

Universidade Federal do Maranhão
Centro de Ciências Exatas e Tecnologia
Programa de Pós Graduação em Engenharia Elétrica

Identificação do Padrão de Hipsarritmia em Eletroencefalogramas
Utilizando Decomposição de Sinais em Pequenas Ondas

Gean Carlos Lopes de Sousa

São Luis
2019

Universidade Federal do Maranhão
Centro de Ciências Exatas e Tecnologia
Programa de Pós Graduação em Engenharia Elétrica

Gean Carlos Lopes de Sousa

Identificação do Padrão de Hipsarritmia em Eletroencefalogramas
Utilizando Decomposição de Sinais em Pequenas Ondas

Tese de doutorado apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Maranhão, como requisito parcial para obtenção do título de Doutor em Engenharia de Eletricidade

Orientador: Prof. Allan Kardec Duailibe Barros Filho

São Luis
2019

Ficha gerada por meio do SIGAA/Biblioteca com dados fornecidos pelo(a) autor(a).
Núcleo Integrado de Bibliotecas/UFMA

Lopes de Sousa, Gean Carlos.

Identificação do Padrão de Hipsarritmia em
Eletroencefalogramas Utilizando Decomposição de Sinais em
Pequenas Ondas / Gean Carlos Lopes de Sousa. - 2019.
103 f.

Orientador(a): Allan Kardec Duailibe Barros Filho.
Tese (Doutorado) - Programa de Pós-graduação em
Engenharia de Eletricidade/ccet, Universidade Federal do
Maranhão, São Luis, 2019.

1. Funções de Gabor. 2. Gabor Sintonizadas. 3.
Hipsarritmia. 4. Wavelet. I. Duailibe Barros Filho,
Allan Kardec. II. Título.

**IDENTIFICAÇÃO DO PADRÃO DE HIPSARRITMIA EM
ELETROENCEFALOGRAMAS UTILIZANDO DECOMPOSIÇÃO DE SINAIS
EM PEQUENAS ONDAS**

Gean Carlos Lopes de Sousa

Tese aprovada em 18 de dezembro de 2019

Prof. Allan Kardec Duailibe Barros Filho, Ph.D., UFMA
(Orientador)

Prof. João Viana da Fonseca Neto, Dr., UFMA
(Membro da Banca Examinadora)

Prof. Francisco das Chagas de Souza, Dr., UFMA
(Membro da Banca Examinadora)

Prof. Antonio Oseas de Carvalho Filho, Dr., UFPI
(Membro da Banca Examinadora)

Prof. Eduardo Bouth Sequerra, Dr., UFRN
(Membro da Banca Examinadora)

Dedicatória

Dedico,

Em primeiro lugar dedico à Deus minha fortaleza, meu refúgio e proteção. A Ele que em seu imenso amor nos concedeu o dom de desvendar os segredos da vida. Muito obrigado!!!

À Alcineide Dutra Pessoa minha auxiliadora, minha amada esposa, sempre presente e disposta a me ajudar no que preciso for. A minha mãe Luzia Lopes de Sousa, mulher de garra que nunca mediu esforço para conferir aos filhos uma educação de qualidade, ao meu Pai (*in memoriam*) José de Sousa “seu maroca”, exemplo de ser humano, pessoa inteligente amiga que dedicou grande parte de seus dias na terra para dialogar comigo a fim de me educar para a vida. Meus irmãos Eunice, Izaias, Francisca e Jéssica por me darem a a honra de compôr essa família abençoada.

Agradecimentos

Agradeço,

Ao professor Allan Kardec Duailibe Barros Filho, por tudo que pacientemente tem me ensinado, pelas riquíssimas contribuições a minha formação acadêmica. Gostaria de lhe agradecer também por dividir comigo e com meus colegas de laboratório um pouquinho do seu amplo conhecimento, mais do que um agradecimento deixo minha admiração.

Ao professor Cláudio Queiroz, que foi de fundamental importância para o desenvolvimento dessa pesquisa.

À professora Patrícia Souza que não mediu esforços para que esse trabalho fosse possível.

Ao professor Ewaldo Eder Carvalho pela motivação e dedicação, fundamentais para o desenvolvimento deste trabalho de pesquisa.

À PPGEE/UFMA pela estrutura que oferece aos estudantes e pesquisadores.

E aos meus colegas de laboratório que direta ou indiretamente contribuíram para a elaboração deste trabalho.

A todos meu sincero, Obrigado!

*“Às vezes eu acho que todo preto como eu só quer
um terreno no mato só seu”.*

Brown

Resumo

Título: Identificação do Padrão de Hipsarritmia em Eletroencefalogramas Utilizando Decomposição de Sinais em Pequenas Ondas.

A epilepsia consiste em uma alteração temporária do funcionamento do cérebro. É possível, durante uma crise epilética, verificar alterações acentuadas. Em lactentes, como por exemplo, destacam-se algumas encefalopatias como a síndrome de West. A síndrome de West caracteriza-se principalmente por espasmos infantis e hipsarritmia. A hipsarritmia, descrita pela primeira vez em 1952, é caracterizada por padrão eletroencefalográfico composto de ondas lentas e espículas de projeção anárquica e de elevada tensão. O padrão de hipsarritmia também se faz presente em EEG's de crianças com microcefalia causada pelo vírus Zika. Apesar das características de hipsarritmia no eletroencefalograma estarem bem definidas, a identificação desse padrão ainda causa discordância entre especialistas. Propõe-se neste trabalho o desenvolvimento de métodos matemáticos e computacionais capazes de identificar as características básicas desse sinal elétrico patológico, a fim de auxiliar no diagnóstico e no prognóstico de pacientes epiléticos. Durante esse trabalho é apresentada a formulação matemática da Transformada de Wavelet, contínua e discreta, das funções de Gabor que serão utilizadas como núcleo dessa transformação e das funções de Gabor sintonizadas. Aplicou-se o cálculo dos três índices em sinais de eletroencefalograma de lactentes portadores do vírus Zika para identificação de hipsarritmia. Esses sinais foram adquiridos junto ao projeto do Governo do Estado do Maranhão denominado Casa Ninar. Os resultados podem evidenciar uma eficácia da metodologia em identificar o padrão de hipsarritmia, pois apresentam taxas de acerto que entre 95 a 100% na classificação de eletroencefalogramas. Na tarefa mais difícil, que é classificar trechos do EEG, as taxas de acerto variaram entre 75 a 90%.

Palavras-chave: Hipsarritmia, Wavelet, Funções de Gabor, Gabor Sintonizadas.

Abstract

Title: Identification of the Pattern of Hypsarrhythmia in Electroencephalograms Using Signal Decomposition in Small Waves.

Epilepsy is a temporary change in brain function. It is possible during an epileptic seizure to see marked changes. In infants, for example, some encephalopathies such as West syndrome stand out. West syndrome is mainly characterized by infantile spasms and hypsarrhythmia. Hypsarrhythmia, first described in 1952, is characterized by an electroencephalographic pattern composed of slow waves and high-voltage anarchic projection spicules. The pattern of hypsarrhythmia is also present in EEG's of children with microcephaly caused by Zika virus. Although the characteristics of hypsarrhythmia in the electroencephalogram are well defined, the identification of this pattern still causes disagreement among specialists. This work proposes the development of mathematical and computational methods capable of identifying the basic characteristics of this pathological electrical signal, in order to assist in the diagnosis and prognosis of epileptic patients. During this work we present the mathematical formulation of the continuous and discrete Wavelet Transform of the Gabor functions that will be used as the core of this transformation and the tuned Gabor functions. The calculation of the three indices in electroencephalogram signals of infants with Zika virus was applied to identify hypsarrhythmia. These signals were acquired from the Maranhão State Government project called Casa Ninar. The results may show the effectiveness of the methodology in identifying the pattern of hypsarrhythmia, since they present rates of accuracy that range from 95 to 100 % in the classification of electroencephalograms. In the most difficult task, which is to classify sections of the EEG, the hit rates ranged from 75 to 90 %.

Keywords: Hypsarrhythmia, Wavelet, Gabor Functions, Tuned Gabor.

Lista de Figuras

1.1	Identificação de hipsarritmia no EEG. (A) Ausência de hipsarritmia (B) Presença de hipsarritmia (Figura adaptada de Hussain et al. (2014)).	3
1.2	Identificação de hipsarritmia no EEG. (A) Ausência de hipsarritmia (B) Sem consenso sobre a presença de hipsarritmia(Figura adaptada de Hussain et al. (HUSSAIN et al., 2014)).	4
2.1	O sistema de eletrodos 10-20 e a nomenclatura dos eletrodos de EEG nesse sistema. (Figura adaptada de Geerts (2012))	12
2.2	Montagem referencial em uma página de 10 segundos	12
2.3	Montagem bipolar em uma página de 10 segundos	13
2.4	Quatro ritmos normais típicos do cérebro.(Figura adaptada de Sanei e Chambers. (2007))	14
2.5	Classificação Internacional das Crises Epiléticas da ILAE de 1981 (YACUBIAN; KOCHEN, 2014)	15
2.6	Frequência dos sintomas premonitórios, adaptadp de (SCARAMELLI et al., 2009)	16
2.7	Presença de espículas no EEG (PACOLA et al., 2012)	19
2.8	Presença de ondas agudas no EEG (PACOLA et al., 2012)	19
2.9	Classificação de espículas segundo a amplitude (adaptado de Janca et al. (2014)	20
2.10	Registro de EEG com hipsarritmia fragmentada (YACUBIAN; KOCHEN, 2014)	21
2.11	Hipsarritmia com aumento da sincronização inter-hemisférica (YACUBIAN; KOCHEN, 2014)	22
2.12	Hipsarritmia assimétrica (YACUBIAN; KOCHEN, 2014)	23
2.13	Hipsarritmia com anormalidade focal consistente (YACUBIAN; KOCHEN, 2014)	23
2.14	Hipsarritmia com episódios de atenuação de tensão (YACUBIAN; KOCHEN, 2014)	24
2.15	Hipsarritmia com atividade lenta de amplitude elevada, bilateral e assíncrona (HRA-CHOVY; FROST; KELLAWAY, 1984).	25
2.16	Ciclos de transmissão do vírus Zika (RUSSO; JUNGSMANN; BRAGA, 2017).	26
3.1	Hiperplano de separação e semiespaços. Adaptado de (BOYD; VANDENBERGHE, 2004)	29
3.2	Hiperplano Ótimo (TAKAHASHI, 2012)	30

3.3	Primeiro caso de violação da condição 3.15 (adaptado de Takahashi (2012)) . .	32
3.4	Segundo caso de violação da condição 3.15 (adaptado de (LORENA; CARVAHO, 2003; TAKAHASHI, 2012))	33
3.5	Terceiro caso de violação da condição 3.15 (adaptado de (LORENA; CARVAHO, 2003; TAKAHASHI, 2012))	33
3.6	Exemplo de variáveis soltas (LORENA; CARVAHO, 2003; TAKAHASHI, 2012)	34
3.7	Mapeamento através de uma função kernel (adaptado de (TAKAHASHI, 2012))	35
4.1	Exemplos de Wavelet de Haar (MONTEIRO, 2019)	39
4.2	Tipos de wavelets (adaptado de Stéphane (2009a))	39
4.3	Caixas de Heisenberg de duas wavelets (adaptado de Stéphane (2009a))	40
4.4	Exemplos de funções de Gabor. (A),(B) e (C) deslocamento do parâmetro $\bar{t}$ no tempo; (D),(E) e (F) impacto da mudança do parâmetro σ	42
4.5	Decomposição em funções de Gabor sintonizadas	46
5.1	Delineamento da pesquisa.	49
5.2	EEG com identificação de hipsarritmia e descargas multifocais.	51
5.3	Janelamento do sinal de EEG	52
5.4	Extração de características e classificação do EEG.	53
5.5	Comparação entre uma espícula e uma função de Gabor.	53
5.6	Extração de características e classificação do EEG.	54
5.7	Extração de características e classificação do EEG.	55
5.8	Método de validação cruzada	57
6.1	Boxplot com as médias dos índices de hipsarritmia: A) coif B) db3 C) Haar D) sym4	60
6.2	Boxplot com as log energia dos índices de hipsarritmia: A) coif B) db3 C) Haar D) sym4	61
6.3	Boxplot com as médias dos índices de hipsarritmia. Parâmetros da função de Gabor: A) $\sigma^2 = 20$, $\omega = 1,02$ e $\phi = 5$, B) $\sigma^2 = 5$, $\omega = 1,40$ e $\phi = 5$, C) $\sigma^2 = 15$, $\omega = 2,68$ e $\phi = 5$ e D) $\sigma^2 = 10$, $\omega = 3,20$ e $\phi = 5$	63
6.4	Boxplot com as log energia dos índices de hipsarritmia. Parâmetros da função de Gabor: A) $\sigma^2 = 20$, $\omega = 1,02$ e $\phi = 5$, B) $\sigma^2 = 5$, $\omega = 1,40$ e $\phi = 5$, C) $\sigma^2 = 15$, $\omega = 2,68$ e $\phi = 5$ e D) $\sigma^2 = 10$, $\omega = 3,20$ e $\phi = 5$	64
6.5	Exemplo de decomposição em funções de Gabor: A) comparação entre o sinal original e o aproximado para 100 funções de Gabor sintonizadas B) representação esparsa do sinal original	66
6.6	Boxplot com as médias e log energia dos índices baseados na decomposição de funções de Gabor sintonizadas: A) Média dos índices B) Log energia dos índices	67

