathbf{x}_1$ ) situado no plano  $H1$ . Deste modo, podem ser escritas as seguintes igualdades:

$$\begin{aligned} \mathbf{w} \cdot \mathbf{x}_0 + b &= 0 \\ \mathbf{w} \cdot \mathbf{x}_1 + b &= -1 \end{aligned} \quad (4.5)$$

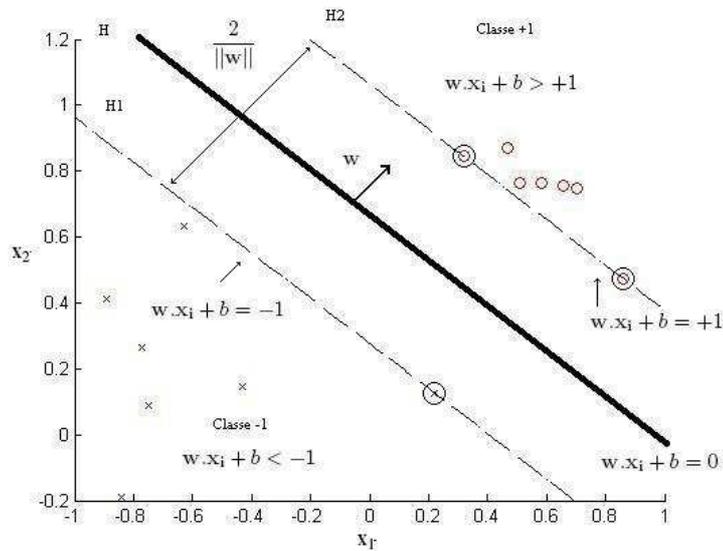


Figura 4.8: Definição dos hiperplanos

, subtraindo a segunda equação pela primeira obtemos a seguinte:

$$\mathbf{w} \cdot (\mathbf{x}_0 - \mathbf{x}_1) = 1 \quad (4.6)$$

Resolvendo a equação calcula-se a distância entre o hiperplano H e o hiperplano H1:

$$\|(\mathbf{x}_1 - \mathbf{x}_0)\| = \frac{1}{\|\mathbf{w}\|} \quad (4.7)$$

Procedendo de forma similar e utilizando o hiperplano H em relação ao hiperplano H2 com um ponto ( $\mathbf{x}_2$ ) situado nesse mesmo hiperplano, acabar-se-á por obter a mesma distância, chegando assim à seguinte equação:

$$\|(\mathbf{x}_0 - \mathbf{x}_2)\| = \frac{1}{\|\mathbf{w}\|} \quad (4.8)$$

Pode-se dizer então que através da equação dos dois hiperplanos é possível definir a margem do classificador, ou seja, a margem que delimita o conjunto de hiperplanos, sendo essa margem dada pela distância entre os dois hiperplanos  $H1$  e  $H2$ :

$$\frac{2}{\|\mathbf{w}\|} \quad (4.9)$$

Os vectores de suporte encontram-se portanto a  $1/\|\mathbf{w}\|$  do limite de decisão( $H$ ), isto é, nos hiperplanos paralelos. Estes exercem uma papel fundamental na operação de aprendizagem, já que têm influência directa na localização da superfície de separação.

Para o hiperplano H que define o classificador linear ( onde  $\mathbf{w} \cdot \mathbf{x} + b = 0$  ), a distância entre a origem e o hiperplano é:

$$\frac{|b|}{\|\mathbf{w}\|} \quad (4.10)$$

Deste modo é possível traçar um hiperplano de separação dos dados de treino. No entanto o SVM é caracterizado por apresentar diversas vantagens especialmente no que diz respeito a maximização da margem, ou seja, maximizar a distância dos dados de treino, havendo assim inúmeros hiperplanos possíveis com diversas distâncias de margem. Sendo que o hiperplano óptimo corresponde a uma igualdade de distância do hiperplano em relação aos dados positivos e dados negativos.

A diferença entre o SVM linear e outros classificadores baseados em hiperplanos é que este permite uma flexibilidade da margem.

## Aprendizagem

Considerando um teste  $(\mathbf{x}_i, y_i)$  em que  $\mathbf{x}_i$  é um vector de características e  $y_i$  a classe a que pretence.

A fase de formação do modelo SVM envolve assim estimar os parâmetros  $\mathbf{w}$  e  $b$  para isso devem ser escolhidos de forma a verificar as seguintes condições:

$$\begin{aligned} \mathbf{w} \cdot \mathbf{x}_i + b &\geq +1 & se & y_i = +1 \\ \mathbf{w} \cdot \mathbf{x}_i + b &\leq -1 & se & y_i = -1 \end{aligned} \quad (4.11)$$

Embora a condição anterior seja aplicada a qualquer classificador linear, o SVM impõe o requisito adicional definindo uma margem do seu limite de decisão.

A forma de aprendizagem no classificador SVM pode ser resumida a um problema de optimização, em que se maximiza a margem, através da minimização de 4.12, com as restrições adjacentes 4.13. Simplificando a equação 4.11 :

$$\text{minimizando } f(\mathbf{w}) = \frac{\|\mathbf{w}\|^2}{2} \quad (4.12)$$

$$\text{sujeita a } y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1, \quad i = 1, 2, \dots, N \quad (4.13)$$

O problema de optimização das equações 4.12 e 4.13 representa assim a minimização da função quadrática sob restrições lineares.

No entanto esta equação não representa o problema real da optimização, para a descoberta do melhor hiperplano. Baseado no uso da função de Lagrange, as equações 4.12 e 4.13 passam a ter uma transformação dupla, sendo uma forma conveniente de resolver o problema de minimização.

Recorrendo então às funções de Lagrange [19]:

$$Lp(\mathbf{w}, b, \mathbf{\Lambda}) = \frac{\|\mathbf{w}\|^2}{2} - \sum_{i=1}^n \lambda_i (y_i(\mathbf{w} \cdot \mathbf{x}_i + b) - 1) \quad (4.14)$$

, em que  $\mathbf{\Lambda}$  corresponde aos multiplicadores de Lagrange ( $\mathbf{\Lambda} = (\lambda_1, \lambda_2, \dots, \lambda_n)$ ) e  $n$  corresponde ao tamanho do conjunto de treino.

O primeiro termo representa a função  $f(\mathbf{w})$ , enquanto que o segundo termo representa as restrições impostas onde os multiplicadores de Lagrange  $\lambda_i \geq 0$ . Desta forma não há qualquer

tipo de infracção, podendo somente aumentar o valor da função de Lagrange.

Para minimizar o Lagrangiano, deve-se ajustar a derivada de  $Lp$  em relação a  $\mathbf{w}$  e a  $b$  igualando a zero [19]. Trata-se de uma forma de lidar com as desigualdades das restrições, transformando-as, num conjunto de igualdades.

A alternativa passa pela utilização das restrições *Karush-Kuhn-Tucker* (KKT) [19, 18], passando-as para a forma dual:

$$Ld(\Lambda) = \sum_{i=1}^n \lambda_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j y_i y_j \mathbf{x}_i \cdot \mathbf{x}_j, \quad (4.15)$$

Desta forma o modelo SVM passa a ser definido por uma relação entre os vectores de suporte e o  $\lambda_i$  correspondente, passando o  $\mathbf{w}$  a ser definido por:

$$\mathbf{w} = \sum_{i=1}^n \lambda_i y_i \mathbf{x}_i \quad (4.16)$$

$$\sum_{i=1}^n \lambda_i y_i = 0 \quad (4.17)$$

Podendo ser encontradas duas situações distintas:

$$\begin{aligned} \lambda_i &> 0 \\ \lambda_i &= 0 \end{aligned}$$

Os vectores de características  $\lambda_i = 0$  podem ser encontrados na parte externa aos hiperplanos  $H1$  ou  $H2$ . Enquanto que os vectores de características  $\lambda_i > 0$  situam-se exactamente num dos hiperplano  $H1$  ou  $H2$  designado-se por vectores de suporte.

Assim, o modelo SVM, pode ser caracterizado por  $\lambda_i > 0$ , isto no teste, pois todos os elementos  $\lambda_i = 0$ , podem ser retirados da equação 4.16, resultando daí a seguinte equação:

$$\mathbf{w} = \sum_{i=1}^{ns} \lambda_i y_i \mathbf{x}_i \quad \text{com } ns < n \quad (4.18)$$

Para calcular o valor de  $b$  do hiperplano, consideramos as condições KKT acoplada à expressão anterior e com a condição  $\lambda_i \geq 0$  obtemos a seguinte expressão:

$$b = y_i - \sum_{i=1}^n \sum_{j=1}^n \lambda_i y_i \mathbf{x}_i \cdot \mathbf{x}_j \quad (4.19)$$

### Classes não linearmente separáveis

Na secção anterior os dados apresentados eram linearmente separáveis, ou seja, os dados de treino foram separados correctamente através de um hiperplano de separação.

Há muitas razões pelas quais um conjunto de treino é não linearmente separável.

Na figura 4.9, é apresentado um problema de classificação não linear, permitindo que hajam objectos mal classificados entre os hiperplanos  $H1$  e  $H2$  e/ou  $H3$  e  $H4$  havendo desta forma uma relação entre maximização das margens e objectos situados entre as mesmas.

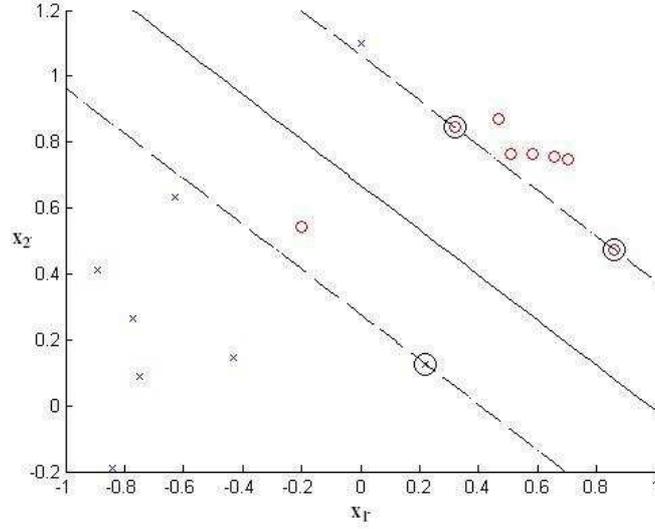


Figura 4.9: Dados linearmente não linearmente separáveis

Quando se trata deste tipo de classes (classes não linearmente separáveis), os métodos utilizados para classes linearmente separáveis, não produzirão um solução viável na obtenção de um hiperplano óptimo de separação. No entanto as equações definidas para classes linearmente separáveis podem ser adaptados para classes não linearmente separáveis.

O problema de otimização é formulado de modo a que no interior dos hiperplanos  $H1$  e  $H2$  possam existir exemplares do conjunto de treino. A formulação é dada pela equação:

$$f(\mathbf{w}) = \frac{\|\mathbf{w}\|^2}{2} + C \left( \sum_{i=1}^n \xi_i \right)^k. \quad (4.20)$$

Os parâmetros  $C$  e  $\xi$  representam então um compromisso entre o número de padrões  $\mathbf{x}$  que podem estar no interior da margem e a maximização da respectiva largura (figura 4.10). Recorrendo à formulação dual [19] o problema reduz-se à equação:

$$Ld(\mathbf{\Lambda}) = \sum_{i=1}^n \lambda_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j y_i y_j \mathbf{x}_i \cdot \mathbf{x}_j, \quad \text{com } 0 \leq \lambda_i \leq C \quad (4.21)$$

Esta equação é muito semelhante à equação obtida em 4.16. A única diferença é que os  $\lambda_i$  tem agora valor inferior a  $C$ . As figuras 4.11 e 4.12 para dois valores de  $C$ , permitem verificar que a margem tem largura maior ou menor consoante o valor de  $C$  escolhido para o processo de aprendizagem.

#### 4.2.2 Considerações finais sobre SVM

Após a determinação dos parâmetros ( $\mathbf{w}$  e  $b$ ) do SVM, para cada novo padrão  $\mathbf{x}_j$  tem-se:

$$\begin{aligned} f(\mathbf{x}_j) &= \mathbf{w} \cdot \mathbf{x}_j + b \\ &= \sum_{i=1}^n \lambda_i y_i (\mathbf{x}_i \cdot \mathbf{x}_j) + b \end{aligned} \quad (4.22)$$

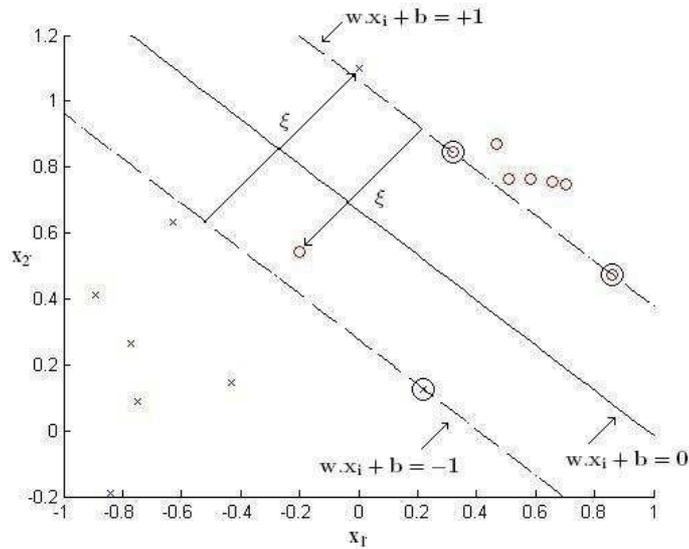


Figura 4.10: Hiperplano de separação de dados não linearmente separáveis

A equação anterior envolve o produto interno  $\mathbf{x}_i$  e  $\mathbf{x}_j$ . O produto interno pode ser calculado utilizando as funções de kernel[18]. As funções de kernel permitem calcular  $k(\mathbf{x}_i, \mathbf{x}_j) = \phi^T(\mathbf{x}_i)\phi(\mathbf{x}_j)$ . Portanto permitem calcular em simultâneo com o produto interno a transformação  $\phi$ .

No caso de ser o kernel linear a operação de produto interno é feita sem transformação, isto é:

$$k(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i \cdot \mathbf{x}_j \quad (4.23)$$

Note-se que no caso de utilizar outro tipo de funções de kernel (por exemplo *Radial Basis Function* (RBF)) as superfícies de separação deixam de ser hiperplanos[18].

A classificação é feita da seguinte forma:

$$\text{sign}\left(\sum_{i=1}^n \lambda_i y_i k(\mathbf{x}_i, \mathbf{x}_j) + b\right) \quad (4.24)$$

Assim sendo se a saída for igual a  $-1$  então o ponto de teste encontra-se na região  $-1$ , se a saída for  $+1$ , a saída encontra-se na região  $+1$ , caso o ponto de teste esteja correctamente classificado.

### 4.2.3 Organização do Treino e do Teste

Na organização do treino e teste, procedeu-se à selecção de alguns vectores de características. O objectivo passa por dividir os dados (conjunto total de vectores de características) em partições (treino e teste). Sendo essa divisão feita de forma aleatória.

Para a organização do conjunto de treino estabelece-se um valor para a quantidade de vectores de características desejados, estes são escolhidos de forma aleatória do conjunto total de vectores de características, enquanto que os restantes dados serão colocados no conjunto de teste.

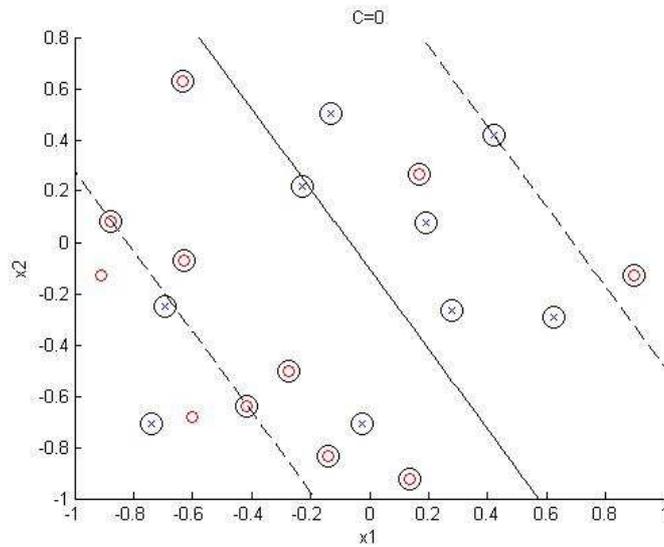


Figura 4.11: Aprendizagem considerando  $C=1$

A escolha aleatória do conjunto de treino é portanto preponderante, já que por cada escolha aleatória é possível a formação de um classificador com parâmetros diferentes. Assim sendo, uma formação de diversos conjuntos de treino permite um estudo do desempenho do classificador.

#### 4.2.4 Implementação do classificador

No processo de classificação recorreu-se ao uso de varias funções provenientes da toolbox `sprtool` [20]. Trata-se de uma toolbox com alguns algoritmos de apoio no reconhecimento de padrões. Funções da toolbox `sprtool` utilizadas:

- Função `evalsvm`  
Esta função avalia o classificador SVM. Usa o *cross-validation* para estimar os melhores valores do parametro de entrada `options`.

$$[model, Errors] = evalsvm(trn, options)$$

Parâmetros de entrada:

- `trn`: Conjunto de treino constituído por duas classes
- `options`:
  - \* `options.ker`: 'linear' ou 'rbf'
  - \* `options.C`: [1 10 100] (os valores foram os utilizados)
  - \* `options.solver`: 'svm-light' ou 'smo'

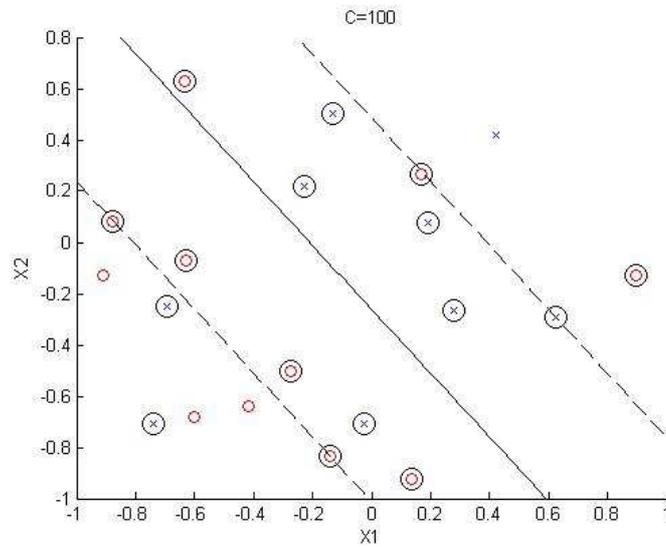


Figura 4.12: Aprendizagem considerando  $C=100$

\* *options.arg*: [0.1 0.5 1 5] (parâmetros só utilizado quando o kernel não for linear)

Parâmetros de saída:

- *model*: Modelo em análise
- *Errors*: Erro associado ao modelo em análise (erro estimado através de *cross – validation*).

Assim, optou-se pela utilização de um Kernel linear com o parâmetro  $C=100$ .

- Função *smo*

Esta função é designada por *Sequential Minimal Optimization*. Trata-se de um algoritmo de aprendizagem para SVM binário.

$$model = smo(data, options)$$

Parâmetros de entrada:

- *data* Conjunto de treino
- *options* Conjunto de opções já definidas na função *evalsvm*

Parâmetros de saída:

- *model*: Conjunto de parametros que permitem a elaboração de um hiperplano de separação entre os dados de uma classe e outra.

- Função *svmlight*  
Trata-se de um algoritmo de aprendizagem para SVM binário baseado nos mesmos princípios da função *smo*.

$$model = svmlight(data, options)$$

Parâmetros de entrada:

- *data*: Conjunto de treino
- *options*: Conjunto de opções já definidas na função *evalsvm*

Parâmetros de saída:

- *model*: Conjunto de parametros que permitem a elaboração de um hiperplano de separação entre os dados de uma classe e outra.

- Função *svmclass*  
Esta função é designada por *Support Vector Machines Classifier*. Estabelece a classificação segundo o modelo elaborado no algoritmo de aprendizagem.

$$[y, dfce] = svmclass(X, model)$$

Parâmetros de entrada:

- *X*: Dados de teste
- *options*: Parâmetros do modelo elaborado no algoritmo de aprendizagem

Parâmetros de saída:

- *y*: Etiquetas atribuídas no processo de classificação
- *dfce*: Valor da função discriminante

#### 4.2.5 Software desenvolvido

Todo o software desenvolvido, resulta de uma aplicação de funções implementadas em ambiente matlab [21]. No Apêndice A encontra-se o código associado a cada função.

- Função *eeg\_extracao\_funcao\_Classe* (ver apêndice A1)  
Esta função tem como objectivo extrair as características do sinal em análise, elaborando desta forma uma matriz contendo os vectores de características de cada subsegmento.

$$[A] = eeg\_extracao\_funcao\_Classe(ficheiro, niveis, L, niveis\_d, ordem\_filtro, tipo, energia)$$

Parâmetros entrada:

- *ficheiro*: Sinal de entrada

- *L*: Número de divisões desejadas para cada segmento de sinal
- *niveis*: Número de níveis desejados para a decomposição
- *niveis\_d*: Quantidade de ultimos níveis desejáveis
- *ordem\_filtro*: Ordem do filtro daub
- *tipo*: Tipo de classe pertencente (1 ou 2)
- *energia*:
  - \* 0: Para a extracção das 4 características
  - \* 1: Para a extracção exclusiva da energia média

Parâmetros saída:

- *A*: matriz contendo os vectores de características do segmento em análise
- Função *eeg gerar\_treino\_testes\_def* (ver apêndice A2)  
Esta função destina-se organizar o conjunto de teste e o conjunto de treino, seleccionados de forma aleatória .

$$[treino\_def \quad teste\_def \quad qt\_tipos3 \quad qt\_tipos4] = eeg\_gerar\_treino\_testes\_def(comprimento\_tipo\_treino, treino, teste, C)$$

Parâmetros entrada:

- *comprimento\_tipo\_treino*: Quantidade de dados de cada classe para elaboração do conjunto de treino
- *treino*: Totalidade de dados do conjunto de treino
- *teste*: Totalidade de dados do conjunto de teste
- *C*:
  - \* 0: Se no conjunto de teste não estiverem contidos dados do mesmo tipo do conjunto de treino
  - \* 1: Se no conjunto de teste estiverem contidos dados do mesmo tipo do conjunto de treino

Parâmetros saída:

- *treino\_def*: Conjunto de treino organizado segundo o comprimento desejado
  - *teste\_def*: Conjunto de teste organizado com os restantes dados não seleccionados
  - *qt\_tipos3*: Quantidade de tipos de dados da classe 1 no conjunto de teste
  - *qt\_tipos4*: Quantidade de tipos de dados da classe 2 no conjunto de teste
- Função *classificacao\_eeg* (ver apêndice A3)  
Esta função destina-se a determinar a classificação o conjunto de teste.

$$[model1, ypred1, dfce1] = classificacao\_eeg(tipo\_class, treino\_def, teste\_def)$$

Parâmetros entrada:

- *tipo\_class*:
  - \* 1: Escolha do *smo* como algoritmo de aprendizagem
  - \* *Outrovalor*: Escolha do *svm*light como algoritmo de aprendizagem
- *treino\_def*: Conjunto de treino já organizado
- *teste\_def*: Conjunto de teste já organizado

Parâmetros saída:

- *model1*: conjunto de parametros que permitem a elaboração de um hiperplano de separação entre os dados de uma classe e outra.
  - *ypred1*: Etiquetas atribuidas no processo de classificação
  - *dfce1*: Valor da função discriminante
- Função *erros\_eeg* (ver apêndice A4)  
Esta função tem como objectivo calcular o erro de classificação de cada tipo de dados.

$$[c1, c2] = \text{erros\_eeg}(ypred1, teste\_def, qt\_tipos3, qt\_tipos4)$$

Parâmetros entrada:

- *ypred1*: Etiquetas atribuidas no processo de classificação
- *treino\_def*: Conjunto de treino já organizado
- *qt\_tipos3*: Quantidade de tipos de dados da classe 1 no conjunto de teste
- *qt\_tipos4*: Quantidade de tipos de dados da classe 2 no conjunto de teste

Parâmetros saída:

- *c1*: Erro de classificação associado a cada tipo da classe 1
- *c2*: Erro de classificação associado a cada tipo da classe 2



## Capítulo 5

# Resultados

Neste capítulo são apresentados os resultados da aplicação dos algoritmos SVM estudados, em características extraídas de segmentos de sinais EEG. Começa-se por caracterizar os sinais EEG usados nas experiências. Posteriormente são apresentados os resultados, bem como a sua discussão.