Lista de Tabelas

3.1	Principais Kernels	35
5.1	Características das crianças que tiveram os EEG's utilizados nas simulações: PC- Perímetro Cefálico; M- Masculino; F- Feminino; SD- Sem Dados; SL- Sem Acesso ao Laudo.	50
5.2	Tipos de erros em testes de hipóteses	56
6.1	Teste de Hipótese.	62
6.2	Classificação quando as médias de I_k^w são utilizadas como entrada da SVM. . . .	62
6.3	Classificação quando as log energias de I_k^w são utilizadas como entrada da SVM.	62
6.4	Avaliação dos resultados para diferentes funções de Gabor usando somente as médias dos índices como entrada da SVM.	65
6.5	Avaliação dos resultados para diferentes funções de Gabor usando somente as E_{LgEn} dos índices como entrada da SVM.	65
6.6	Avaliação dos resultados de classificação	67
6.7	Duração total dos segmentos de cada grupo identificado	68
6.8	Resultados da classificação de Hipsarritmia e Descargas focais	68
6.9	Resultados da classificação de Descargas focais e Sem marcação	69
6.10	Resultados da classificação de Hipsarritmia e Sem marcação	69

Sumário

1	Introdução	2
1.1	Trabalhos Relacionados	4
1.1.1	Trabalhos Relacionados - Área Clínica	4
1.1.2	Trabalhos Relacionados - Área Computacional	7
1.2	Motivação	9
1.3	Objetivos	9
1.3.1	Objetivo Geral	9
1.3.2	Objetivos Específicos	9
1.4	Organização do Trabalho	10
2	O Eletroencefalograma e a Hipsarritmia	11
2.1	Introdução	11
2.2	O Eletroencefalograma	11
2.3	Epilepsia	14
2.4	Hipsarritmia	18
2.4.1	Espículas e ondas agudas	18
2.4.2	Variações do padrão de hipsarritmia	20
2.4.3	Hipsarritmia e a Síndrome congênita do Zika Vírus	25
2.5	Considerações	27
3	Máquina de Vetores de Suporte - SVM	28
3.1	Introdução	28
3.2	Classes Linearmente Separáveis	29
3.2.1	Hiperplano Ótimo	30
3.3	Classes Não Linearmente Separáveis	32
3.3.1	Função Kernel	35
3.4	Considerações	36
4	Decomposição de Sinais em Pequenas Ondas	37
4.1	Introdução	37
4.2	Wavelets	38

4.3	Transformadas de Wavelet	40
4.4	Funções de Gabor	42
4.5	Funções de Gabor Sintonizadas	43
4.5.1	O Algoritmo Matching Pursuit-MP	45
4.5.2	Decomposição em Funções de Gabor Sintonizadas	45
4.6	Considerações	46
5	Materiais e Métodos para Identificação de Hipsarritmia	48
5.1	Introdução	48
5.2	Delineamento da Pesquisa	49
5.2.1	Organização da Base de dados	50
5.2.2	Índices de Hipsarritmia	51
5.2.3	Métricas de Avaliação	55
5.3	Considerações	58
6	Resultados	59
6.1	Identificação de EEG's com Hipsarritmia	59
6.1.1	Resultados do cálculo e aplicação do índice de Hipsarritmia baseado na DWT	60
6.1.2	Resultados do cálculo e aplicação do índice baseado na DWT com núcleo de Gabor	63
6.1.3	Resultados do cálculo e aplicação do índice Baseado na Decomposição em Funções de Gabor Sintonizadas	65
6.2	Identificação de Hipsarritmia e Descargas Focais em Eletroencefalogramas	67
6.3	Considerações	69
7	Discussões	70
8	Conclusões e Perspectivas	73
	APÊNDICE	80

Introdução

Estima-se que a epilepsia acometa aproximadamente 70 milhões de pessoas em todo o mundo e que sua incidência ocorra principalmente em bebês e em pessoas idosas (THIJS et al., 2019). A definição de epilepsia está relacionada a predisposição duradoura para gerar crises epiléticas. As crises epiléticas, por sua vez, são definidas a partir de três elementos: modo de início e término, manifestações clínicas e ocorrência de sincronização neuronal aumentada. O modo de início e término de uma crise, são marcos que podem ser delimitados pelo comportamento do indivíduo ou pelo eletroencefalograma (EEG) (SAKAMOTO, 2007). O EEG consiste no registro da atividade elétrica cerebral capaz de expressar tanto os potenciais elétricos normais quanto aqueles característicos dos diversos tipos de crises e síndromes epiléticas.

O início das crises epiléticas podem ocorrer em diversos estágios da vida de uma pessoa e, dependendo da idade na qual as crises se iniciam, a epilepsia recebe uma classificação específica. Em lactentes, por exemplo, destaca-se a síndrome de West (SPENNER et al., 2019). Esta síndrome é uma encefalopatia epilética relacionada especificamente a crianças com menos de 1 ano de idade e caracteriza-se por um tipo específico de crise denominada “espasmos epiléticos” e anormalidades grosseiras no EEG, o chamado padrão hipsarrítmico ou hipsarritmia (KWAN; SANDER, 2004; MUIR et al., 2019).

Os espasmos epiléticos consistem em movimentos musculares agrupados que podem ser de flexão ou extensão e que, normalmente, tem duração menor do que 2 segundos (YACUBIAN; KOCHEN, 2014). Já o padrão eletrencefalográfico denominado hipsarritmia consiste em uma atividade anormal no EEG observada em algumas formas de epilepsia. Esse padrão é caracterizado por descargas de ondas com alta amplitude, multifocais e desorganizadas (PHILIPPI et al., 2008; MARCHI et al., 2015; LIMA; LONGO; MOREIRA, 1974). Dependendo da quantidade, frequência e localização dessas ondas, a hipsarritmia recebe uma classificação específica. A hipsarritmia foi descrita pela primeira vez por Gibbs e Gibbs em 1952 no simpósio de Marselha e vem sendo estudada até os dias atuais (GIBBS; GIBBS, 1952).

Recentemente, Carvalho et al. (2017) verificaram que o EEG de algumas crianças com microcefalia e Síndrome congênita do Zika Vírus apresentaram o padrão de hipsarritmia. Nesses casos, segundo Carvalho et al. (2017), a identificação de hipsarritmia é ainda mais delicada pois esses EEG's são repletos de anormalidades.

A Figura 1.1 é um exemplo de ausência e presença de hipsarritmia. Neste caso, em específico,

é visualmente perceptível a diferença entre um sinal de EEG sem o padrão hipsarrítmico e um sinal com hipsarritmia. No entanto, vale ressaltar que, mesmo para especialistas, nem sempre é uma tarefa fácil identificar quando está ocorrendo a hipsarritmia apenas observando o Eletroencefalograma.

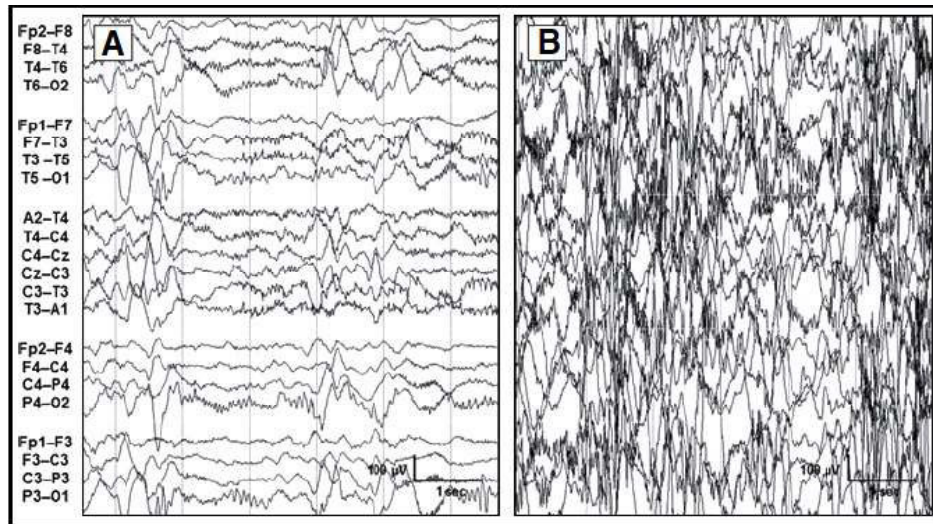


Figura 1.1: Identificação de hipsarritmia no EEG. (A) Ausência de hipsarritmia (B) Presença de hipsarritmia (Figura adaptada de Hussain et al. (2014)).

Apesar da hipsarritmia ser objeto de estudo de muitos especialistas e atualmente esse termo ser amplamente utilizado na prática clínica, o consenso entre avaliadores sobre sua identificação em sinais de EEG é baixo, principalmente quando a diferença entre um sinal dito sem alterações e um com alterações é pequena (HUSSAIN et al., 2014). Em alguns casos, como por exemplo o apresentado na Figura 1.2, a presença de hipsarritmia em B não é consenso entre os especialistas que avaliaram esse sinal (HUSSAIN et al., 2014).

Tendo em vista que a identificação de hipsarritmia tende a ser uma tarefa laborosa, e que em alguns casos implica em discordância entre os especialistas é que propõe-se neste trabalho o desenvolvimento de métodos matemáticos e computacionais capazes de identificar as características básicas desse sinal elétrico patológico, a fim de auxiliar no diagnóstico e no prognóstico de pacientes epiléticos. A Figura 1.2 ilustra trechos de sinais de eletroencefalogramas que causam discordância sobre a presença e ausência do padrão de hipsarritmia.

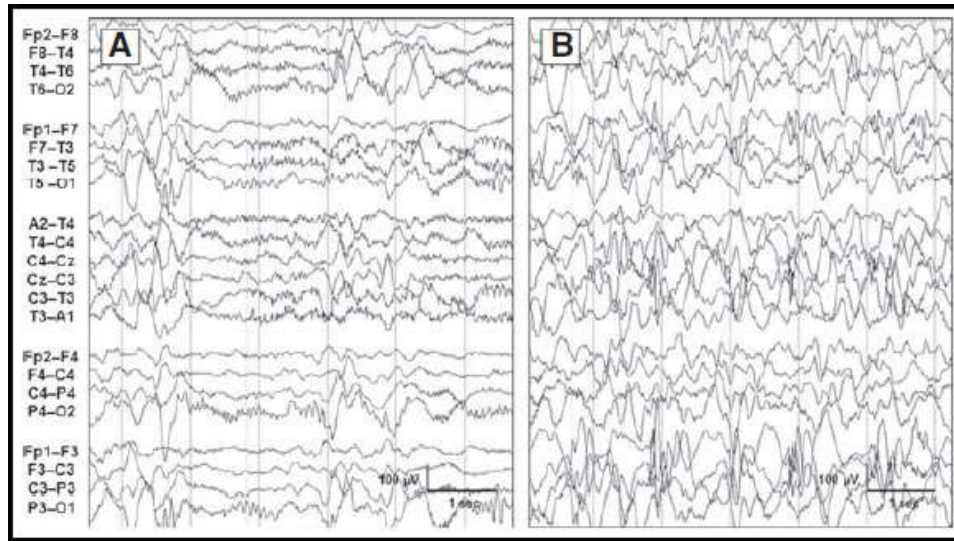


Figura 1.2: Identificação de hipsarritmia no EEG. (A) Ausência de hipsarritmia (B) Sem consenso sobre a presença de hipsarritmia (Figura adaptada de Hussain et al. (HUSSAIN et al., 2014)).

1.1 Trabalhos Relacionados

Como esse trabalho tem um viés interdisciplinar com uma integração entre a engenharia e a neurociência, os trabalhos relacionados foram divididos em dois subgrupos: os referentes a área clínica e aqueles referentes a área matemática ou computacional. Destaque para os trabalhos desenvolvidos na área computacional, pois em apenas Smith et al. (2017) foi encontrada pesquisa com objetivo semelhantes aos propostos aqui.

1.1.1 Trabalhos Relacionados - Área Clínica

Os primeiros trabalhos a apontarem o padrão de hipsarritmia em eletroencefalogramas datam de 1952. No trabalho de Gibbs e Gibbs (1952) foi realizada uma descrição das principais características desse tipo de traçado e as mesmas denominou-se de hipsarritmia típica. Esse padrão de ondas consiste em um traçado eletroencefalográfico que, em vigília, se apresenta como uma sucessão contínua de ondas lentas e pontas de grande amplitude.

As características do padrão de hipsarritmia típica foram bem definidas por Gibbs e Gibbs (1952) no simpósio de Marselha. Porém, ao longo do tempo diversos pesquisadores verificaram a existência de variações desse tipo de padrão, ou seja, sinais eletroencefalográficos que carregavam consigo alguns aspectos apontados na descrição original de hipsarritmia também continham características variáveis (KASINATHAN et al., 2019). Nesse contexto então, foram definidas diversos tipos de hipsarritmia modificada (HRACHOVY; FROST; KELLAWAY, 1984).

Mytinger, Weber e Vidaurre (2018) avaliaram a presença de ondas lentas e de alta amplitude no EEG de cem crianças sem epilepsia ou hipsarritmia, com idades variando entre 3 a 18 meses. Três eletroencefalógrafos realizaram a avaliação dos EEGs. O padrão 10-20 foi utilizado e uma montagem bipolar. A concordância entre os especialistas foi avaliada através do coeficiente de

correlação intraclasse $\tau = 0,97$. Foram encontradas ondas lentas recorrentes de $200 \mu V$ em 17%, 200 a $299 \mu V$ em 49%, 300 a $399 \mu V$ em 30%, 400 a $499 \mu V$ em 3% e $500 \mu V$ em 1%. Mytinger et al.(2018) concluíram então que, apesar da hipsarritmia geralmente incluir ondas lentas de fundo e ondas de alta amplitude, 83% das crianças normais apresentaram ondas lentas de $200 \mu V$ e 34% de $300 \mu V$.