### 5.1 Caracterização dos sinais EEG

Nas experiências efectuadas, são utilizados os sinais descritos em [1].

O conjunto completo de sinais é organizado em 5 subconjuntos (A até ao E). Cada um destes subconjuntos é composto por 99 segmentos de sinal EEG. Estes sinais foram seleccionados e retirados a partir de gravações em multi-canal EEG. Todos estes sinais são apresentados sem artefactos.

Os subconjuntos A e B correspondem a segmentos de gravações EEG que foram realizadas em cinco voluntários saudáveis, sendo o subconjunto A correspondente a sinal com olhos abertos e o subconjunto B correspondente a sinal com olhos fechados. Os subconjuntos C, D e E correspondem a arquivos de diagnóstico pré cirúrgico. Os subconjuntos C e D correspondem a pacientes epiléticos sem crise epilética, com D a corresponder à zona de ocorrência de lesão e C ao hemisfério oposto à ocorrência de crise [1]. O subconjunto E corresponde somente a actividade de crise, tabela 5.1.

Na figura 5.1 é possível observar um segmento de cada tipo de sinal (A a E). A conversão analógico-digital foi efectuada com 12 bits e frequência de amostragem 173,61 Hz. A duração de cada segmento é de 23.6 segundos.

SET A	Sinal normal olhos abertos
SET B	Sinal normal olhos fechados
SET C	Sinal normal do hemisfério oposto à zona de ocorrência de crise epilética
SET D	Sinal normal da zona de ocorrência de crise epilética
SET E	Sinal com crise epilética

Tabela 5.1: Caracterização dos segmentos EEG

Tendo em conta que a aquisição dos dados foi obtida utilizando 12 bits de resolução procedeu-se à normalização dos segmentos (A a E). Cada segmento foi dividido por  $2^{11} = 2048$ , obtendo-

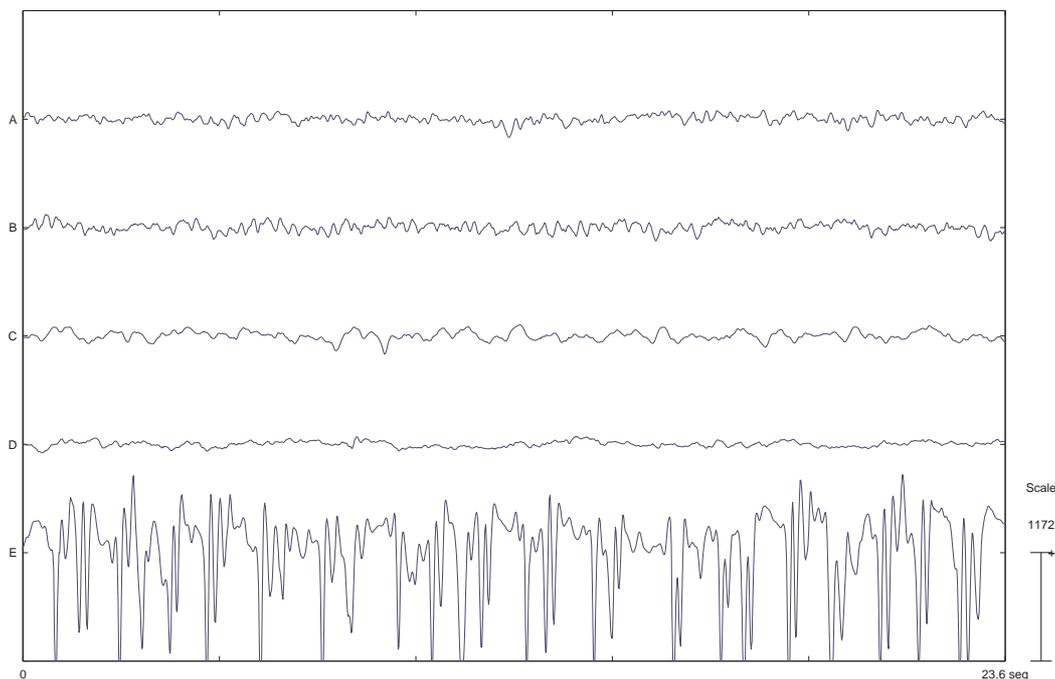


Figura 5.1: Representação dos diversos tipos de segmentos EEG (A a E)

se os dados delimitados segundo uma amplitude compreendida entre  $[-1,1]$ .

## 5.2 Extracção de características

Para facilitar a aplicação da função  $wt$  descrita em 4.1.2 truncou-se a dimensão de cada segmento para  $N=4096$ . Na prática cada segmento apresenta 4097 amostras, logo o sinal foi truncado ao  $2^x$  maior valor possível que neste caso é  $2^{11} = 4096$  amostras.

Depois de efectuada a restrição do número de elementos do sinal procedeu-se a extracção de características, com recurso à TDW (figura 4.4), escolhendo o número de níveis para decomposição, bem como a ordem a utilizar nos filtros daub (tabela 5.2).

Decomposição do sinal	Banda de Frequência ( $Hz$ )
D1	43.4 - 86.8
D2	21.43 - 43.4
D3	10.8 - 21.43
D4	5.4 - 10.8
D5	2.7 - 5.4
D6	1.35 - 2.7
D7	0.0675 - 1.35

Tabela 5.2: Frequências correspondentes aos níveis de decomposição

Neste trabalho a escolha das bandas de interesse corresponderam ao intervalo entre o  $3^o$

nível e o 6º nível, tendo em atenção os seguintes parâmetros:

- frequência amostragem =  $173.61 Hz$
- frequência superior do nível 3 =  $21.7 Hz$  (próxima de  $25 Hz$ )
- frequência inferior do nível 6 =  $1.35 Hz$  (próxima de  $0.5 Hz$ )

Desta forma há uma aproximação aos limites de frequência de interesse ( $21.7 Hz$  a  $1.35 Hz$ ), permitindo obter uma gama muito próximo dos valores usados na literatura [3] ( $25 Hz$  a  $0.5 Hz$ ).

### 5.2.1 Vector de características

Após a escolha do intervalo da banda de interesse, procedeu-se à elaboração de um vector de características, organizando-o segundo a análise da figura 5.2.

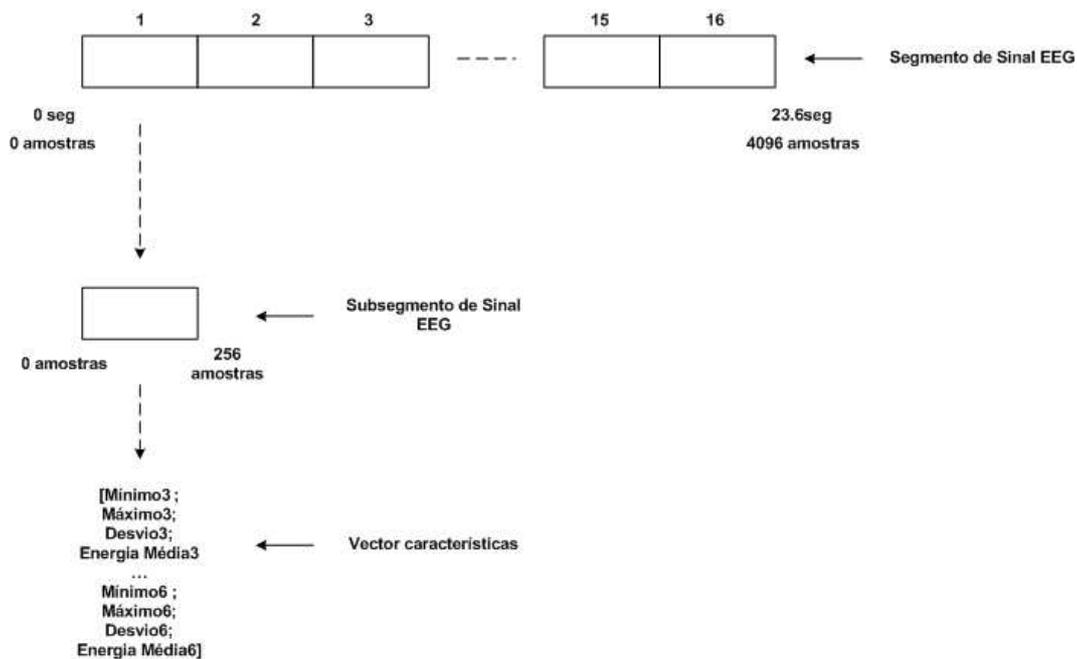


Figura 5.2: Dados organizados à saída da extração de características

Cada um dos 99 segmentos em análise é dividido em subsegmentos de 256 amostras. Para cada um destes subsegmentos foram extraídas as diversas características do sinal correspondente a cada banda de interesse. Assim sendo a informação de cada subsegmento é representada através de um vector que contém as características das bandas correspondentes. As características usadas são:

- máximo( $x^1$ ),
- mínimo( $x^2$ ),
- desvio( $x^3$ ),

- energia média( $x^4$ ),

ficando a saída da decomposição em Wavelet caracterizada por essas medidas. A equação 5.1 demonstra a organização do vector extraído.

$$\mathbf{x}_j = [x_3^1; x_3^2; x_3^3; x_3^4; \dots; x_6^1; x_6^2; x_6^3; x_6^4], \quad j=1, \dots, N \quad (5.1)$$

Onde  $x_i$  com  $i = 3, \dots, 6$  corresponde à característica na banda de interesse.

### 5.2.2 Organização do treino e do teste

Os conjuntos de treino e teste foram organizados em matrizes, onde cada coluna contém as características das bandas de interesse de cada subsegmento do sinal EEG.

A organização matricial quer do conjunto de treino, quer do conjunto de teste pode ser descrita segundo 5.2.

$$\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N] \quad (5.2)$$

Em que o N corresponde ao tamanho do conjunto, e cada  $\mathbf{x}$  representa um vector de características extraído a partir de um subsegmento de 256 amostras (equação 5.2).

## 5.3 Validação das características

Com o objectivo de verificar quais as características relevantes, procedeu-se à elaboração de gráficos que evidenciam a relação entre características dos extremos das sub bandas relevantes (3º (figura 5.3), 6º (figura 5.4)), com sinal normal (Tipo A) e com sinal de crise epiléptica (Tipo E).

Estes gráficos (figuras 5.3 e 5.4) permitem verificar que não há uma característica que se destaque sobre as outras. No entanto, em [3] a Energia Média é utilizada como única característica no classificador.

Os resultados obtidos permitem verificar que a Energia Média, tal como as outras características, não apresentam uma acentuada distinção entre dados do tipo A e E. No entanto, para o nível 6 da decomposição wavelet a Energia Média é aquela que apresenta uma melhor separação entre dados. Assim optou-se pela realização de algumas experiências, utilizando a Energia Média como única característica no processo de classificação.

## 5.4 Avaliação do classificador

Para avaliar a performance do classificador SVM efectuaram-se duas estratégias (A e B), mencionadas na literatura:

- **Experiência A**

Sinal Normal vs Sinal epiléptico. Na base de dados corresponde ao Sinal A vs Sinal E. Observam-se resultados publicados recentemente com este tipo de estratégia [22].

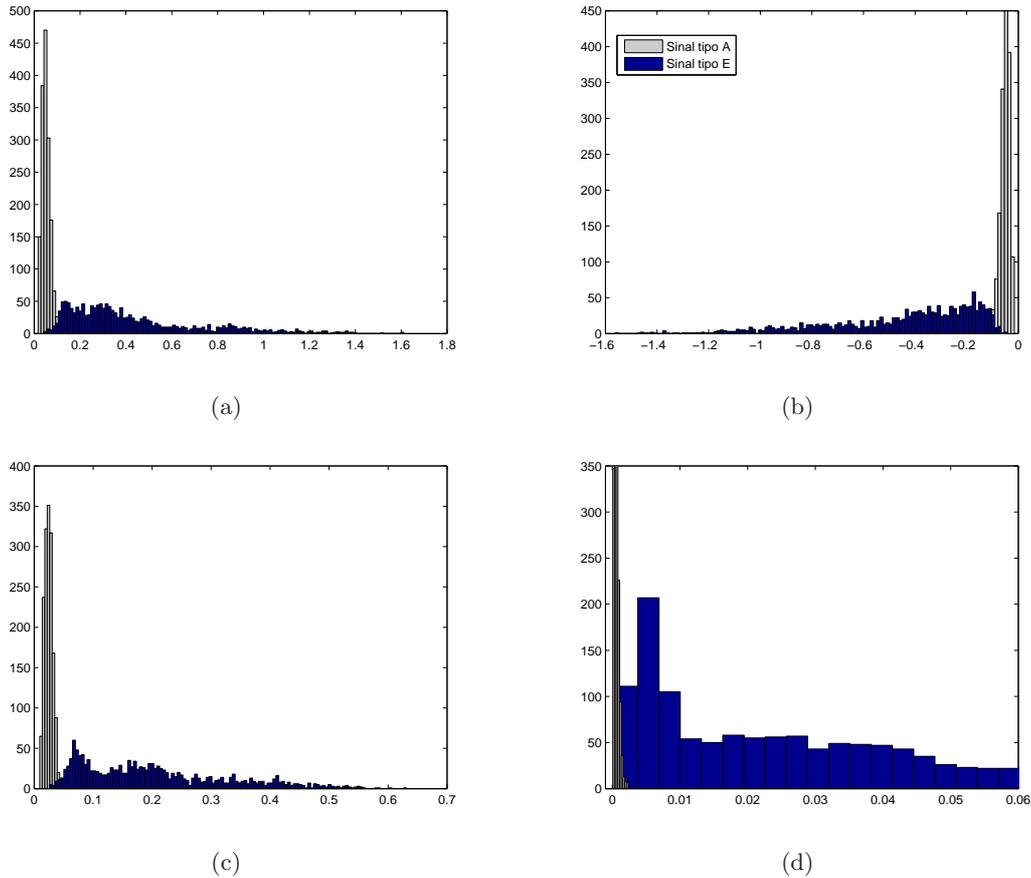


Figura 5.3: Relação entre características do nível 3 (a) Máximo, (b) Mínimo, (c) Desvio, (d) Energia Média

- **Experiência B**

Sinal não epiléptico vs Sinal epiléptico. Na base de dados corresponde ao A, B, C e D vs E. Observam-se resultados publicados recentemente com este tipo de estratégia [22].

Na avaliação e desempenho do classificador são utilizados no conjunto de teste vectores de características de todos os tipos de sinais.

Cada conjunto de treino foi seleccionado 30 vezes de forma aleatória e testado com os restantes vectores de características (conjunto de teste). Em cada iteração foi calculado o erro associado à classificação.

Neste estudo foram efectuadas duas experiências (A e B). Em cada experiência são distinguidos dois casos, classificação baseada em 4 características por banda e classificação baseada na Energia Média por banda.

Os erros associados à classificação dos dados em cada experiência são apresentados em box-plots.

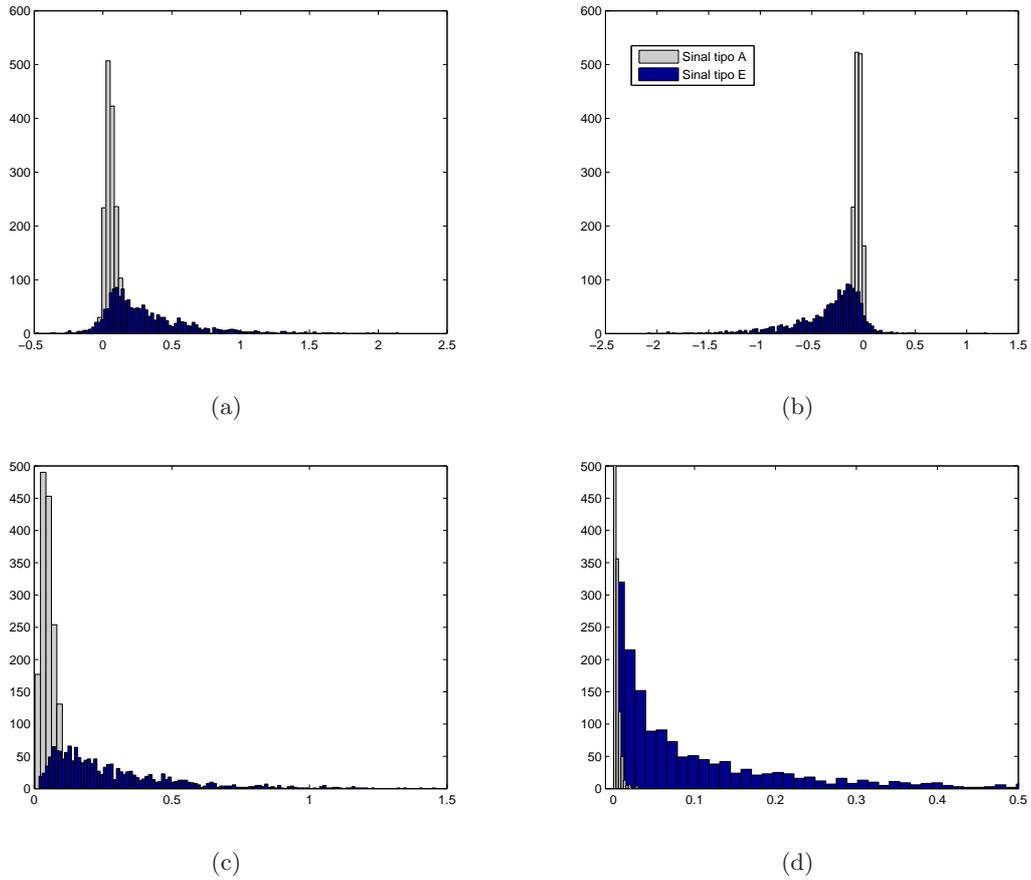


Figura 5.4: Relação entre características do nível 6 (a) Máximo, (b) Mínimo, (c) Desvio, (d) Energia Média

### 5.4.1 Experiência A

Nesta experiência utilizou-se um conjunto de treino contendo duas classes (classe 1 e classe 2). A classe 1 é composta por dados do tipo A e a classe 2 composta por dados do tipo E. Também se estudou a influência do tamanho do conjunto de treino no desempenho do classificador. Assim foram criados e 3 treinos, designados por A1, A2 e A3 tal como descrito na tabela 5.3.

	Classe 1	Classe 2
Treino A1	100	100
Treino A2	500	500
Treino A3	1000	1000

Tabela 5.3: Dimensão dos diversos treinos na experiência A

**Caso 1:** Classificação baseada em 4 características por banda

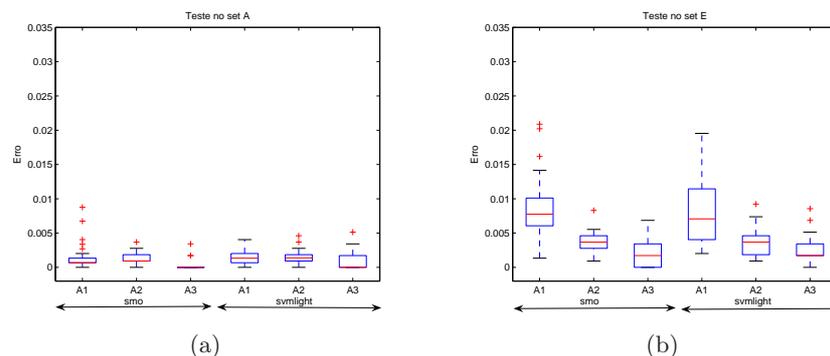


Figura 5.5: Boxplot do erro associado ao Teste no conjunto A e ao Teste no conjunto E utilizando as 4 características (Experiência A caso 1)

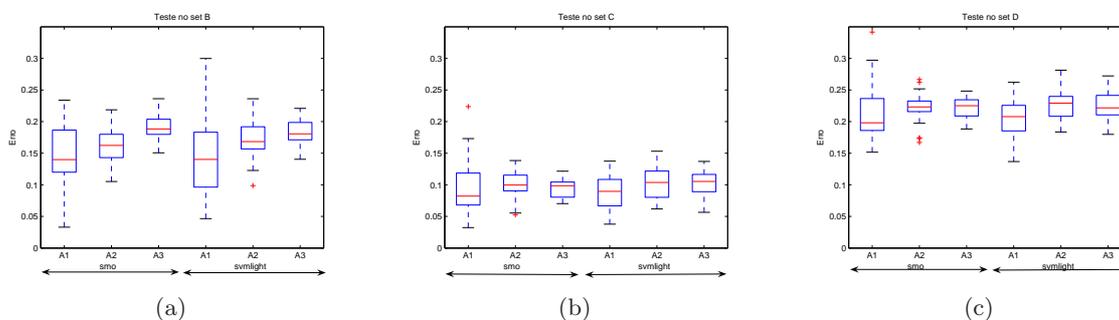


Figura 5.6: Boxplot do erro associado ao Teste no conjunto B, Teste no conjunto C e Teste no conjunto D, utilizando as 4 características (Experiência A caso 1)

Aspectos a ter em conta na análise aos resultados obtidos experimentalmente (Experiência A caso 1):

- **Conjunto de Treino vs Conjunto de Teste**

A realização desta experiência permite verificar um melhor desempenho por parte do classificador quando testado com dados do mesmo tipo do conjunto de treino (tal como se pode observar nas figuras 5.5 e 5.6).

O facto da aprendizagem resultar de um conjunto de treino constituído por dados de apenas dois tipos (A e E), faz com que taxa de erro de classificação nos dados do tipo A e E apresente valores inferiores aos dados do tipo B, C e D. Enquanto que nos gráficos da figura 5.5 a taxa de erro de classificação é menor que 0.020, nos gráficos da figura 5.6 varia entre 0.025 e 0.3.

- **Dimensão do conjunto de treino**

É possível constatar que para um conjunto treino e teste constituídos por dados do mesmo tipo, uma variação da dimensão do conjunto de treino não provoca uma variação brusca na taxa de erro de classificação. No caso da figura 5.5a o erro é menor que 0.005 qualquer que seja o tamanho do conjunto de treino (A1, A2 e A3).

Na figura 5.5b, verifica-se uma diminuição da taxa de erro de classificação, à medida

que se aumenta a dimensão do conjunto de treino. Para o treino A1 a taxa de erro é de aproximadamente 0.0075, aproximando-se de zero à medida que as dimensões do treino aumentam (casos A2 e A3).

Nos gráficos da figura 5.6, verifica-se um aumento do erro, à medida que é aumentada a dimensão do conjunto de treino. Assim para uma aprendizagem de um treino com dados exclusivamente do tipo A e E, quando se aumenta o conjunto de treino com dados do mesmo tipo, o classificador generaliza pior.