O trabalho realizado por Mytinger, Weber e Vidaurre (2018) ressalta um problema importante, pois a detecção em grande quantidade de ondas lentas em EEG's de crianças saudáveis indica a possibilidade de diagnóstico incorreto sobre a presença de hipsarritmia.

Altunel, Sever e Altunel (2015) analisaram a evolução de EEG pré-hipsarrítmicos para paroxismos esporádicos de hipsarritmia. Utilizaram Registros de EEGs de 42 pacientes antes e subsequente aos espasmos infantis. Nesse trabalho foi estabelecido um índice com base na frequência de paroxismos esporádicos de hipsarritmia. Esse índice foi proposto como mecanismo de auxílio na verificação da recorrência de espasmos e deterioração do desenvolvimento psicomotor.

A descoberta de um índice capaz de identificar padrões evolutivos em EEG's que ainda não apresentaram o padrão de hipsarritmia mas que evoluíram para tal patologia é sobretudo, um achado importante, pois pode auxiliar no tratamento precoce das patologias relacionadas, ou que culminam, no padrão hipsarrítmico.

Philippi et al. (2008) realizaram análise de eletroencefalografias pré-hipsarrítmicas com o objetivo de permitir um tratamento precoce, com relação a síndrome de West. Foram utilizados EEGs de 39 neonatos com síndrome de West . Um sistema de classificação de três tipos de EEG correspondendo aos estados clínicos anteriores à ocorrência da síndrome de West foi desenvolvido. Três estados clínicos e os correspondentes tipos de EEG foram classificados. O estado clínico 1 apresentou descargas epiléticas (multi-focais) menor que 50% do tempo de registro de EEG. O estado clínico 2 apresentou descargas epiléticas bi-hemisféricas maior do que 50% do tempo de registro. O estado clínico 3 foi caracterizado por hipsarritmia. A concordância entre os especialistas foi avaliada através do coeficiente de correlação intraclasse de $\tau = 0,67$ em média. Concluíram nesse trabalho que, bebês com síndrome de West podem ser identificados, com segurança, várias semanas antes da ocorrência de hipsarritmia por um padrão típico de EEG (tipo 2).

O trabalho de Philippi et al.(2008), assim como o de Altunel, Sever e Altunel (2015) podem representar um avanço no que tange a possibilidade de estudos de intervenção precoce sobre a possível evolução de um quadro cínico que tenha como base eletroencefalográfica a hipsarritmia. Além disso esses trabalhos podem servir como inspiração para elaboração de mecanismos computacionais capazes de identificar uma tendência à hipsarritmia.

Humphrey et al. (2008) desenvolveram a Escala de Gravidade da Epilepsia da Primeira Infância, para quantificar a gravidade da epilepsia em lactentes. Nesse trabalho Foram utilizados EEGs de vinte lactentes com idades entre 11-36 meses. Foram utilizadas seis potenciais marcadores de gravidade da epilepsia: período de tempo durante o qual ocorreram convulsões; frequência de crises; número e tipos de crises; ocorrência e duração do status epiléticos; número de medicações anticonvulsivantes utilizadas e resposta ao tratamento. As seis variáveis receberam um escore (0 a 3). Para cada criança, essas variáveis foram pontuadas por três avali-

adores independentes. O escore proposto, quando avaliado aos 12 meses, foi significativamente preditivo dos escores avaliados aos 24 e 36 meses.

Kramer, Sue e Mikati (1997) propuseram um escore para determinar a frequência e o significado das características do EEG com hipsarritmia. Foram analisado registros de 53 pacientes com espasmos infantis, verificando e determinando pontuação para as seguintes anormalidades: desorganização de fundo, lentidão, alta amplitude e atividade de espículas. Foi verificado nesse trabalho a presença ou ausência de cada um dos seguintes padrões: atividade normal do sono, assimetria inter-hemisférica e sincronização. Com base nessa quantificação foi calculada uma pontuação total indicando a gravidade da hipsarritmia para cada registro. Os padrões de hipsarritmia ocorreram com frequência em até 69% dos registros.

Os artigos de Humphrey et al. (2008) e Kramer, Sue e Mikati (1997) são bastante semelhantes quanto ao objetivo de propor um escore (ou escala) capaz de de auxiliar na investigação sobre a gravidade da epilepsia no desenvolvimento infantil. Nesses trabalhos percebe-se também que ambas metodologias são bastante trabalhosas e exigem muito tempo de análise por parte do especialista. Esse fato instiga o desenvolvimento de trabalhos que visem automatizar esse processo.

Uma aplicação da metodologia de Kramer, Sue e Mikati (1997) pode ser encontrada no trabalho de Sehgal et al. (2016), que por sua vez, avaliou o impacto da gravidade da epilepsia clínica e da severidade da hipsarritmia em crianças tratadas com síndrome de West. Foram analisados EEGs de trinta e três crianças, com idades entre 1 e 5 anos, com espasmos infantis.

A identificação de hipsarritmia através de inspeção visual do eletroencefalograma nem sempre é uma tarefa simples, pois em algumas situações o diagnóstico dessas características eletroencefalográficas causa discordância entre especialistas. Com o objetivo de abordar essa problemática dois artigos se destacam. No primeiro Hussain et al. (2014) verificaram a confiabilidade inter-avaliador na identificação de hipsarritmia. Nesse trabalho os autores confirmaram a hipótese de que existem padrões no EEG que causam discordância.

Hussain et al. (2014) entregaram 22 amostras de EEG de pacientes com espasmos infantis a seis especialistas diferentes. Os avaliadores determinaram se cada EEG era anormal e se a hipsarritmia estava presente ou ausente e avaliaram as seguintes características: tensão, organização, descargas epileptiformes, lentificação, atenuações interictais, simetria e sincronia. Além disso, os avaliadores indicaram seu nível de confiança para cada avaliação. As estatísticas multiponto kappa κ foram calculadas. Os resultados obtidos foram de $\kappa = 0.89$ para classificação de normal ou alterado, $\kappa = 0.40$ para presença ou ausência de hipsarritmia, $\kappa = 0.37$ para alta voltagem, $\kappa = 0.22$ para desorganização, $\kappa = 0.68$ para descargas epileptiformes multifocais, $\kappa = 0.20$ para atenuações da tensão interictal, $\kappa = 0.26$ para assimetria e $\kappa = 0.08$ para assincronia. Com esses resultados os autores concluíram que há uma grande necessidade de medidas práticas, confiáveis e imparciais, capazes de colaborar na identificação de hipsarritmia.

Os resultados obtidos por Hussain et al. (2014) evidenciaram um problema preocupante pois a identificação errada de uma patologia pode submeter um indivíduo a um tratamento desnecessário, caso o diagnóstico seja um falso positivo, ou ainda eximir de tratamento indivíduos que de fato necessitam de tratamento, caso este seja um falso negativo. Nesse contexto, Mytinger et al. (2015) apresentaram um método chamado de "BASED" (Burden of Amplitudes and

Epileptiform Discharges) a fim de aumentar a concordância interavaliador na identificação do padrão de hipsarritmia (ver anexo I).

O método apresentado por Mytinger et al. (2015) consiste na aplicação de um questionário referente a identificação de descargas epileptiformes interictais e as amplitudes de fundo. Para tais características é atribuída uma pontuação que posteriormente é utilizada para determinação de um escore. O escore obtido no questionário é comparado com uma tabela de valores e só então define-se a existência ou não do padrão de hipsarritmia. Utilizando esse método os autores verificaram que os especialistas concordaram sobre presença ou ausência de hipsarritmia em 37 dos 40 registros ($\kappa = 92.5\%$).

Apesar da metodologia apresentada por Mytinger et al. (2015) culminar em um aumento da concordância entre avaliadores, é importante salientar que para utilização do mesmo o especialista tem que dispor de bastante tempo de análise pois necessita fazer contagem de diversos tipos de ondas, preencher um questionário e computar um escore.

O padrão de hipsarritmia está ligado principalmente à síndrome de West (ZHANG et al., 2019; BRINCIOTTI; MATRICARDI, 2019). Porém, recentemente foi encontrado esse padrão em exames de EEG de crianças com microcefalia infectadas pelo vírus Zika. Os primeiros achados correspondem a descrição de Carvalho et al. (2017) que encontraram esse padrão em 11 exames de crianças com microcefalia infectadas pelo vírus Zika.

Além de Carvalho et al. (2017), outros trabalhos também apontaram para presença de hipsarritmia em EEG's de crianças com microcefalia associada a infecção pelo vírus Zika. Entre esses trabalhos podem ser citados, por exemplo, Kanda et al. (2018) e Linden et al. (2018). Nesses trabalhos fica evidente a existência de uma relação entre microcefalia/Zika e o padrão de hipsarritmia. Vale ressaltar porém que tal relação não consiste em uma implicação lógica, ou seja, nem todos os casos de microcefalia/Zika culminam na existência de hipsarritmia.

1.1.2 Trabalhos Relacionados - Área Computacional

Janca et al. (2014) desenvolveram um algoritmo de detecção de descargas. O algoritmo é baseado na modelagem das distribuições estatísticas do envelope do sinal. O objetivo desse trabalho foi identificar os sinais que contém descargas epiléticas e as diferenciar de sinais com atividade de segundo plano.

O detector proposto por Janca et al. (2014) foi aplicado a dados intracranianos não epiléticos de pacientes com dor facial intratável. Os experimentos revelaram a existência de transientes agudos com formas de onda remanescentes de descargas interictais, e essas formas de ondas podem representar fontes biológicas de detecções de falso positivo. A identificação desses transientes permitiu aos autores desenvolver e propor etapas de processamento secundário, e assim excluir esses transientes, melhorando a especificidade do detector.

A identificação computadorizada de descargas epileptiformes é uma tarefa importante no auxílio ao diagnóstico de diversas patologias. Nesse sentido o trabalho de Janca et al. (2014) aparenta ser bastante promissor. No entanto, percebe-se algumas lacunas na apresentação dos resultados em Janca et al. (2014), pois carece de comparações com outros trabalhos. Ressalta-se ainda o fato da metodologia ser aplicada apenas a sinais intracranianos, ou seja, para aplicação da mesma são necessários procedimentos clínicos bastante invasivos.

Lopes et al. (2013) apresentaram o uso de Análise de Componentes Independentes (ICA) para extração de características no processo de clusterização de espículas. O percentual médio de acertos obtidos utilizando ICA foi de 90.86 %. Esse trabalho tem sua importância por ter sido aplicado em EEG's posicionados no couro cabeludo diferentemente de Janca et al. (2014) e ainda por possibilitar a identificação dos diferentes tipos de espículas presentes em um EEG.

É comum encontrar na literatura o desenvolvimento de algoritmos que tem por objetivo auxiliar os especialistas no que tange a problemas referentes a epilepsia, porém quando os problemas são restringidos aos padrões de hirsarritmia três trabalhos se destacam.

No primeiro trabalho, desenvolvido por Trairruengsakul et al. (2015), é utilizada decomposição em tempo e frequência e fatoração de matriz não negativa para extrair características, e assim localizar descargas associadas a espasmos infantis na presença de hirsarritmia. Os resultados obtidos nos testes com lactentes foram de 86 % de verdadeiros positivos, 14 % de falsos negativos e 53 % de falsos positivos.

Com o objetivo de resolver o problema referente a alta taxa de falsos positivos do primeiro algoritmo, Trairruengsakul et al. (2017) desenvolveram uma metodologia baseada em representação de tempo-frequência de alta resolução, combinada com um método de decomposição matricial para caracterizar e localizar descargas epileptiformes associadas a espasmos infantis na presença de hirsarritmia. Essa metodologia foi aplicada a sinais de EEG de cinco lactentes com espasmos infantis e os resultados obtidos foram de 94,30 % de acurácia com especificidade de 93,83% e sensibilidade de 94,76%.

Os trabalhos desenvolvidos por Trairruengsakul et al. (2015) e por Trairruengsakul et al. (2017) se completam pois o segundo é um complemento do primeiro e ambos são aplicados em um estágio posterior a identificação de hirsarritmia. Nesses artigos não se faz, por exemplo, uma classificação prévia dos EEG's com e sem o padrão hirsarrítmico.

Recentemente Smith (2019) definiram biomarcadores computacionais para EEG's, afim de identificar a presença de hirsarritmia. A proposta do autor consistiu em quantificar características básicas do sinal como amplitude, espectro de potência e entropia de Shannon. Com a quantificação dessas características foi possível, através de testes estatísticos, identificar diferenças significativas entre os pacientes com espasmos epilépticos/hirsarritmia e os indivíduos do grupo controle. Nesse trabalho não foram aplicadas técnicas de aprendizado de máquina a fim de realizar classificação dos EEG's. A metodologia apresentada limitou-se apenas a aplicação de testes de hipótese.

Por outro lado foram encontrados em Smith et al. (2017) procedimentos de classificação de EEG's. Nesse trabalho os autores utilizaram autocorrelação e análise de flutuação retificada para verificar diferenças entre os EEGs de 21 pacientes com espasmos infantis e 21 EEGs controle. O processo de classificação (presença ou ausência de hirsarritmia) consistiu na utilização de uma máquina de vetores de suporte. O resultado da classificação foi de 92% de acertos. Esse trabalho é importante pois tem uma temática semelhante ao que se propõe no desenvolvimento desta tese, com o diferencial de que os EEG's dos dois trabalhos (artigo de Smith et al. (2017) e esta tese) serem oriundos de patologias diferentes.