### Caso 2: Classificação baseada na energia média por banda

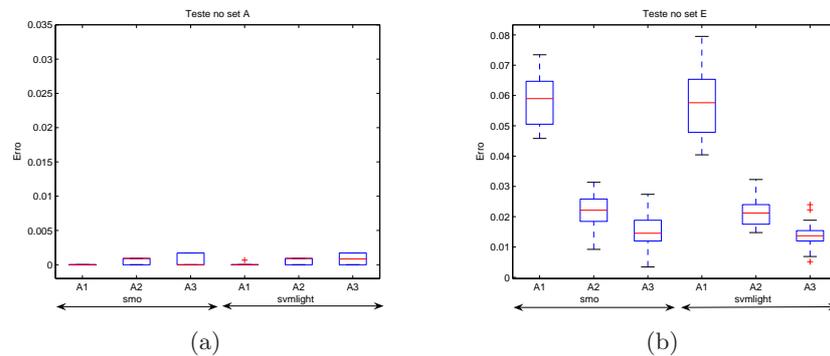


Figura 5.7: Boxplot do erro associado ao Teste no conjunto A e ao Teste no conjunto E utilizando a Energia média como característica (Experiência A caso 2)

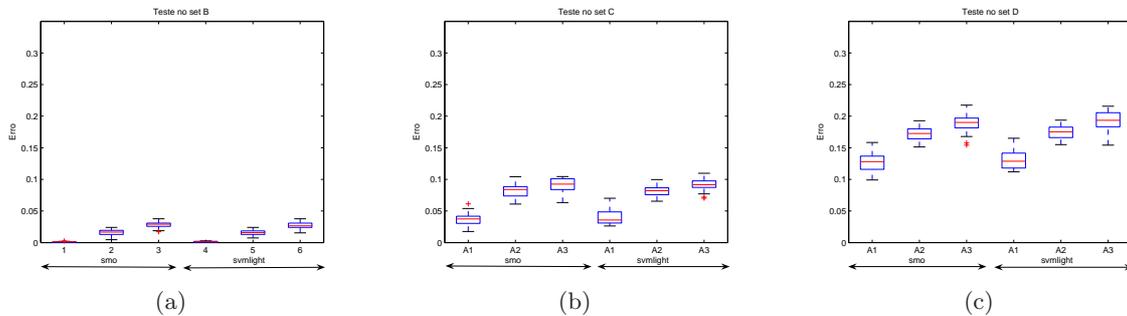


Figura 5.8: Boxplot do erro associado ao Teste no conjunto B, Teste no conjunto C e Teste no conjunto D utilizando a Energia média como característica (Experiência A caso 2)

Aspectos a ter em conta na análise aos resultados obtidos experimentalmente (Experiência A caso 2):

- **Conjunto de Treino vs Conjunto de Teste**

A utilização de um conjunto de teste contendo dados do mesmo tipo do conjunto de treino, proporciona um desempenho mais eficiente do classificador, tal como no caso 1 da experiência A. Enquanto que na figura 5.7a a taxa de erro de classificação é menor que 0.005, na figura 5.7b é menor que 0.08.

Nos gráficos da figura 5.8 a taxa de erro apresenta valores inferiores a 0.2.

- **Dimensão do conjunto de treino**

A variação da dimensão do conjunto de treino não provoca uma variação brusca na taxa de erro de classificação, sendo esta inferior a 0.05 no caso da figura 5.7a.

Constata-se pela análise da figura 5.7b, uma diminuição da taxa de erro de classificação à medida que se aumenta a dimensão do conjunto de treino. Para o treino A1 a taxa de erro é de aproximadamente 0.06, diminuindo drasticamente à medida que as dimensões do treino aumentam (casos A2 e A3).

Nos gráficos da figura 5.8, verifica-se um aumento do erro à medida que se aumenta a dimensão do conjunto de treino. Tal como no **Caso 1**, para uma aprendizagem de um treino com dados exclusivamente do tipo A e E, quando se aumenta o conjunto de treino com dados do mesmo tipo, o classificador generaliza pior.

### Considerações finais acerca da Experiência A

Um dos objectivos principais desta experiência, consistiu na observação das diferenças entre uma classificação utilizando 4 características por banda e uma classificação utilizando a energia média por banda.

Quando a energia é utilizada como única característica no processo de aprendizagem do classificador, verifica-se um aumento da taxa de erro na classificação de sinal epiléptico. Enquanto que na figura 5.5b a taxa de erro é menor que 0.0015, na figura 5.7b é menor que 0.06.

No entanto, há uma diminuição do erro no que se refere à classificação de sinal não epiléptico. Comparando as figuras 5.5a e 5.7a com os gráficos das figuras 5.6 e 5.8, observa-se uma diminuição da taxa de erro quando se utiliza a energia média como característica.

### 5.4.2 Experiência B

Nesta experiência utilizou-se um conjunto de treino contendo dados do tipo A, B, C, D e E. Elaboraram-se 2 treinos com diferentes dimensões; como se observa na tabela 5.4.

A classe 1 é composta por dados do tipo A, B, C e D, distribuídos equitativamente, enquanto que a classe 2 é constituída por dados do tipo E.

	Classe 1	Classe 2
Treino B1	100	100
Treino B2	500	500

Tabela 5.4: Dimensão dos diversos treinos na experiência B

#### **Caso 1:** Classificação baseada em 4 características por banda

Aspectos a ter em conta na análise aos resultados obtidos experimentalmente (Experiência B caso 1):

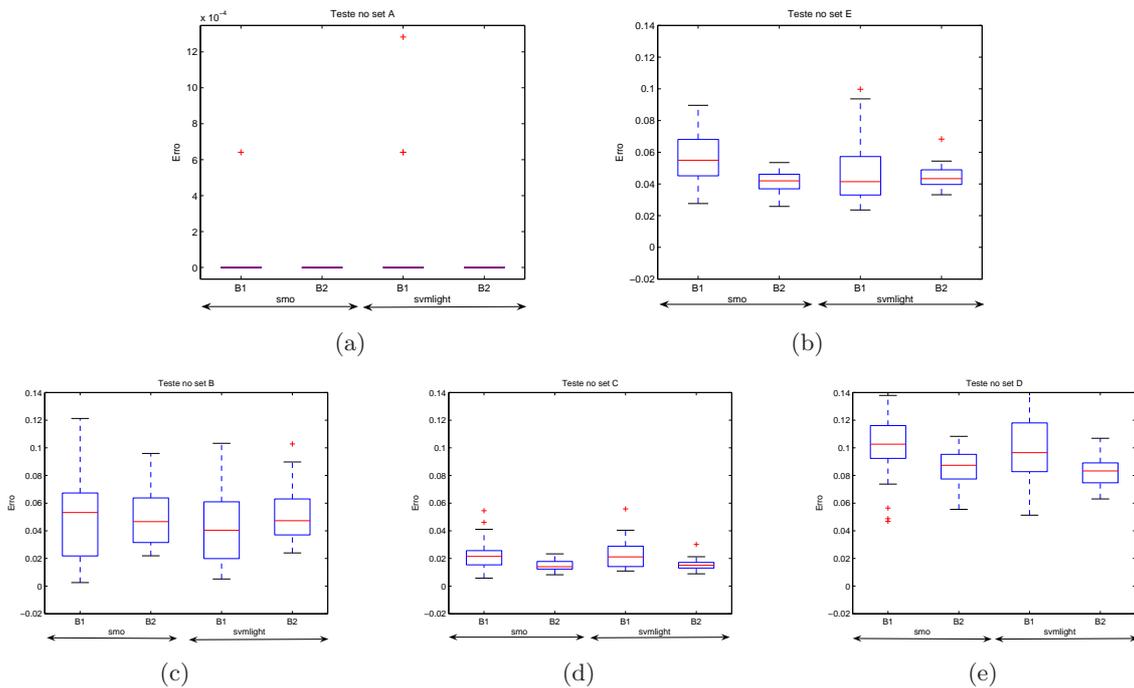


Figura 5.9: Boxplot do erro associado ao Teste no conjunto A, Teste no conjunto B, Teste no conjunto C, Teste no conjunto D, Teste no conjunto E utilizando as 4 características (Experiência B caso 1)

- **Conjunto de Treino vs Conjunto de Teste**

Um conjunto de treino composto por dados de todos os tipos (A, B, C, D e E) evidencia uma melhoria na classificação de sinal não epiléptico (comparando os gráficos da figura 5.9 com as figuras 5.5a ,5.7a, gráficos da figura 5.6 e gráficos da figura 5.8). Na figura 5.9a a taxa de erro é nula, evidenciando assim um bom desempenho na classificação de dados do tipo A.

Contudo é verificado um aumento da taxa de erro na classificação de sinal epiléptico (comparando as figuras 5.5b e 5.7b, com a figura 5.9b).

- **Dimensão do conjunto de treino**

Na generalidade dos gráficos da figura 5.9 é possível observar uma diminuição da taxa de erro à medida que se aumenta a dimensão do conjunto de treino. Esta diminuição deve-se essencialmente ao facto do conjunto de teste ser constituído por dados do mesmo tipo do conjunto de treino. Assim para uma aprendizagem de um treino com dados de todos os tipos (A, B, C, D e E), quando se aumenta o conjunto de treino com dados do mesmo tipo, o classificador generaliza melhor.

**Caso 2:** Classificação baseada na energia média por banda

Aspectos a ter em conta na análise aos resultados obtidos experimentalmente (Experiência B caso 2):

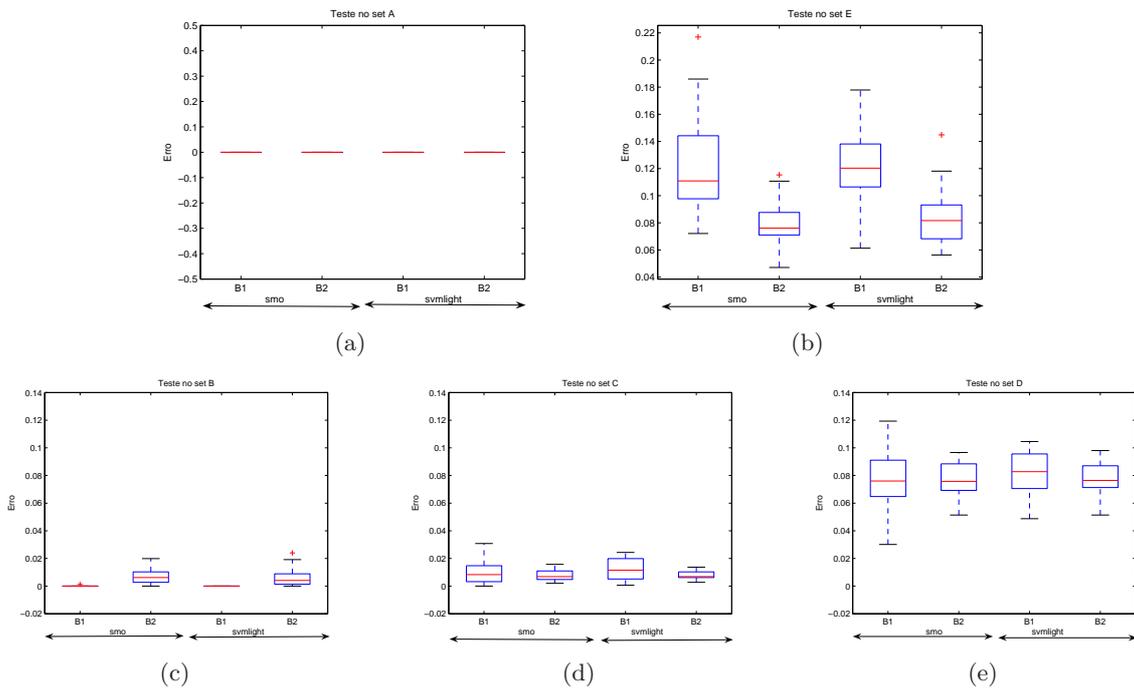


Figura 5.10: Boxplot do erro associado ao Teste no conjunto A, Teste no conjunto B, Teste no conjunto C, Teste no conjunto D e Teste no conjunto E utilizando a Energia média como característica(Experiência B caso 2)

- **Conjunto de Treino vs Conjunto de Teste**

No caso 2 da experiência B, há igualmente uma melhoria na detecção de sinal não epiléptico com um conjunto de treino composto por dados de todos os tipos (A, B, C, D e E) (comparando as figuras 5.10 com as figuras 5.5a ,5.7a, gráficos da figura 5.6 e os gráficos da figura 5.8). Na figura 5.9a a taxa de erro é nula, evidenciando assim um bom desempenho na classificação de dados do tipo A, tal como o sucedido no caso 1 da experiência B.

Contudo é verificado um aumento da taxa de erro na detecção de sinal epiléptico (comparando as figuras 5.5b e 5.7b, com a 5.10b), à semelhança do caso 1 da experiência B.

- **Dimensão do conjunto de treino**

É igualmente observável na maioria dos gráficos da figura 5.10, uma diminuição da taxa de erro à medida que se aumenta a dimensão do conjunto de treino, comparativamente ao caso 1 da experiência B. Esta diminuição deve-se mais uma vez, ao facto do conjunto de teste ser constituído por dados do mesmo tipo do conjunto de treino. Tal como no caso 1 da Experiência B para uma aprendizagem de um treino com dados de todos os tipos (A, B, C, D e E), quando se aumenta o conjunto de treino com dados do mesmo tipo, o classificador generaliza melhor.

## Considerações finais acerca da Experiência B

O principal objectivo desta experiência consistiu na elaboração de um conjunto de treino contendo dados de todos os tipos (A, B, C, D e E). Assim, comparando com a experiência A (conjunto de treino com dados do tipo A e E) observa-se a influência de conjuntos de treino compostos por diferentes tipos de dados na classificação.

Tal como na experiência A, quando a energia é utilizada como única característica no processo de aprendizagem do classificador, verifica-se um aumento da taxa de erro na detecção de sinal epiléptico. Comparando a figura 5.9b com a figura 5.10b, há um aumento da taxa de erro quando se utiliza a energia média como característica única no processo de aprendizagem.

No entanto há uma diminuição do erro no que se refere à detecção de sinal não epiléptico.

Fazendo uma comparação entre os gráficos a, b, c e d da figura 5.9 e 5.10, observa-se um diminuição do erro quando se utiliza a energia média como característica.

## 5.5 Análise global

Nesta secção é feita uma análise global dos resultados expostos anteriormente.

Com objectivo de expor os resultados de forma quantificável, utilizou-se uma medida muito usada na literatura [22], denominada por *Accuracy*. Esta medida estatística é frequentemente utilizada em problemas de classificação binária para calcular a precisão de classificação dos dados, podendo ser descrita pela equação 5.3.

$$P = \frac{\#(Vp + Vn)}{\#(Vp + Vn + Fp + Fn)} \quad (5.3)$$

Esta equação representa a percentagem de resultados correctamente classificados.

A atribuição de positivo (p) ou negativo (n) refere-se à decisão do algoritmo de classificação. Enquanto que o verdadeiro (V) ou falso (F) é uma relação entre a atribuição feita na classificação e o estado clínico previamente definido (Tabela 5.5).

		Estado		
		Verdadeiro	Falso	
Teste	Positivo	Vp	Fp	
	Negativo	Fn	Vn	
				Precisão(P)

Tabela 5.5: Matriz Confusão

Assim, evidenciam-se 4 tipos de situações, o verdadeiro positivo, o verdadeiro negativo, o falso positivo e o falso negativo. O verdadeiro positivo (Vp), indica que a classificação foi a correcta, detectando anormalidade. O verdadeiro negativo (Vn), indica que a classificação foi a correcta, não detectando anormalidade. O falso positivo (Fp), indica que a classificação foi incorrecta, detectando anormalidade. O falso negativo (Fn), indica que a classificação foi incorrecta, não detectando anormalidade. Para cada experiência (A e B) foram calculados os valores de  $P$ , que estão descritos nas tabelas 5.6 e 5.7. Nesta análise houve uma escolha do tipo de teste, elaborando-se assim dois conjuntos. Um conjunto de teste que é composto por dados do tipo A e E e outro conjunto de teste que é composto pela totalidade de tipos de

dados (A, B, C, D e E).

Experiência A: Treino(A,E)						
	Caso 1			Caso 2		
Teste	A1	A2	A3	A1	A2	A3
(A,E)	99,47%	99,76%	99,87%	97,07%	98,85%	99,23%
(A,B,C,D,E)	90,76%	90,27%	89,74%	95,52%	94,15%	93,50%

Tabela 5.6: Accuracy associada ao Caso 1 e 2 da experiência A

Experiência B: Treino (A,B,C,D e E)				
	Caso 1		Caso 2	
Teste	B1	B2	B1	B2
(A,E)	97,26%	98,24%	94,12%	96,57%
(A,B,C,D,E)	95,39%	96,20%	95,89%	96,82%

Tabela 5.7: Accuracy associada ao Caso 1 e 2 da experiência B

Através da observação das tabelas 5.6 e 5.7, verifica-se que aumentando a dimensão do conjunto de treino, há uma melhoria de  $P$ , mas apenas quando, no conjunto de teste estão presentes dados do mesmo tipo do conjunto de treino.

Para um conjunto de treino e teste constituído por dados exclusivamente do tipo A e E (experiência A), obteve-se um valor para  $P$  próximo de 100%, sendo considerada como uma classificação eficaz. Para um conjunto de treino e teste constituído por dados do tipo A, B, C, D e E, obteve-se como melhor valor para  $P$ , 96.82%.

Comparando o caso 1 com o caso 2 da experiência A verifica-se uma diminuição ligeira de  $P$ , quando se utiliza a energia como única característica. No entanto na experiência B, obtém-se exactamente o contrário, ou seja, verifica-se um aumento de  $P$ , quando se utiliza a energia como única característica, para um treino constituído por dados de todos os tipos.

## 5.6 Análise Temporal

Nesta secção é feita uma ilustração temporal dos resultados das experiências A e B, utilizando os diferentes conjuntos de treino. Para cada caso ilustrado, seleccionou-se um segmento de cada tipo de sinal (A, B, C, D e E), dividindo-o em subsegmentos de 256 amostras, classificando-o posteriormente.

Para caracterizar o resultado da classificação utilizaram-se dois símbolos:

- $\square$ , quando o sinal classificado pertence à classe 1
- $*$ , quando o sinal classificado pertence à classe 2

Desta forma é possível uma observação das relações entre conjuntos de treino e características utilizadas, de uma forma temporal.

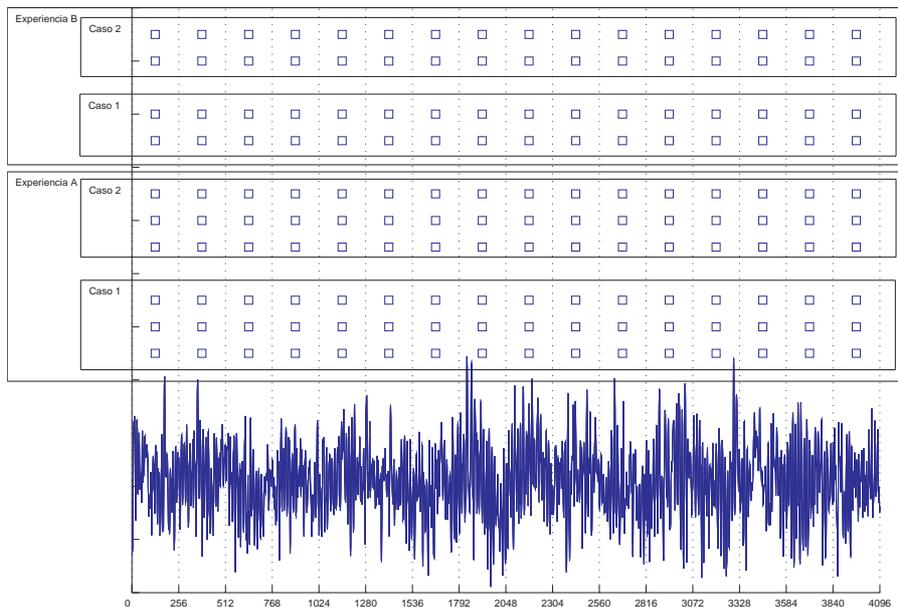


Figura 5.11: Análise a um segmento do tipo A

Pela observação da figura 5.11, tanto para os casos experiência A como para os casos da experiência B, a classificação foi a correcta em todos os subsegmentos analisados. A utilização de diferentes tipos de conjuntos de treino, evidenciou-se como irrelevante.

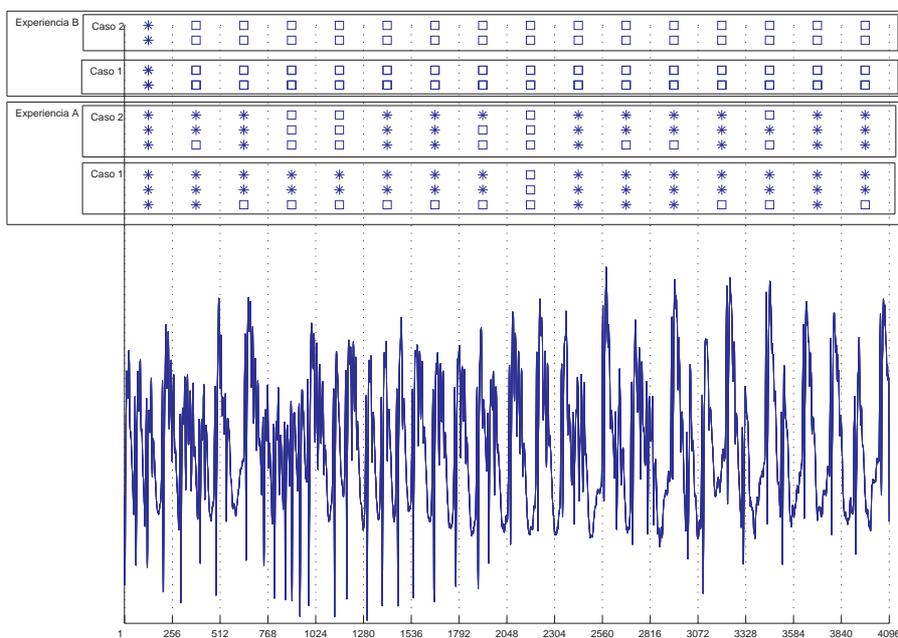


Figura 5.12: Análise a um segmento do tipo E

Quando o caso em análise é um segmento do tipo E, observa-se (figura 5.12), que na experiência A, o caso 1 apresenta uma melhoria na classificação relativamente ao caso 2. No

entanto em ambos os casos, à medida que se aumenta o conjunto de treino há uma melhoria na classificação, tal como o descrito nas análises às experiências. A experiência B não se revela tão eficaz na classificação evidenciando um elevado erro.

Assim, a experiência A tornou-se a mais adequada na classificação deste segmento revelando uma classificação mais eficaz do que a experiência B.

Com a utilização de um segmento do tipo B (figura 5.13), à imagem do que aconteceu nas

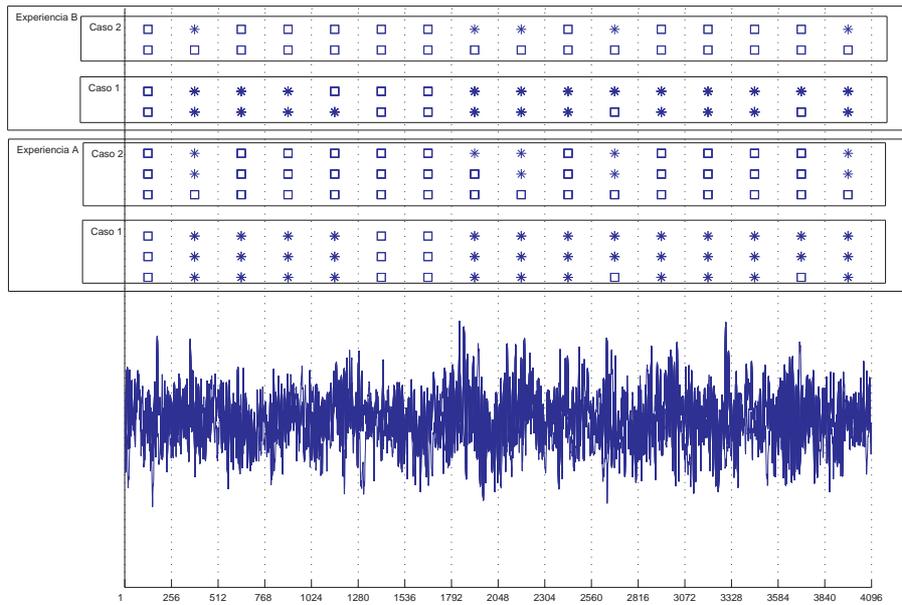


Figura 5.13: Análise a um segmento do tipo B

análises das boxplots, em todos os casos, houve um aumento do erro quando de aumentou a dimensão do conjunto de treino. No caso 2 da experiência B chega-se mesmo a ter, a totalidade do segmento bem classificado, estando por isso de acordo com os resultados obtidos anteriormente (figura 5.10c).

A utilização de um segmento do tipo C (figura 5.14), evidenciou os resultados obtidos nas análises às experiências, estando completamente de acordo com o obtido anteriormente. Realça-se o caso 2 da experiência B, onde se obtém uma classificação correcta da totalidade do segmento em análise.