É importante observar que os trabalhos de Smith et al. (2017) e de Smith (2019), apesar de apresentarem boas taxas de acerto na identificação de hirsarritmia não foram aplicados em

EEG's de crianças com microcefalia/Zika. Acredita-se que a identificação de padrões de onda em EEG's de crianças portadoras da Síndrome Congênita do Zika Vírus não é uma tarefa fácil pois mesmo as crianças sem epilepsia apresentam modificações importantes no eletroencefalograma.

1.2 Motivação

A ausência de mecanismos computacionais capazes de auxiliar o especialista em eletroencefalografia, na análise e classificação do padrão de hipsarritmia no EEG, é um fator importante no que diz respeito a motivação de realização desse trabalho.

É perceptível que a dificuldade em caracterizar esse padrão é um fator relevante pois o melhor método é a contagem de espículas combinada a um escore produzido por um questionário, o que com certeza demanda um tempo razoável para se chegar a um diagnóstico consensual (MYTINGER et al., 2015).

Uma outra dificuldade para identificar o padrão de hipsarritmia no EEG é o fato de existirem pelo menos seis padrões diferentes, com características que, de maneira sutil, representam a presença ou não de espasmos infantis (YACUBIAN; KOCHEN, 2014).

Portanto, é a possibilidade de auxiliar o profissional que analisa o EEG hipsarítmico que motiva a realização desse trabalho. Vale ressaltar que os métodos que serão apresentados nesse trabalho vislumbram a possibilidade de oferecer uma alternativa computacional como auxílio ao especialista em eletroencefalografia.

1.3 Objetivos

Tendo em vista a problemática descrita, os trabalhos relacionados tanto na área clínica quanto na área computacional e as questões de pesquisa apresentadas anteriormente, este trabalho possui como objetivos gerais e específicos os pontos relatados a seguir:

1.3.1 Objetivo Geral

Apresentar uma metodologia de decomposição de sinais em pequenas ondas que seja capaz de auxiliar na identificação do padrão eletroencefalográfico de hipsarritmia.

1.3.2 Objetivos Específicos

- Propor um mecanismo de representação de sinais em tempo, frequência e de forma esparsa, utilizando funções de Gabor sintonizadas ;
- Propor índices de hipsarritmia que sejam capazes de evidenciar diferenças significativas entre os EEG's com e sem hipsarritmia;
- Aplicar a metodologia proposta em EEGs de lactentes diagnosticados com microcefalia infectados com o vírus Zika;
- Classificar os EEGs em com hipsarritmia e sem hipsarritmia;

- Aplicar a metodologia proposta em EEGs identificados com hirsarritmia para classificação de trechos com hirsarritmia e com descargas multifocais;
- Testar a aplicabilidade dos índices propostos na identificação de trechos com descargas multifocais e trechos sem alteração.

1.4 Organização do Trabalho

No capítulo 2 são apresentados os conceitos de EEG, Epilepsia e Hirsarritmia. São apresentadas nesse capítulo as variações do padrão hirsarrítmico e um pequeno histórico de como a hirsarritmia e o vírus Zika estão relacionados. No capítulo 3 é apresentada uma breve abordagem sobre Máquina de Vetores de Suporte, pois essa técnica tem apresentado, em diversos tipos de problema de classificação, uma boa capacidade de generalização, ou seja, apresenta um bom desempenho em dados novos. Conceitos e exemplos de Wavelet, Transformada de Wavelet, funções de Gabor e funções de Gabor sintonizadas são discutidos no capítulo 4. Ainda no capítulo 4 é apresentada uma proposta de decomposição de sinais em pequenas ondas. No capítulo 5 são apresentados mecanismos metodológicos propostos para solução da problemática e os critérios de avaliação das possíveis soluções.

Os resultados obtidos são expostos no capítulo 6, através de gráficos e tabelas comparativas. No capítulo 7 uma discussão sobre os resultados é apresentada. O último capítulo desse trabalho, a saber o capítulo 8, aborda as conclusões observadas e os desafios a serem superados em trabalhos futuros.

O Eletroencefalograma e a Hipsarritmia

2.1 Introdução

Em 1875 o cientista Richard Caton (1842-1926) colocou dois eletrodos sobre o couro cabeludo de um ser humano e registrou, pela primeira vez, a atividade cerebral na forma de sinais elétricos. Desde então, os conceitos de eletroencefalografia (eletro - atividades elétricas, encefalo - cabeça e grama ou grafia - desenho ou escrita), foram combinados de modo que o termo eletroencefalograma (EEG) passou a ser usado para designar atividade neural do cérebro (SANEI; CHAMBERS., 2007).

O EEG é o registro da atividade elétrica cerebral, captada no couro cabeludo, ou intracranial (nesse caso usa-se o termo iEEG). Esse exame é bastante utilizado na investigação e diagnóstico das epilepsias, pois expõe os potenciais elétricos anormais característicos dos diversos tipos de crises e síndromes epiléticas (SANEI; CHAMBERS., 2007).

Apesar de diversas técnicas de neuroimagem estrutural e funcional terem seu advento nas ultimas décadas, o EEG nunca perdeu seu papel no auxílio à avaliação de diversas síndromes, pois conta com a vantagem da simplicidade técnica e do baixo custo, se comparado a outros métodos complementares de diagnóstico (CUKIERT et al., 1999).

O uso do EEG no diagnóstico de síndromes epiléticas infantis também é fator importante. As relações entre EEG e espasmos infantis pode ser identificado através do padrão de hipsarritmia. Os conceitos e as relações entre EEG e hipsarritmia são os principais objetivos deste capítulo.

2.2 O Eletroencefalograma

A atividade elétrica do cérebro é medida por eletrodos colocados no couro cabeludo, como mostrado na Figura 2.1, na qual é ilustrado o sistema de eletrodos 10-20. O sistema 10-20 é baseado na estratégia geral de medir a distância entre dois pontos anatômicos fixos, como o *nasion* (ponto onde a ponta do nariz encontra a testa) e o *inion*(ponto no occipício). Determina-se uma linha imaginária entre esses dois pontos e, em seguida, coloca-se eletrodos em intervalos de 10% ou 20% ao longo dessa linha .

Os nomes dos eletrodos estão relacionados ao lóbulo ou área na qual o eletrodo está posicionado: F - frontal, Fp - frontal polar, T - temporal, C-central, P - parietal, O - occipital, A - auricular.

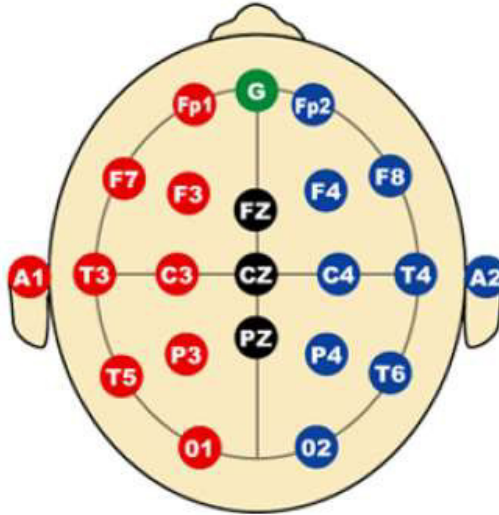


Figura 2.1: O sistema de eletrodos 10-20 e a nomenclatura dos eletrodos de EEG nesse sistema. (Figura adaptada de Geerts (2012))

A ordem e escolha dos canais exibidos no EEG é conhecido como montagem. As duas montagens mais utilizadas são a referencial e a bipolar. Uma montagem referencial compara cada eletrodo com um ponto de referência em algum outro lugar do corpo. Espera-se que esse ponto de referência seja neutro, podendo ser um eletrodo colocado no nariz, no queixo ou lóbulos da orelha (GEERTS, 2012; SANEI; CHAMBERS., 2007). Em algumas vezes a referência (ou sinal de referência) é a média de todos os sinais dos demais eletrodos do couro cabeludo (Figura 2.2).

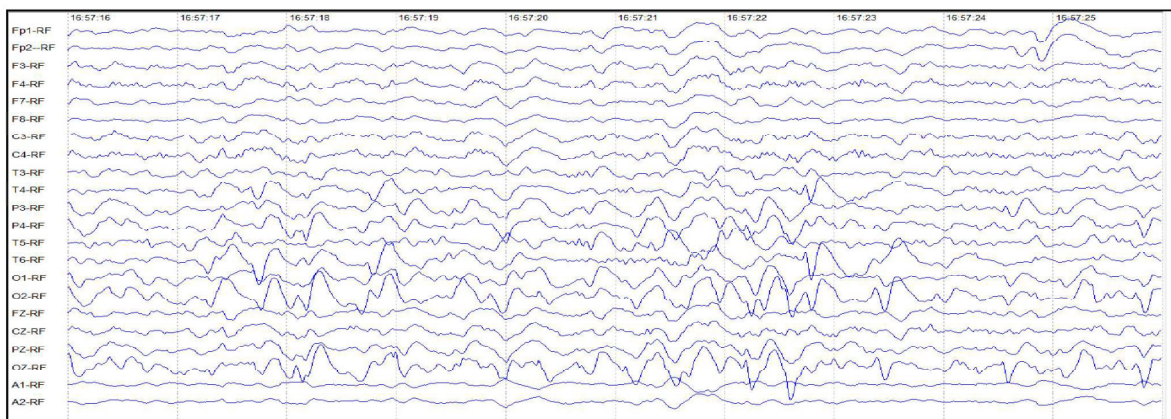


Figura 2.2: Montagem referencial em uma página de 10 segundos

Por outro lado, em uma montagem bipolar, cada canal representa a diferença de tensão entre dois eletrodos adjacentes ($F_{p1} - F_3, F_3 - C_3, P_3 - O_1$ e etc) (GEERTS, 2012; SANEI; CHAMBERS., 2007). Um exemplo de uma página EEG usando uma montagem bipolar é mostrado na Figura 2.3. O comprimento da página é de 10 segundos, o que é normalmente usado ao analisar um registro.

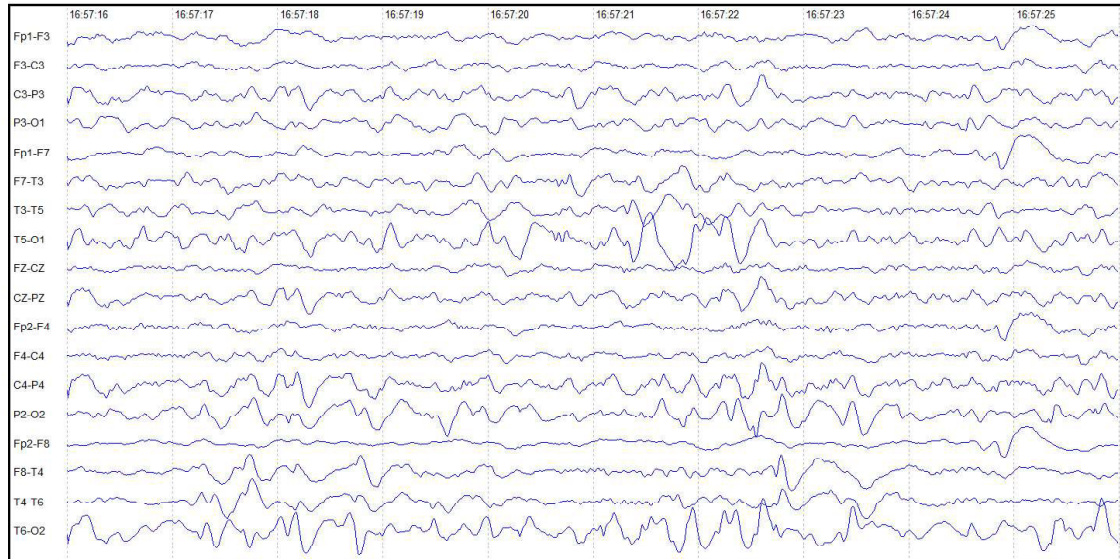


Figura 2.3: Montagem bipolar em uma página de 10 segundos

Alguns distúrbios cerebrais como a epilepsia, podem ser diagnosticados através de inspeção visual do sinal de EEG. Algumas características importantes, intrínsecas ao EEG, são comumente identificadas (caso presentes) pelos especialistas. Entre essas características destacam-se cinco grandes ondas cerebrais distintas (GEERTS, 2012; SANEI; CHAMBERS., 2007). Essas ondas são caracterizadas por suas diferentes faixas de frequência e são denominadas como alpha (α), theta (θ), beta (β), delta (δ) e gama (γ).

As ondas alpha e beta foram identificadas em 1929. Em 1938 foram identificadas ondas acima de 30 Hz e essas ondas ficaram conhecidas como ondas "gama". O ritmo delta foi introduzido por Walter (1964), essas ondas referem-se a todas as frequências abaixo do intervalo alfa. Walter (1964) também introduziu as ondas theta definidas como sendo aquelas que têm frequências dentro da faixa de 4-7,5 Hz (WALTER, 1964).

As ondas delta estão dentro do intervalo de 0.5 a 4 Hz. Essas ondas estão associadas principalmente ao sono profundo mas podem aparecer no estado de vigília. As ondas teta encontram-se no intervalo de 4 a 7.5 Hz e aparecem quando a consciência "desliza" em direção à sonolência (GEERTS, 2012; SANEI; CHAMBERS., 2007).

Ondas alfa aparecem na metade posterior da cabeça e geralmente são encontradas na região occipital do cérebro. Essas ondas podem ser detectadas em todas as partes dos lobos posteriores do cérebro. A frequência das ondas alphas geralmente encontra-se dentro do intervalo de 8 a 13 Hz, e essas ondas aparecem como um sinal arredondado ou sinusoidal. No entanto, em casos

raros, pode se manifestar como ondas agudas. Nesses casos, o componente negativo parece ser agudo e o componente positivo parece ser arredondado (GEERTS, 2012; SANEI; CHAMBERS., 2007).