A utilização de um segmento do tipo D (figura 5.15), evidenciou que na experiência B a classificação foi mais eficaz do que na experiência A. No entanto há um aumento do erro de classificação quando se aumenta a dimensão do conjunto de treino na experiência B, caso 1. Analisando a figura 5.9c verifica-se que deveria haver uma diminuição do erro, quando se aumenta a dimensão do conjunto de treino. No entanto como o desvio associado ao treino B1 abrange o valor médio do erro no treino B2, admite-se um aumento do erro neste caso.

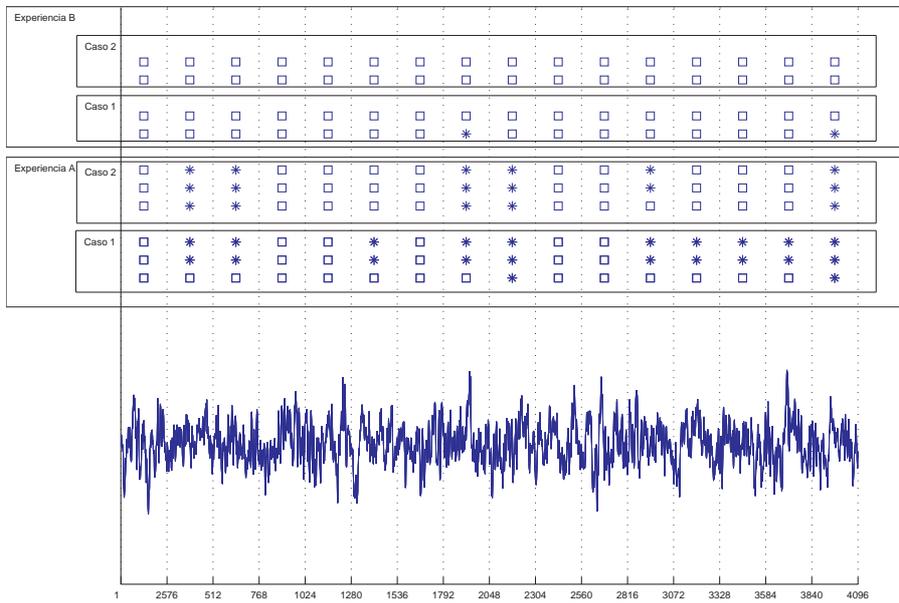


Figura 5.14: Análise a um segmento do tipo C

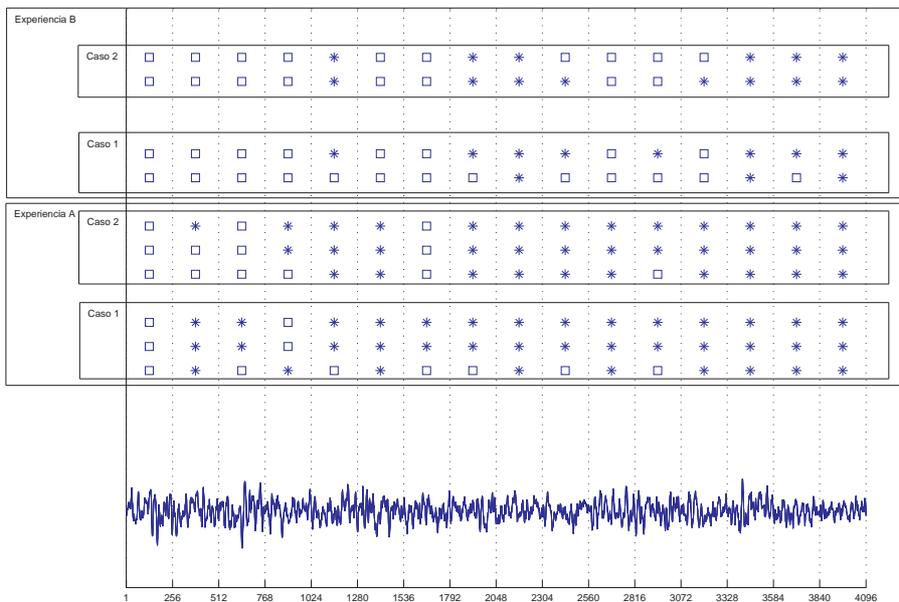


Figura 5.15: Análise a um segmento do tipo D

## Capítulo 6

# Conclusões

Nesta tese, apresentou-se um sistema de classificação de sinais EEG. O sistema descrito pode ser dividido em dois blocos, o primeiro que trata da extração de características do sinal EEG, com apoio à Transformada Discreta de Wavelet, e o segundo é baseado num algoritmo de classificação linear do tipo SVM. Este método passa por dividir um sinal em subsegmentos, onde são extraídas características. O processo de extração de características, é fundamental, pois uma selecção das melhores características permite uma classificação mais eficiente. As características seleccionadas foram o máximo, mínimo, desvio e energia média.

- **Energia vs 4 Características**

É importante referir que a utilização de 4 características (máximo, mínimo, desvio e energia média) ou de apenas uma (energia média), não apresentaram resultados muito díspares na classificação. No entanto, quando se utilizaram as 4 características, houve um melhoramento na classificação de sinal epiléptico. Quando se utilizou a energia média como única característica no processo de classificação houve um melhoramento na detecção de sinal não epiléptico.

Poderiam ser escolhidas outros tipos de características para a classificação de sinal EEG. No entanto os resultados obtidos nas experiências apresentaram uma baixa percentagem de erro de classificação, sendo por isso, demonstrada uma viabilidade na escolha de características relacionadas com a amplitude do sinal.

- **Influência do tipo de dados na classificação**

Outro dos aspectos em estudo foi o tipo de conjunto de treino. Uma variação do tipo de conjunto de treino poderá provocar efeitos na classificação, daí a elaboração conjuntos constituídos por diferentes tipos de sinais, para uma observação do seu efeito na classificação.

Um conjunto de treino constituído pela totalidade dos tipos de sinais (A, B, C, D e E) evidenciou um melhoramento na detecção de sinal não epiléptico. No entanto, quando o conjunto de dados foi exclusivamente constituído por dados do tipo A e E houve um melhoramento na detecção de sinal epiléptico.

- **Dimensão do conjunto de treino**

Um dos aspectos a ter em conta neste trabalho teve haver com a dimensão do conjunto

de treino e a sua influência na classificação. De uma forma geral, quando o conjunto de teste é constituído por elementos do mesmo tipo do conjunto de treino evidencia-se um melhoramento quando se aumenta a dimensão do conjunto de treino. Quando o conjunto de teste é constituído por elementos de tipo diferente do conjunto de treino, de uma forma geral pioram os resultados da classificação à medida que se aumenta a dimensão do conjunto de treino.

- **Comparação com estudos recentes**

Fazendo uma comparação entre este estudo e os mais recentes [22], observa-se que com a utilização de um classificador linear, é possível uma aproximação aos melhores resultados descritos até ao momento na literatura [22].

Neste trabalho obteve-se um valor para  $P$  de 99,87% quando treinado com dados do tipo A e E e testado com dados do mesmo tipo, muito próximo do valor descrito na literatura (100%). Contudo o trabalho exposto na literatura [22], recorre ao uso de classificação baseada em redes neuronais.

Quando o classificador foi treinado com dados do tipo A, B, C, D e E e testado com um conjunto de dados do mesmo tipo obteve-se um valor de  $P$  de 96.82%, próximo dos 97.73% obtidos na literatura [22].

Na globalidade, os valores de  $P$ , apresentam uma percentagem elevada, nunca sendo inferior a 89%. Assim, o algoritmo elaborado apresenta uma elevada percentagem de resultados correctamente classificados, mesmo quando treinado com conjuntos de treino compostos por pequenas dimensões.

É importante referir que quando o conjunto de teste é constituído por elementos do mesmo tipo do conjunto de treino há um melhoramento de  $P$ , quando se aumenta a dimensão do conjunto de treino. Quando no conjunto de teste estão presentes tipos de dados que não estão contidos no conjunto de treino, há uma diminuição de  $P$  à medida que se aumenta a dimensão do conjunto de treino.

Com um conjunto treino constituído por dados do tipo A e E e testado com um conjunto com dados constituído por elementos de todos os tipos (A, B, C, D e E), obteve-se como pior resultado de  $P$ , 89,4% (utilizando 4 características) e 93,50% (utilizando a energia como única característica). Estes resultados demonstram, que mesmo com um conjunto de treino constituído com apenas dados exclusivamente do tipo A e E, é possível obter uma elevada percentagem de acertos na classificação.

- **Análise temporal**

Neste estudo, foi também elaborada uma análise temporal de um segmento de cada tipo, dividindo-o em subsegmentos e classificando-os. Assim, é possível observar uma evolução dos resultados da classificação de um determinado segmento, quer no tempo, quer quando são utilizados conjuntos de treino constituídos por diferentes tipos de dados e com dimensões diferentes.

Não sendo um método sistemático, permite uma ilustração das experiências efectuadas (A e B).

A nível geral, os resultados obtidos apresentam-se como satisfatórios, sendo de enaltecer o facto de toda a classificação ser efectuada com a utilização de um classificador linear.

No entanto fica a ideia que os dados utilizados são demasiado simples.

## 6.1 Propostas futuras

Nesta secção apresenta-se duas possibilidades para um trabalho futuro:

- **Utilização de dados reais**

Um trabalho interessante seria, a utilização de sinais EEG reais, com recurso ao uso do algoritmo elaborado. O sistema em mente trabalharia em multi-canal, procedendo-se desta forma a uma classificação canal a canal.

A localização seria um factor de interesse, pois desta forma podiam-se delimitar regiões cerebrais com problemas epilépticos.

- **Início de crise**

Outro trabalho de interesse seria a utilização do algoritmo em tempo real.

Esta estratégia poderia ser elaborada, não para uma detecção de crises, mas sim para uma detecção de início de crise. Teria de haver um aprimoramento do algoritmo de modo a detectar apenas o início de crise, ou ate mesmo encontrar padrões que poderiam relacionar-se com esse início. Assim, o algoritmo ajudaria a prever uma futura crise.

Uma possibilidade seria, criar um sistema de comunicação entre o paciente e o médico, permitindo uma intervenção imediata.



# Bibliografia

- [1] R. G. A. Lehnertz and Klaus, “Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state,” 2001.
- [2] S. Shigemura, T. Nishimura, M. Tsubai, and H. Yokoi, “An investigation of EEG artifacts elimination using a neural network with non-recursive 2nd order volterra filters,” in *Engineering in Medicine and Biology Society, 2004. IEMBS '04. 26th Annual International Conference of the IEEE*, vol. 1, pp. 612–615, 2004.
- [3] A. Shoeb, H. Edwards, J. Connolly, B. Bourgeois, S. Ted Treves, and J. Guttag, “Patient-specific seizure onset detection,” *Epilepsy and Behavior*, vol. 5, no. 4, pp. 483–498, 2004.
- [4] <http://www.abcdasaude.com.br/artigo.php?95> (Último acesso a 4/11/2008).
- [5] <http://www.robots.ox.ac.uk/parg/projects/bci/rev1.html> (Último acesso a 4/11/2008).
- [6] <http://www.geocities.com/Itallis/Omanual/Placement.htm> (Último acesso a 4/11/2008).
- [7] P. Jahankhani, K. Revett, and V. Kodogiannis, “Data mining an eeg dataset with an emphasis on dimensionality reduction,” in *Computational Intelligence and Data Mining, 2007. CIDM 2007. IEEE Symposium on*, pp. 405–412, 2007.
- [8] J. B. Ochoa, “EEG signal classification for brain computer interface applications,” *Ecole Polytechnique Federale De Lausanne*.
- [9] M. Selim R Benbadis, “<http://www.emedicine.com/neuro/topic678.htm>,”
- [10] P. Gloor and J. Gotman, “Automatic recognition and quantification of interictal epileptic activity in the human scalp EEG.,” *Electroencephalography and Clinical Neurophysiology*, vol. 41, p. 513:529, 1976.
- [11] J. Gotman, “Automatic recognition of epileptic seizures in the EEG,” *Electroencephalogr Clin Neurophysiol*, vol. 54, no. 5, pp. 530–40, 1982.
- [12] Q. Hao and J. Gotman, “A patient-specific algorithm for the detection of seizure onset in long-term eeg monitoring: possible use as a warning device,” *Biomedical Engineering, IEEE Transactions on*, vol. 44, no. 2, pp. 115–122, 1997.
- [13] L. Tarassenko, Y. U. Khan, and M. R. G. Holt, “Identification of inter-ictal spikes in the EEG using neural network analysis,” *Science, Measurement and Technology, IEE Proceedings -*, vol. 145, no. 6, pp. 270–278, 1998.