Uma onda beta é observada no EEG variando dentro da faixa de 14 a 26 Hz. Essa onda representa o ritmo normal do cérebro em estado de vigília, associado ao pensamento ativo, atenção ativa, foco no mundo externo ou solução de problemas concretos, e é encontrado em adultos normais. As ondas gamas correspondem as com frequências acima de 30 Hz (principalmente até 45Hz). Essas ondas são, algumas vezes, chamadas de onda beta rápida (GEERTS, 2012; SANEI; CHAMBERS., 2007). A Figura 2.4 ilustra as principais ondas cerebrais.

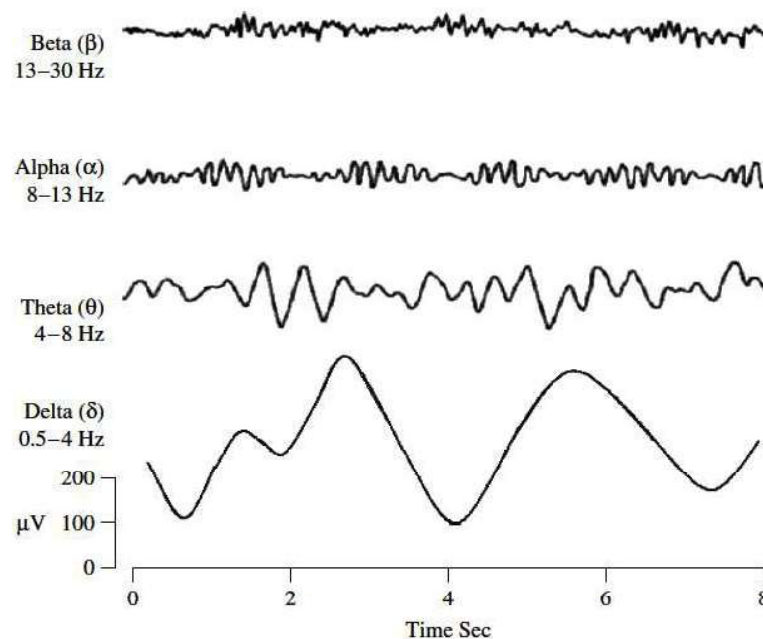


Figura 2.4: Quatro ritmos normais típicos do cérebro.(Figura adaptada de Sanei e Chambers. (2007))

2.3 Epilepsia

As definições e classificações de epilepsia sofreram algumas alterações, principalmente nas últimas décadas. Em 2005, uma força-tarefa da Liga Internacional Contra a Epilepsia (ILAE) formulou definições de crises epiléticas e epilepsia. A ILAE definiu crise epileptica como sendo uma ocorrência transitória de sinais e/ou sintomas devido à atividade neuronal anormal, excessiva ou síncrona no cérebro” (FISHER et al., 2005). Já epilepsia, foi definida como um distúrbio do cérebro caracterizado por uma predisposição duradoura para gerar convulsões epiléticas e

por consequências neurológicas, cognitivas, psicológicas e sociais (FISHER et al., 2005). Segundo Fisher et al. (2005) a definição de epilepsia requer a ocorrência de pelo menos uma crise epilética.

Em 2010 a ILAE publicou o relatório da Comissão de Terminologia da gestão 2005-2009, com a introdução de novos conceitos em epileptologia. A fim de tornar a definição mais operacional e coerente com a forma como os epileptologistas pensam sobre o assunto, a ILAE definiu então epilepsia como sendo uma doença do cérebro caracterizada por qualquer uma das seguintes condições (YACUBIAN; KOCHEN, 2014):

1. Pelo menos duas convulsões não provocadas (ou reflexas) ocorrendo em tempo maior que 24 horas.
2. Uma crise não provocada (ou uma crise reflexa) e chance de ocorrência de uma nova crise estimada em pelo menos 60 %.
3. Diagnóstico de uma síndrome de epilepsia (FISHER et al., 2005).

Apesar das pequenas mudanças aplicadas à definição de epilepsia, as classificações oficiais da ILAE ainda são as de 1981 e considera três grupos de crises: as crises parciais ou focais, as crises generalizadas e as crises não classificáveis (YACUBIAN; KOCHEN, 2014) Figura 2.5.



Figura 2.5: Classificação Internacional das Crises Epiléticas da ILAE de 1981 (YACUBIAN; KOCHEN, 2014)

Crises parciais (focais) são aquelas que se originam em uma rede neuronal limitada a um hemisfério e estão subdivididas em (YACUBIAN; KOCHEN, 2014):

1. Crises parciais simples
 - Com sinais motores.

- Com alucinações somatossensitivas ou sensoriais especiais.
- Com sinais e sintomas autonômicos

2. Crises parciais complexas

- Com início parcial simples seguido de alteração da consciência.
- Com comprometimento da consciência desde o início.

3. Crises parciais com generalização secundária

- Crises parciais simples que se generalizam.
- Crises parciais complexas que se generalizam.
- Crises parciais simples que evoluem para complexas e se generalizam.

As crises que se originam em um ponto e rapidamente envolvem redes neuronais distribuídas de forma bilateral são chamadas de crises epiléticas generalizadas (YACUBIAN; KOCHEN, 2014). Algumas crises parecerem localizadas quando analisadas individualmente, a lateralização não é consistente de uma crise para outra. Segundo Berg et al. (2010) as Crises generalizadas podem ser assimétricas. Hauser e Kurland (1975) realizaram um estudo epidemiológico no qual concluíram que mais da metade das pessoas com epilepsia estudadas apresentaram crises tônico-clônicas generalizadas.

As crises tônico-clônicas generalizadas são a forma mais comum dentre todas as outras crises epiléticas e são caracterizadas por perda abrupta da consciência, apneia, mordedura de língua e etc. As crises tônico-clônicas generalizadas podem se manifestar clinicamente em cinco fases: sinais e sintomas premonitórios; fase pré-ictal imediata; fase ictal; fase pós-ictal imediata e período de recuperação pós-ictal (YACUBIAN; KOCHEN, 2014).

- Fase de sintomas premonitórios: essa fase é caracterizada por alguns sintomas que podem anteceder as crises em horas ou até mesmo dias. Scaramelli et al. (2009) realizaram um estudo com 100 pacientes epiléticos a fim de analisar sintomas que antecedem crises epiléticas. Esses sintomas foram encontrados em 39 dos 100 pacientes avaliados. A Figura 2.6 apresenta os percentuais de ocorrência desses sintomas nos pacientes estudados.

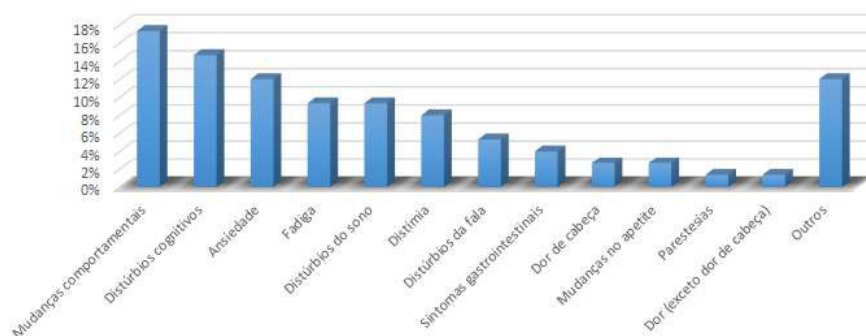


Figura 2.6: Frequência dos sintomas premonitórios, adaptado de (SCARAMELLI et al., 2009)

- Fase pré-ictal imediata: nessa fase das crises tônico-clônicas generalizadas (TCGs) ocorre um aumento da respiração e da frequência cardíaca, além de crises de ausência, surtos de polispícula no EEG e movimentos corporais;
- Fase ictal: uma diversidade de manifestações pode ser observada e algumas delas possuem variações mesmo durante o período ictal. Nesse contexto, é possível distinguir dois períodos característicos da fase ictal: o período tônico e o clônico. O período tônico tem duração de 10 a 20 segundos aproximadamente e caracteriza-se por contração tônica muscular axial, desvio ocular para cima e dilatação da pupila. Nesse período ocorre ainda uma duplicação da frequência cardíaca e da pressão arterial e apneia. Já na fase clônica ocorrem espasmos flexores e uma redução gradativa da frequência cardíaca e da pressão arterial, porém, a apneia permanece. Esse período tem duração aproximada de 40 segundos.
- Fase pós-ictal: após o período clônico ocorre um reestabelecimento da respiração iniciando, portanto, a fase pós-ictal. Nessa fase, além da normalização da respiração ocorre um cessar dos movimentos tônico-clônicos e um retorno gradual da consciência acompanhado de esquecimento da convulsão.
- Fase de recuperação: o término da crise, ou fase de recuperação pós-ictal, é marcado pelo coma, ou pelo sono pós-ictal ou despertar com confusão mental. Ou ainda podem ocorrer nessa fase o chamado comportamento automático acompanhado de cansaço, as dores musculares e cefaleia (Kotagal, 2000).

A Liga Internacional Contra a Epilepsia, através do grupo de trabalho em classificação e terminologia, publicou em setembro de 2006, um comunicado no qual contém uma lista das síndromes epiléticas e o correspondente período de vida no qual se inicia as crises (YACUBIAN; KOCHEN, 2014). Os períodos relacionados pela liga compreendem o neonatal, lactentes e a adolescência. Essa descrição pode ser observada na lista a seguir:

1. Período Neonatal

- Crises neonatais benignas familiares.
- Encefalopatia mioclônica precoce
- Síndrome de Ohtahara.

2. Lactentes

- Crises migratórias parciais do lactente.
- Síndrome de West
- Epilepsia mioclônica do lactente
- Crises infantis benignas
- Síndrome de Dravet
- Encefalopatia mioclônica em desordens não progressivas Infância

- Epilepsia occipital benigna da infância de início precoce (tipo Panayiotopoulos)
- Epilepsia com crises mioclono-astáticas.
- Epilepsia benigna da infância com descargas centrotemporais
- Epilepsia occipital da infância de início tardio tipo (tipo Gastaut)
- Epilepsia com ausências mioclônica
- Síndrome de Lennox-Gastaut
- Encefalopatia epiléptica com descargas de espícula-onda contínuas
- durante o sono incluindo a síndrome de Landau-Kleffner
- Epilepsia ausência da infância

3. Adolescência

- Epilepsia ausência juvenil
- Epilepsia mioclônica juvenil
- Epilepsia mioclônica progressivas

Alguns tipos de epilepsia não tem relação específica com a idade, dentre as quais podem ser citadas: epilepsia do lobo frontal noturna autossômica dominante, epilepsias do lobo temporal familiar, epilepsia do lobo temporal mesial com esclerose hipocampal, síndrome de Rasmussem e crises gelásticas com hamartoma hipotalâmico (YACUBIAN; KOCHEN, 2014).

2.4 Hipsarritmia

Gibbs e Gibbs (1952) utilizaram o vocábulo “hipsarritmia” para descrever um padrão eletroencefalográfico interictal. Esse padrão hipsarrítmico, normalmente é acompanhado por espasmos epilépticos infantis. Segundo Yacubian e Kochen (2014) os espasmos epilépticos consistem na contração de músculos axiais, ocasionando a flexão do tronco e abdução e elevação de ambos os membros superiores numa posição saudatória. Com um rápido movimento inicial, o paciente permanece na atitude saudatória por poucos segundos antes de alguns relaxamentos.

A tendência é que os espasmos epilépticos ocorram ao adormecer ou ao despertar. Em cerca de 75% dos casos os espasmos iniciam-se entre 3 e 7 meses de idade, porém podem ocorrer desde o nascimento até mais ou menos 5 anos (YACUBIAN; KOCHEN, 2014).

Os espasmos são característicos da síndrome de West. Esta síndrome tem como características físicas os espasmos e algumas características eletroencefalográficas. As características encefalográficas da síndrome de West são o que Gibbs e Gibbs (1952) definiram como “padrão de hipsarritmia” (HAGA et al., 1995).

2.4.1 Espículas e ondas agudas

A hipsarritmia consiste em um padrão eletroencefalográfico variável e dinâmico que, dentre outras características, possui uma grande quantidade de espículas e ondas agudas. Espículas são

marcadores eletrográficos cuja quantificação é utilizada no planejamento cirúrgico de pacientes com epilepsia. Um pico é clinicamente definido como um transitório agudo com duração entre 20 e 70 ms e claramente distinguível da atividade de fundo (IVES; GLOOR, 1977; CASPERS, 1983) (Figura 2.7).

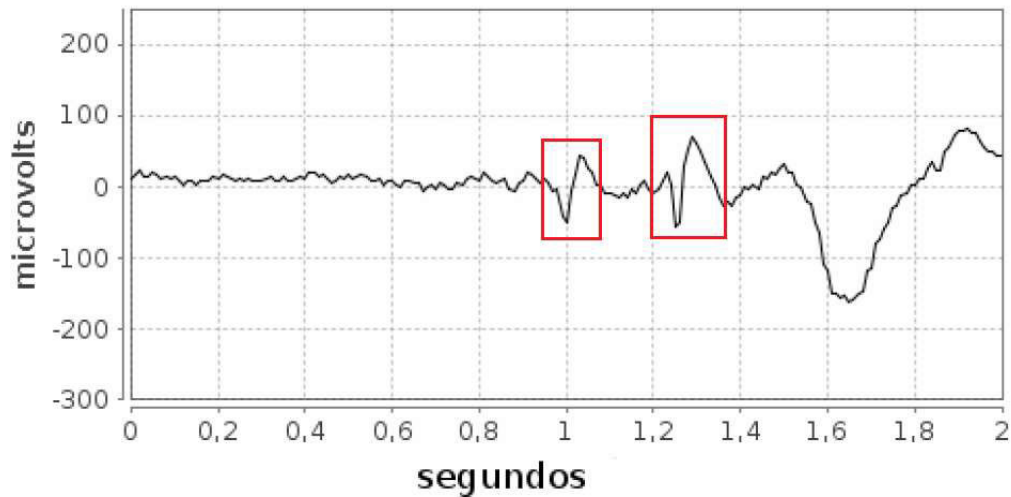


Figura 2.7: Presença de espículas no EEG (PACOLA et al., 2012)

Apesar da onda aguda ter forma semelhante a espícula, sua duração varia entre 70 a 200 milissegundos. Tanto espícula quanto ondas agudas também podem aparecer em registros com traçado normal, porém em menor quantidade (PACOLA et al., 2012) (Figura 2.8).

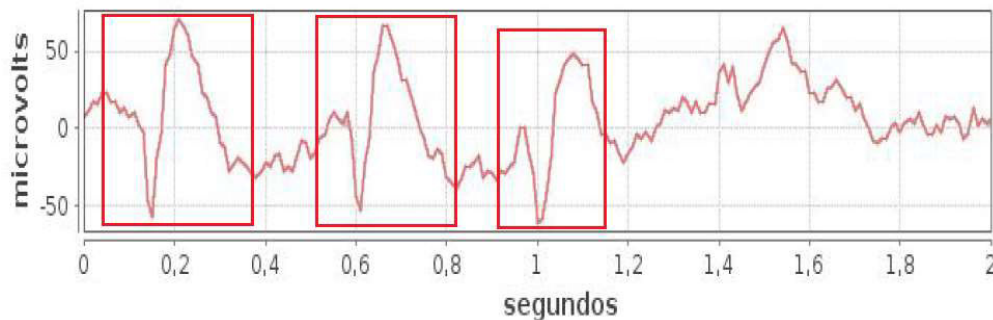


Figura 2.8: Presença de ondas agudas no EEG (PACOLA et al., 2012)

A amplitude das espículas é uma característica utilizada para classificação das mesmas. Janca et al. (2014) classifica as espículas, de acordo com amplitude, da seguinte forma: As espículas óbvias apresentaram amplitude acima de $300\mu V$ e as espículas ambíguas tem amplitude variando entre 190 a $210\mu V$ (JANCA et al., 2014) (Figura 2.9).



Figura 2.9: Classificação de espículas segundo a amplitude (adaptado de Janca et al. (2014))

Gastaut et al. (1965) ao investigarem as relações entre a esclerose tuberosa e hipsarritmia propuseram a seguinte classificação para hipsarritmia:

1. Hipsarritmia clássica (típica);
2. Hipsarritmia atípica (modificada):
 - hipsarritmia clássica fragmentada;
 - Hipsarritmia muito lenta;
 - Hiperarritmia muito rápida ;
 - Hipsarritmia assimétrica;
 - Hipsarritmia associada ao foco irritativo.

2.4.2 Variações do padrão de hipsarritmia

Hrachovy, Frost e Kellaway (1984) realizaram gravações prolongadas de EEG's e observaram que algumas características do padrão de hipsarritmia podem mudar drasticamente de minuto para minuto e de hora para hora. Os autores observaram ainda que tais características podem ser influenciadas por fatores como a duração de gravação do EEG, o estado clínico do paciente e a presença de anormalidades estruturais do cérebro.

Apesar dos EEG's observados por Hrachovy, Frost e Kellaway (1984) apontarem para aspectos que diferem da descrição inicial de Gibbs e Gibbs (1952), duas características fazem com que os mesmos sejam incluídos dentro da definição de hipsarritmia. A primeira consiste no fato de que os padrões listados por Hrachovy, Frost e Kellaway (1984) são compostos por assincronias e alta tensão. A segunda característica refere-se a associação aos mesmos tipos de crises clínicas e mesmos padrões de EEG ictal. Tendo em vista que os EEG's analisados e descritos por Hrachovy, Frost e Kellaway (1984) apresentaram diversas variações do conceito inicial de hipsarritmia, definiu-se então essas características como "hipsarritmia modificada".

Hrachovy, Frost e Kellaway (1984) classificaram os padrões encontrados em suas pesquisas da seguinte forma: hipsarritmia fragmentada, hipsarritmia com aumento da sincronização

interhemisférios, hipsarritmia assimétrica, hipsarritmia com anormalidade focal consistente, hipsarritmia com episódios de atenuação de tensão generalizada e hipsarritmia com atividade lenta de amplitude elevada, bilateral e assimétrica. A seguir são apresentados exemplos de EEGs com os diversos tipos de hipsarritmia.

- Hipsarritmia Fragmentada

O padrão de hipsarritmia fragmentada é observada principalmente durante a segundo estágio do sono. Nesse estágio o eletroencefalograma mostra um aumento nas pontas e uma tendência à sincronização, com uma fragmentação da hipsarritmia (YACUBIAN; KOCHEN, 2014).

As duas características do padrão de hipsarritmia fragmentada são a atenuação de trechos curtos do registro de EEG, e a interrupção em um curto intervalo de tempo da hipsarritmia. Esses dois fatores tem como origem o aumento da sincronia entre os hemisférios durante o sono (HRACHOVY; FROST; KELLAWAY, 1984).

Segundo Yacubian e Kochen (2014) esse padrão pode ser totalmente suprimido durante o sono REM fazendo com que o traçado se torne quase normal (HRACHOVY; FROST; KELLAWAY, 1984; YACUBIAN; KOCHEN, 2014). Um exemplo de EEG com esse tipo de traçado pode ser observado na (Figura 2.10).

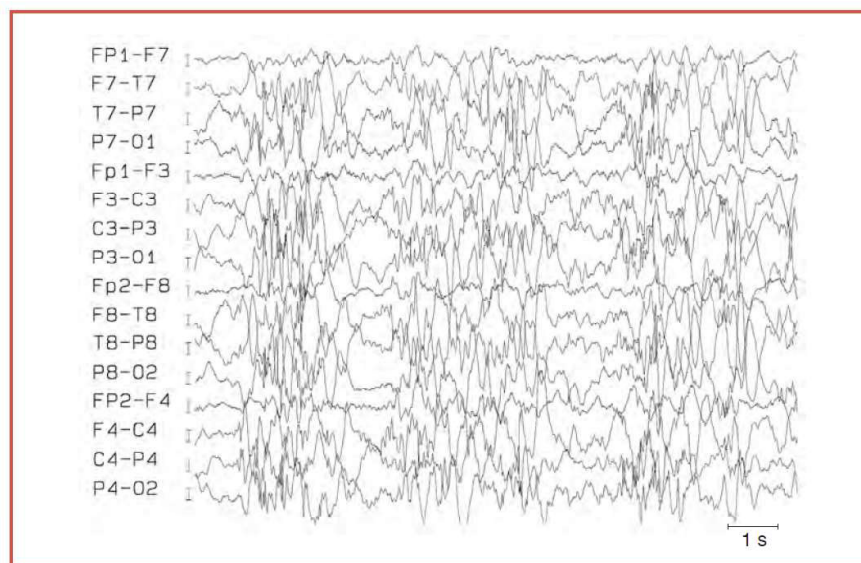


Figura 2.10: Registro de EEG com hipsarritmia fragmentada (YACUBIAN; KOCHEN, 2014)

- Hipsarritmia com aumento da sincronização inter-hemisférica

O tempo de vida de uma criança, cujo EEG tenha hipsarritmia, pode implicar em mudanças no eletroencefalograma da mesma. Padrões típicos descritos por Gibbs e Gibbs (1952), como por exemplo, as características multifocais e anárquicas são, em alguns casos, substituídos por organização, sincronização e simetria inter-hemisférica (HRACHOVY; FROST; KELLAWAY, 1984).

Segundo Hrachovy, Frost e Kellaway (1984), o aumento da sincronização e da simetria inter-hemisférica pode ser caracterizado como grande atividade generalizada de pico e onda ou aumento da sincronização da atividade de fundo.

A ocorrência da sincronização inter-hemisférica não é necessariamente uma característica constante. Em alguns momentos o EEG pode apresentar padrões de ondas que coincidem com a definição inicial de hipsarritmia. Porém, em outros momentos o eletroencefalograma exibe um aumento da sincronização (HRACHOVY; FROST; KELLAWAY, 1984).

Nessa variação de hipsarritmia é possível verificar a presença de dois padrões no traçado. A ocorrência do aumento da sincronização e da simetria evidenciados por surtos de atividades de espículas ou aumento da sincronização da atividade de base (HRACHOVY; FROST; KELLAWAY, 1984; YACUBIAN; KOCHEN, 2014). Na Figura 2.11 é apresentada uma montagem eletroencefalográfica com características de hipsarritmia com aumento da sincronização inter-hemisférica.

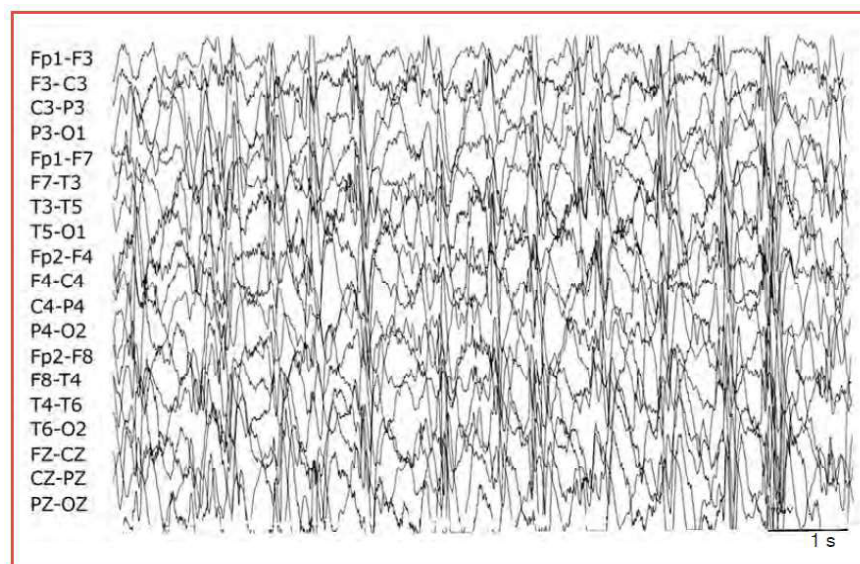


Figura 2.11: Hipsarritmia com aumento da sincronização inter-hemisférica (YACUBIAN; KOCHEN, 2014)

- Hipsarritmia assimétrica.

Hipsarritmia assimétrica consiste em um padrão em que há assimetria de valores de tensão entre os dois lados (YACUBIAN; KOCHEN, 2014). Estes valores de tensão variam e podem configurar uma assimetria regional ou uma hipsarritmia totalmente unilateral. Algumas patologias congênitas estão associadas a esse tipo de variação. Normalmente essas patologias causam algum tipo de deformação em um dos hemisférios (HRACHOVY; FROST; KELLAWAY, 1984). A Figura 2.12 ilustra essa variação de hipsarritmia.

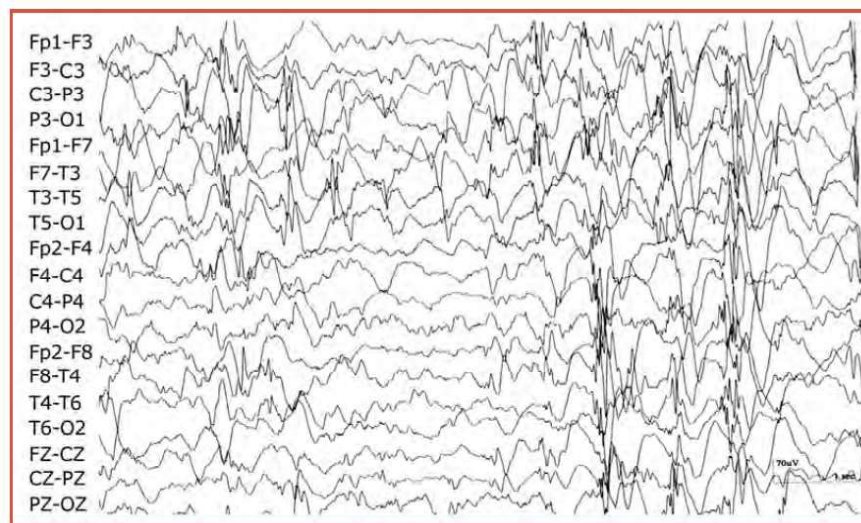


Figura 2.12: Hipsarritmia assimétrica (YACUBIAN; KOCHEN, 2014)

- Hipsarritmia com anormalidade focal consistente.

Hipsarritmia caracterizada por um campo de descarga de espícula ou onda aguda consistente, claramente distinguível das descargas multifocais. Essa variação de hipsarritmia pode ser acompanhada por crises parciais. As descargas focais podem ocorrer durante um conjunto de espasmos infantis (HRACHOVY; FROST; KELLAWAY, 1984; YACUBIAN; KOCHEN, 2014). A Fig. 2.13 ilustra esse tipo de variação.