- [14] A. Subasi, “EEG signal classification using wavelet feature extraction and a mixture of expert model,” *Expert Systems with Applications*, vol. 32, no. 4, pp. 1084–1093, 2007.
- [15] P. Jahankhani, V. Kodogiannis, and K. Revett, “EEG signal classification using wavelet feature extraction and neural networks,” in *Modern Computing, 2006. JVA '06. IEEE John Vincent Atanasoff 2006 International Symposium on* (V. Kodogiannis, ed.), pp. 120–124, 2006.
- [16] I. Guler and E. D. Ubeyli, “Adaptive neuro-fuzzy inference system for classification of EEG signals using wavelet coefficients,” *Journal of Neuroscience Methods*, vol. 148, no. 2, pp. 113–121, 2005.
- [17] [citeseer.ist.psu.edu/672431.html](http://citeseer.ist.psu.edu/672431.html) (Último acesso a 4/11/2008).
- [18] K. Theodoridis, Sergios; Koutroumbas, *Pattern recognition*. academic press ed., 2006.
- [19] M. Tan, Pang-Ning ; Steinbach and V. Kumar, *Introduction to data mining*. addition-wesley ed., 2006.
- [20] <http://cmp.felk.cvut.cz/cmp/software/stprtool/> (Último acesso a 4/11/2008).
- [21] <http://www.mathworks.com/products/matlab/> (Último acesso a 4/11/2008).
- [22] A. T. Tzallas, M. G. Tsipouras, and D. I. Fotiadis, “Automatic seizure detection based on time-frequency analysis and artificial neural networks,” *Intell. Neuroscience*, vol. 7, no. 3, pp. 1–13, 2007.

# Apêndice A

## A1

```
function [A]=eeg_extracao_funcao_Classe...
    (ficheiro ,L,niveis ,niveis_d ,ordem_filtro ,tipo ,energia)

%function [A]=eeg_extracao_funcao_Classe...
%    (ficheiro ,L,niveis ,niveis_d ,ordem_filtro ,tipo ,energia)
%
%Entrada:
%    ficheiro: Sinal de entrada
%            exemplo:
%                ficheiro=[Sinal 1 , Sinal 2 ,...]
%                Nota: Cada Sinal deve ser organizado por coluna
%
%    L: Número de divisões desejadas para cada segmento de sinal
%
%    niveis: Número de níveis desejados para decomposição
%
%    niveis_d:Quantidade de ultimos níveis desejeveis
%
%    ordem_filtro:Ordem do filtro daub
%
%    tipo:Tipo de classe a que pertence 1 ou 2
%
%    energia:
%            0:Para a extracção das 4 Caracteristicas
%            1:Para a extração exclusiva da energia média
%
%Saída:
%
%    A: matriz contendo os vectores de características do segmento em
%        análise
%
x=ficheiro;
```

```

%Normalização
x=x./2048
%Restrição de 2^N elementos
for i=1:14
    if (length(x) > 2^i || length(x)==2^i) && (length(x) < 2^(i+1))
        n_amostras=2^i;
        i=i+1;
    end
end
n_amostras;

x=x(1:n_amostras,:);
%filtro
[h1,g1,rh1,rg1]=daub(ordem_filtro)
M_feat=[]
K=niveis
%Transformada Discreta Wavelet
for i=1:length(x(1,:))
    wx=wt(x(:,i),h1,g1,K);

%baixas frequencias
for k=K:-1:1
    [dummy,ind_inferior_BAIXAS(k),dummy,ind_superior_BAIXAS(k)]...
        =bandsite(wx,K,k,0)
end
%as altas frequencias
for k=K:-1:1
    [dummy,ind_inferior_ALTAS(k),dummy,ind_superior_ALTAS(k)]...
        =bandsite(wx,K,k,3)
end
%array com comprimento dos segmentos em cada escala
for k=1:length(ind_superior_ALTAS)
    N=ind_superior_ALTAS(k)-ind_inferior_ALTAS(k)+1
    s_L(k)=N/L;
    %calcula a feature relacionada com a decomposição para os 16 sub_segmentos
end

for k=1:length(s_L)
    lim=ind_inferior_ALTAS(k)
    for kk=1:L
        sinal=wx(lim:lim+s_L(k)-1);

        lim=lim+s_L(k);
        %as features
        z(1,kk)=max(sinal); %calculo do máximo
        z(2,kk)=min(sinal); %calculo do mínimo
        z(3,kk)=std(sinal); %calculo do desvio
    end
end

```

```

        z(4, kk)=sum(sinal.^2)/length(sinal); %%%calculo do máximo%%
    end
    %copiar as features para a matriz
    M_feat=[M_feat; z]
end
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
A=[];
for i=1:4*niveis:length(M_feat)
A=[A, M_feat(i:(i+(4*niveis-1)), :)]
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
targetsA1=[];

for i=1:(length(A(1, :)))
    targetsA1=[targetsA1; tipo]

end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
f=[];
p=[];
for i=0:(niveis*4):length(A(:, 1))-(niveis*4)
f=A((((niveis*4)-(niveis_d*4)+1)+i):((niveis*4)+i), :)
p=[p; f]
end
p=p
AA=double(p)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
A.X=AA

if energia==1
    A.X=A.X(4:4:end, :)
else

end
A.y=targetsA1'

```

## A2

```

function [treino_def teste_def qt_tipos3 qt_tipos4]=...
    eeg gerar_treino_testes_def (comprimento_tipo_treino ,treino ,teste ,C)

%
%function [treino_def teste_def qt_tipos3 qt_tipos4]=...
%    eeg gerar_treino_testes_def (comprimento_tipo_treino ,treino ,teste ,C)
%
%
%Entrada:
% comprimento_tipo_treino: Quantidade de dados de cada classe para a
% elaboração do conjunto de treino
% treino: Totalidade dos dados do conjunto de treino
% teste: Totalidade dos dados do conjunto de teste
% C:
%    0: Se no conjunto de treino não estiverem contidos dados do mesmo tipo
%    do conjunto de teste
%    1: Se no conjunto de teste estiverem contidos dados do mesmo tipo
%    do conjunto de treino
%
%Saída:
%treino_def: Conjunto de treino organizado segundo o comprimento desejado
%teste_def: Conjunto de teste organizado com os restantes dados não
%            selecionados
%qt_tipos3: Quantidade de tipos de dados da classe 1 no conjunto de teste
%qt_tipos4: Quantidade de tipos de dados da classe 2 no conjunto de teste
%
%EXEMPLO:
% comprimento_tipo_treino=500;
% treino=[A BO CN DF B]
% teste=[A BO CN DF B]
% C=1;
%
%Nota:
%    Caso existam mais que um tipo de sinal no treino ou teste , a dimensão
%    dos sinais deve ser a mesma.

teste2 = [];
treino2 = [];
treino4 = [];
teste4 = [];
target_t2 = [];
target_t4 = [];
qt_tipos1 = 0;
qt_tipos2 = 0;
qt_tipos3 = 0;
qt_tipos4 = 0;

```

```

%Geração de um vector de forma aleatória
W(1).X=randperm(length(treino(1).X(1,:)))
%Calculo do nºde tipos de sinais pertencentes a cada classe no conj.treino
for i=1:(length(treino))
    if treino(i).y==1
        qt_tipos1=qt_tipos1+1;
    else
        qt_tipos2=qt_tipos2+1;
    end
end
end

%Calculo do nºde tipos de sinais pertencentes a cada classe no conj. teste
for j=1:(length(teste))
    if teste(j).y==1
        qt_tipos3=qt_tipos3+1;
    else
        qt_tipos4=qt_tipos4+1;
    end
end
end

%Geração do conjunto de treino de forma aleatória para cada tipo de classe
for i=1:qt_tipos1
    treino1.X=treino(i).X(:,W(1).X(1:comprimento_tipo_treino/qt_tipos1))
    treino2=[treino2 treino1.X]
end

for i=(1+qt_tipos1):(qt_tipos2+qt_tipos1)
    treino3.X=treino(i).X(:,W(1).X(1:comprimento_tipo_treino/qt_tipos2))
    treino4=[treino4 treino3.X]
end

%Verificação de igualdade de tipos de sinais simultaneamente no treino e no
%teste
    if C==0
        W=[];
        comprimento_tipo_treino=0
        W(1).X=randperm(length(teste(1).X))
    else

    end

%Geração do conjunto de treino de forma aleatória para cada tipo de classe
for i=1:qt_tipos3
    teste1.X=teste(i).X(:,W(1).X((comprimento_tipo_treino/qt_tipos1)...
        +1:length(teste(i).X(1,:))))
    teste2=[teste2 teste1.X]
end
end

```

```

for i=(1+qt_tipos3):(qt_tipos4+qt_tipos3)
D=randperm(length(teste(i).X))
teste3.X=teste(i).X(:,W(1).X((comprimento_tipo_treino/qt_tipos2)...
+1:length(teste(i).X(1,:))))
teste4=[teste4 teste3.X]
end

%Organização do treino e teste contendo o mesmo número de elementos de cada
%classe
treino_def.X=[treino2 treino4]
target2(1:length(treino2))=1;
target4(1:length(treino4))=2;
treino_def.y=[target2 target4]
teste_def.X=[teste2 teste4]
target_t2(1:length(teste2))=1;
target_t4(1:length(teste4))=2;
teste_def.y=[target_t2 target_t4]

```

### A3

```

function [modell,ypred1,dfce1] = ...
    classificacao_eeg(tipo_class,treino_def,teste_def)
%
%
%Entrada:
%tipo_class:
%           1: Escolha do smo como algortimo de aprendizagem
%           Outro valor: Escolha do svmlight como algortimo de aprendizagem
%treino_def: Conjunto de treino já organizado
%teste_def: Conjunto de teste já organizado
%
%Saída:
%modell: conjunto de parametros que permitem a elaboração de um hiperplano
%de separação entre os dados de uma classe e outra.
%ypred1: Etiquetas atribuidas no processo de classificação
%dfce1: Valor da função discriminante
%

%Valor de C escolhido para maximizar ou minimizar a margem
C=100;

```

```

arg=0;
if tipo_class == 1
    modell = smo(treino_def, struct('C',C,'arg',arg))
    [ypred1,dfce1] = svmclass( teste_def.X, modell )
else
    modell=svmlight(treino_def, struct('C',C,'arg',arg))
    [ypred1,dfce1] = svmclass( teste_def.X, modell )
end

```

#### A4

```

function [c1,c2]=erros_eeg(ypred1, teste_def, qt_tipos3, qt_tipos4)
%
% function [c1,c2]=erros_eeg(ypred1, teste_def, qt_tipos3, qt_tipos4)
%Entrada:
%
%          ypred1: Etiquetas atribuidas no processo de classificação
%          treino_def: Conjunto de treino já organizado
%          qt_tipos3: Quantidade de tipos de dados da classe 1
%          qt_tipos4: Quantidade de tipos de dados da classe 2
%
%Saída:
%
%          c1: Erro de classificação associado a cada tipo da classe 1
%          c2: Erro de classificação associado a cada tipo da classe 2
%
%

c1=[];
c2=[];

if qt_tipos3 > 0
tamanho_segmento1=sum(teste_def.y==1)/qt_tipos3;
for i=1:tamanho_segmento1:tamanho_segmento1*qt_tipos3

    c1=[c1; cerror(ypred1(i:((i-1)+tamanho_segmento1)), teste_def.y...
        (i:((i-1)+tamanho_segmento1)),1)]
end

```

```

else
end

if qt_tipos4 > 0
tamanho_segmento2=sum(teste_def.y==2)/qt_tipos4;
for j=tamanho_segmento1*qt_tipos3:tamanho_segmento2:length(ypred1)-1

    c2=[c2; cerror(ypred1(j+1:((j)+tamanho_segmento2)), teste_def.y...
        (j+1:((j)+tamanho_segmento2)), 2)]
end
else
end

```