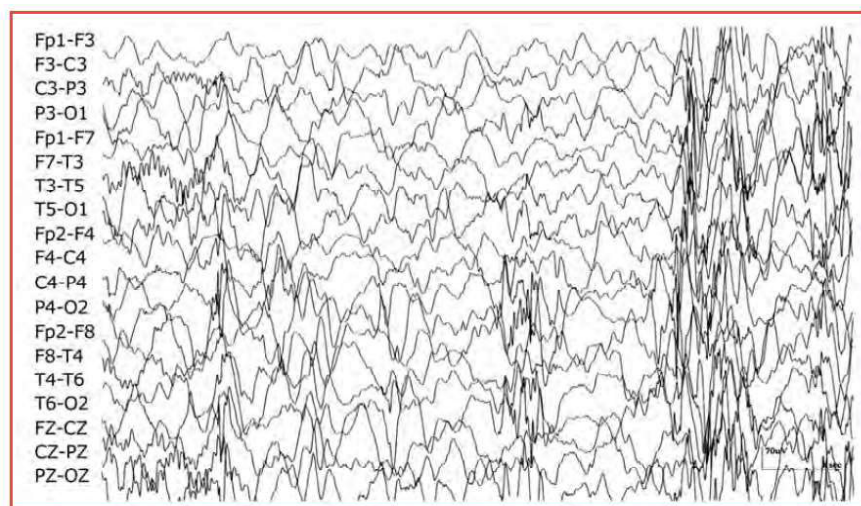


Figura 2.13: Hipsarritmia com anormalidade focal consistente (YACUBIAN; KOCHEN, 2014)

- Hipsarritmia com episódios de atenuação de tensão.

Atenuação de tensão durante um período de dois a dez segundos é a característica principal dessa variação de hipsarritmia (YACUBIAN; KOCHEN, 2014). Esses episódios de atenuação podem ser generalizados, localizados ou regionais e ocorrem de maneira semi-periódica a cada poucos segundos.

Os trabalhos de Hrachovy, Frost e Kellaway (1984) mostraram que os episódios de atenuação são mais evidentes durante o sono Sem Movimentos Oculares Rápidos (NREM) e, em alguns pacientes, podem ser vistos apenas naquele momento. Outra observação feita por Hrachovy, Frost e Kellaway (1984) é que durante um período de 24 horas, um paciente pode ter centenas de episódios de atenuação aleatórios sem serem acompanhados por espasmos clínicos.

Na Figura 2.14 é possível observar que em alguns intervalos de tempo a tensão registrada pelo exame eletroencefalográfico é atenuada, configurando assim um exemplo dessa variação de hipsarritmia.

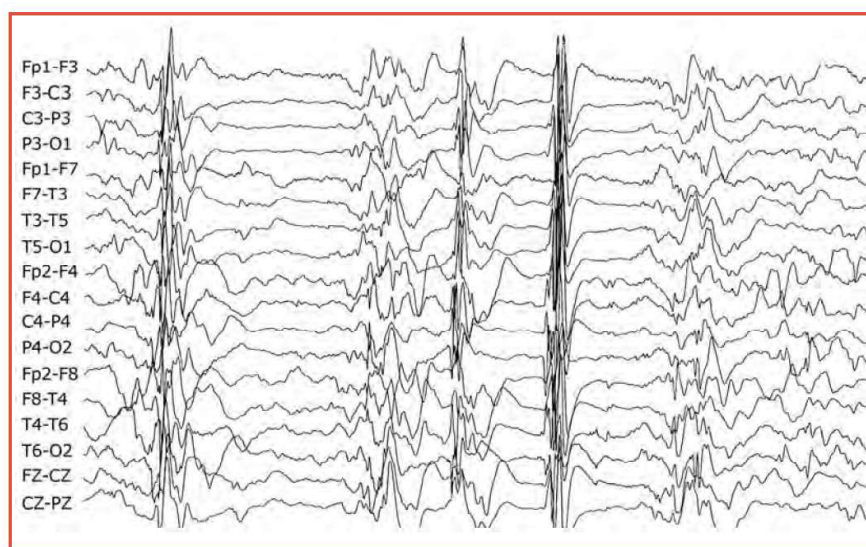


Figura 2.14: Hipsarritmia com episódios de atenuação de tensão (YACUBIAN; KOCHEN, 2014)

- Hipsarritmia com atividade lenta de amplitude elevada, bilateral e assíncrona.

Esse padrão hipsarrítmico é caracterizado por uma predominância de ondas lentas com amplitude elevada, com poucas descargas de espículas ou ondas agudas (YACUBIAN; KOCHEN, 2014). A Figura 2.15 é um exemplo dessa variação de hipsarritmia.

Na Figura 2.15 observa-se exame eletroencefalográfico de uma criança de 13 meses de idade. Nesse exame nota-se uma hipsarritmia modificada caracterizada pela presença de atividade lenta de alta tensão, assíncrona, com pouca atividade de pico ou de ondas agudas (HRACHOVY; FROST; KELLAWAY, 1984).

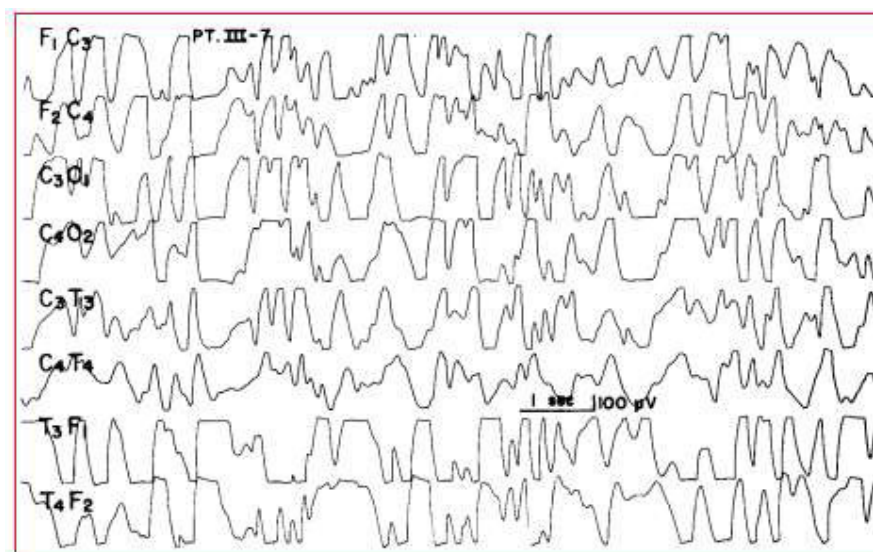


Figura 2.15: Hipsarritmia com atividade lenta de amplitude elevada, bilateral e assíncrona (HRACHOVY; FROST; KELLAWAY, 1984).

2.4.3 Hipsarritmia e a Síndrome congênita do Zika Vírus

O vírus zika é um arbovírus, pois é transmitido pela picada de um mosquito (nesse caso o mosquito fêmea do *Aedes*). O Zika foi descoberto pela primeira vez em Uganda em 1947. Posteriormente, em 1954 na Nigéria esse vírus foi detectado em seres humanos. O perfil clínico da infecção pelo vírus Zika consiste em uma doença febril semelhante à dengue e à chikungunya (GATHERER; KOHL, 2016).

No início do ano de 2015 ocorreu no Brasil um surto de infecções pelo vírus Zika. Nesse mesmo período um aumento na prevalência de bebês nascidos com microcefalia foi observado. A microcefalia é descrita pela Organização Mundial da Saúde como sendo uma redução na circunferência da cabeça (perímetro cefálico) com a medida occipitofrontal de um recém-nascido com 37 semanas de gestação igual ou inferior a 31,9 cm para meninos e 31,5 cm para meninas.

Segundo Russo, Jungmann e Braga (2017) a microcefalia é caracterizada por um perímetro craniano, ao nascer, de 2 desvios-padrões abaixo da média esperada para a idade gestacional e o sexo do bebê. Essa característica implica diretamente em prejuízos para o desenvolvimento cognitivo e comportamental de seus portadores.

Durante o surto de infecção pelo vírus Zika no Brasil verificou-se que o maior volume de casos de microcefalia relatados nesse período ocorreu na região Nordeste. Foi também na região Nordeste que a transmissão do vírus Zika foi relatada pela primeira vez no Brasil (OLIVEIRA; CORTEZ-ESCALANTE; OLIVEIRA, 2016).

Nesse contexto, percebeu-se então que o vírus Zika poderia ser transmitido de outras maneiras, além da maneira clássica via *Aedes*. Além da transmissão via *Aedes*, Russo, Jungmann e Braga (2017) recentemente descreveram outros tipos de transmissão do Zika, a saber, a transmissão vertical (materno/fetal) e sexual. A Figura 2.16 ilustra como ocorrem os três tipos de transmissão do vírus Zika.

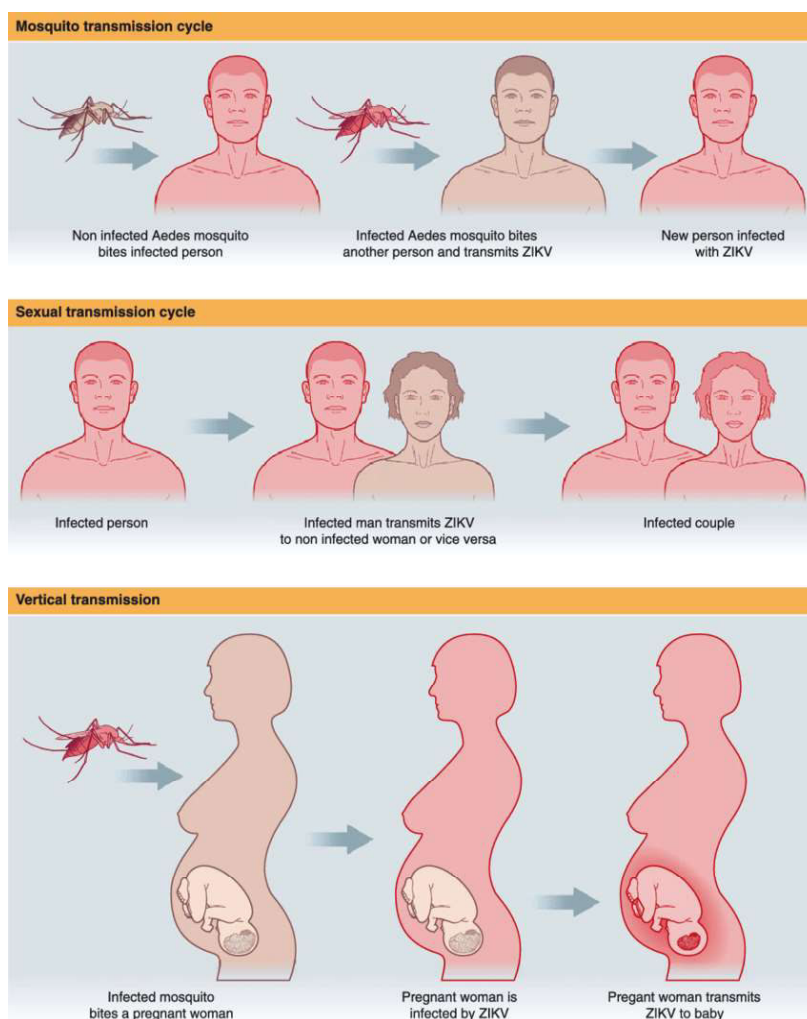


Figura 2.16: Ciclos de transmissão do vírus Zika (RUSSO; JUNGSMANN; BRAGA, 2017).

Oliveira, Cortez-Escalante e Oliveira (2016) analisaram 574 casos de microcefalia detectados por um sistema de vigilância. Os autores identificaram evidências que relacionavam sintomas consistentes com a infecção pelo vírus Zika durante o primeiro trimestre da gravidez e a prevalência de microcefalia no nascimento. Essa relação foi posteriormente confirmada em laboratório.

Desde 2016 até os dias atuais, diversos trabalhos com diferentes abordagens e centrados no tema que envolve microcefalia e o vírus Zika têm sido realizados. Pessoa et al. (2018) realizaram uma revisão de bibliografia e encontraram um total de 36 artigos correlatos com o tema. Nos trabalhos avaliados por Pessoa et al. (2018) verificou-se que entre os bebês com achados clínicos consistentes com a síndrome congênita do Zika, 54% apresentaram epilepsia e 100% apresentaram anormalidades motoras.

Carvalho et al. (2017) realizaram a primeira descrição de achados eletroencefalográficos em exames de crianças com microcefalia infectadas pelo vírus Zika. Em um conjunto de 37 EEG's o padrão de hipsarritmia foi encontrado em 11 exames, o equivalente a 29,72% dos eletroencefalogramas.

Os achados de Carvalho et al. (2017) sobre a relação tríplice entre vírus Zika, microcefalia

e hipsarritmia foram encontrados também por outros pesquisadores. Kanda et al. (2018), por exemplo, encontraram o padrão hipsarrítmico em 8 EEG's de crianças com microcefalia (de um total de 23). Alguns desses casos associados a infecção pelo vírus Zika.

Em um estudo envolvendo 141 bebês com infecção congênita pelo ZIKV, Linden et al. (2018) detectaram anormalidades estruturais no cérebro. Um percentual de 67% dessas crianças apresentaram epilepsia. Ao observar os exames eletroencefalográfico, os autores identificaram o padrão de hipsarritmia clássica ou modificada em 11% dos EEG's. Os casos de hipsarritmia identificados nesse trabalho foram acompanhados de espasmos clínicos.

Os trabalhos publicados desde 2016 evidenciam que casos de microcefalia estão relacionados com a infecção pelo vírus Zika. Essa relação tem como consequência o fato de que uma parcela considerável dessas crianças são acometidas por epilepsia. Em alguns desses casos de epilepsia são encontrados o padrão de hipsarritmia nos respectivos exames de eletroencefalografia.

2.5 Considerações

Este capítulo compõe uma parte do referencial teórico que trata dos conceitos referentes ao exame conhecido como eletroencefalograma e como esse está relacionado a identificação de patologias cerebrais como a epilepsia. Sobre o eletroencefalograma mostrou-se como é estabelecido o padrão de posições dos eletrodos no ato de realização do exame, destacando o padrão 10 - 20. Duas montagens do EEG (uma referencial e uma bipolar) foram exemplificadas.

O conceito de epilepsia foi abordado de forma a apresentar suas principais variações. Entre as variações de epilepsia destaque para aquelas que surgem na infância principalmente no período neonatal. Nesse capítulo o padrão de hipsarritmia foi descrito e suas diversas variações foram definidas e alguns exemplos apresentados.

Tendo em vista que um dos objetivos desse trabalho é aplicar métodos computacionais em EEG's de crianças com microcefalia infectadas pelo vírus Zika, foi reservado um subtópico no qual é discorrido sobre as realções existentes entre a hipsarritmia e a síndrome congênita do Zika.

Máquina de Vetores de Suporte - SVM

3.1 Introdução

As técnicas de Aprendizado de Máquina (AM) compõe uma área de pesquisa pertencente o campo da Inteligência Computacional. O AM estuda o desenvolvimento de mecanismos capazes de extrair conhecimento a partir de amostras de um conjunto dados. Entre os métodos de Aprendizado de máquina consta a Máquina de Vetores de Suporte.

A máquina de vetores de suporte - SVM é uma ferramenta muito popular dentro do contexto de execução de tarefas de aprendizado de máquina, principalmente quando a problemática envolve classificação de padrões ou regressão. O treinamento de uma máquina de vetores de suporte requer a solução de um problema de programação convexa (SWEILAM; THARWAT; MONIEM, 2010; RICHHARIYA; TANVEER, 2018).

As SVM's apresentam algumas características importantes que torna sua utilização mais corriqueira entre os pesquisadores sobre AM (LORENA; CARVAHO, 2003). Entre essas características pode-se destacar:

1. Boa generalização: A eficiência de classificação de novas amostras não pertencentes ao conjunto de treinamento é chamada de capacidade de generalização. Essa eficiência evita situações nas quais o classificador se torna muito especializado em um conjunto de treinamento e baixo desempenho quando testado em novos padrões (LORENA; CARVAHO, 2003).
2. Robustez em grandes dimensões (LORENA; CARVAHO, 2003).
3. Convexidade da função objetivo: O algoritmo das SVM são baseados em uma função objetivo quadrática, o que garante a presença de mínimo global (LORENA; CARVAHO, 2003).
4. Teoria bem definida: A base teórica das SVM's é construída sob um rigor matemático e estatístico bem estabelecido (LORENA; CARVAHO, 2003).

Três diferentes situações são abordadas nesse capítulo. A primeira é o caso em que os padrões que se deseja classificar são linearmente separáveis. Já a segunda situação é quando

esses padrões não são linearmente separáveis. Por último, é apresentada de maneira resumida, a formulação matemática de quando é necessária a utilização de funções kernel para realizar a classificação.

3.2 Classes Linearmente Separáveis

O método de aprendizagem computacional baseado em vetores de suporte é fundamentado, inicialmente, em se obter uma função capaz de separar duas classes. Esta função deve ter boa generalização, ou seja, deve ter um bom desempenho na separação de exemplos não vistos (LORENA; CARVAHO, 2003; TAKAHASHI, 2012; RUFINO, 2011; FILHO, 2010).

Um conjunto é dito linearmente separável se é possível obter pelo menos um hiperplano capaz de separar os padrões das classes diferentes contidos nesse conjunto (TAKAHASHI, 2012). Um hiperplano, segundo Boyd e Vandenberghe (2004), é o conjunto de soluções de uma equação linear não trivial que geometricamente pode ser interpretado como o conjunto de pontos que possui um produto interno constante com um determinado vetor w . Em símbolos, um hiperplano α é dado por:

$$\alpha : \{x | w^T x = b\}, \quad (3.1)$$

em que a constante real b determina o deslocamento do hiperplano a partir da origem.

É importante notar que o hiperplano α divide o R^n em dois semiespaços, $w^T x \geq b$ e $w^T x \leq b$ (BOYD; VANDENBERGHE, 2004). Essa noção de semiespaço é importante para a compreensão dos conceitos de “classes”. A Figura 3.1 ilustra de forma geométrica os semiespaços definidos a partir de um hiperplano.

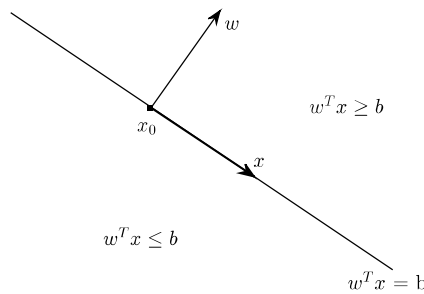


Figura 3.1: Hiperplano de separação e semiespaços. Adaptado de (BOYD; VANDENBERGHE, 2004)

Se para um dado conjunto de amostras existirem vários hiperplanos de separação, então é necessário encontrar aquele capaz de separar os padrões de tal forma que a distância entre o mesmo e o ponto mais próximo de uma classe (margem) seja maximizada (LORENA; CARVAHO, 2003; TAKAHASHI, 2012). Comparando a Figura 3.2 com 3.1, pode-se dizer que, para um problema de classificação binária, uma classe são todos os pontos x que satisfazem $w^T x \geq b$ ou $w^T x \leq b$.

Sejam $\omega = \{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_m, y_m)\}$ o conjunto de dados de treinamento, $\mathbf{x}_i \in R^n$ o vetor com as características de entrada, $y_i \in \{+1, -1\}$ a respectiva classe e δ a distância entre

α e os dados mais próximos, então a seguinte desigualdade é satisfeita (LORENA; CARVAHO, 2003; TAKAHASHI, 2012):

$$\frac{y_i[\mathbf{w}^T \mathbf{x}_i + b]}{\|\mathbf{w}\|} \geq \delta, \quad i = 1, 2, 3, \dots, m, \quad (3.2)$$

Reescalando os parâmetros $\mathbf{w}$ e b de tal forma que $\delta\|\mathbf{w}\| = 1$, a desigualdade (3.2) se reduz a representação do hiperplano de separação para uma amostra $\mathbf{x}_i$.

$$y_i[\mathbf{w}^T \mathbf{x}_i + b] \geq 1, \quad i = 1, 2, 3, \dots, m. \quad (3.3)$$

Como $\delta = \frac{1}{\|\mathbf{w}\|}$, é a menor distância entre α e os dados mais próximos então o hiperplano que separa os dados da melhor forma é aquele que maximiza

$$\Omega(\mathbf{w}) = \frac{1}{\|\mathbf{w}\|}. \quad (3.4)$$

3.2.1 Hiperplano Ótimo

Define-se como hiperplano ótimo aquele capaz de separar um conjunto de dados em classes distintas, tal que a distância entre o mesmo e o ponto mais próximo da margem seja máxima. A Figura 3.2 ilustra alguns exemplos de hiperplanos de separação e destaca o ótimo.

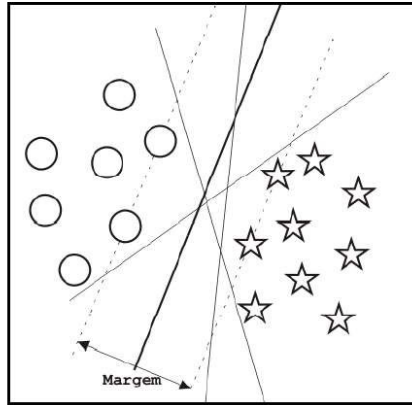


Figura 3.2: Hiperplano Ótimo (TAKAHASHI, 2012)

Em outras palavras, o hiperplano ótimo é a solução do seguinte problema de otimização com restrição (TAKAHASHI, 2012)

$$\min_{w,b} \quad \frac{1}{2} \|\mathbf{w}\|^2 \quad (3.5)$$

$$s.a \quad y_i[\mathbf{w}^T \mathbf{x}_i + b] \geq 1, \quad i = 1, 2, 3, \dots, m.$$

Esse é um problema de otimização que pode ser resolvido utilizando os Multiplicadores de Lagrange. Para solução de (3.5) transforma-se o problema com restrição em uma função sem

restrição equivalente ao problema original da seguinte forma (LORENA; CARVAHO, 2003; TAKAHASHI, 2012)

$$L_p(\mathbf{w}, b, \lambda) = \frac{1}{2} \|\mathbf{w}\|^2 - \sum_{i=1}^m \lambda_i (y_i [\mathbf{w}^T \mathbf{x}_i + b] - 1) \quad (3.6)$$

em que $\lambda = (\lambda_1, \lambda_1, \dots, \lambda_m)$ é o conjunto dos Multiplicadores de Lagrange dos padrões de treinamento, com $\lambda_i \geq 0$.

Como a função L_p deve ser minimizada em relação a $\mathbf{w}$ e b , e maximizada em relação a λ_i então o sinal negativo antes dos Multiplicadores de Lagrange é justificado.

Resolvendo as derivadas parciais em relação a $\mathbf{w}$ e b

$$\frac{\partial}{\partial b} L_p(\mathbf{w}, b, \lambda) = 0 \quad (3.7)$$

e

$$\frac{\partial}{\partial \mathbf{w}} L_p(\mathbf{w}, b, \lambda) = 0, \quad (3.8)$$

tem-se

$$\sum_{i=1}^m \lambda_i y_i = 0, \quad \lambda_i \geq 0 \quad i = 1, 2, 3, \dots, m \quad (3.9)$$

$$\mathbf{w} = \sum_{i=1}^m \lambda_i y_i \mathbf{x}_i, \quad \lambda_i \geq 0 \quad i = 1, 2, 3, \dots, m. \quad (3.10)$$

Substituindo (3.9) e (3.10) em (3.6), obtém-se o seguinte problema de otimização com restrição chamado problema dual (LORENA; CARVAHO, 2003)

$$L_D(\lambda) = \sum_{i=1}^m \lambda_i - \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m \lambda_i \lambda_j y_i y_j \mathbf{x}_i^T \mathbf{x}_j \quad (3.11)$$

$$s.a \quad \lambda_i \geq 0, \quad e \quad \sum_{i=1}^m \lambda_i y_i = 0 \quad i = 1, 2, 3, \dots, m.$$

É importante notar que tanto L_P quanto L_D são oriundas da mesma função objetivo, porém L_P deve ser minimizada em relação a $\mathbf{w}$ e b , enquanto que L_D deve ser maximizada em relação a λ_i (LORENA; CARVAHO, 2003; TAKAHASHI, 2012). Após encontrados os λ_i , é possível determinar $\mathbf{w}$ substituindo o valor de λ_i em (3.10).

O valor do parâmetro b é determinado utilizando a condição de vetor de suporte. Um vetor $(\mathbf{x}_s, \mathbf{y}_s)$ é um vetor de suporte se satisfaz

$$y_s [\mathbf{w}^T \mathbf{x}_s + b] = 1. \quad (3.12)$$

Substituindo (3.10) em (3.13) e isolando b

$$b = y_s - \sum_{i=1}^m \lambda_i y_i [\mathbf{x}_i^T \mathbf{x}_s]. \quad (3.13)$$

Com os valores de $\mathbf{w}$ e b determinados, o classificador é treinado e deverá ser capaz de prever a classe para novos padrões, a partir da seguinte função (LORENA; CARVAHO, 2003; TAKAHASHI, 2012)

$$classe(\mathbf{x}_s) = \begin{cases} +1, & \text{se } \mathbf{w}^T \mathbf{x}_s + b > 0 \\ -1, & \text{se } \mathbf{w}^T \mathbf{x}_s + b < 0 \end{cases} \quad (3.14)$$

Na prática, é comum se obter dados para classificação que não são linearmente separáveis, e portanto a função 3.14 não é a melhor opção de classificação. Para essas situações também é possível obter um hiperplano de separação (LORENA; CARVAHO, 2003; TAKAHASHI, 2012).

3.3 Classes Não Linearmente Separáveis

Um conjunto é dito não linearmente separável quando pelo menos um $\mathbf{x}_i$ não satisfaz a condição (3.15)

$$y_i[\mathbf{w}^T \mathbf{x}_i + b] \geq 1 \quad (3.15)$$

A violação da condição 3.15 pode acontecer em três situações diferentes (TAKAHASHI, 2012):

1. A dupla $(\mathbf{x}_i, y_i)$ se encontrar dentro da região de separação e no lado correto da superfície de decisão (Figura 3.3). Quando isso ocorre, o hiperplano foi escolhido de maneira incorreta.

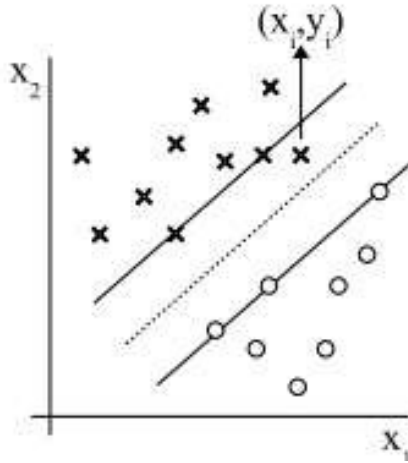


Figura 3.3: Primeiro caso de violação da condição 3.15 (adaptado de Takahashi (2012))

2. A dupla $(\mathbf{x}_i, y_i)$ pertence a região de separação, porém encontra-se no lado incorreto do hiperplano de separação (Figura 3.4).

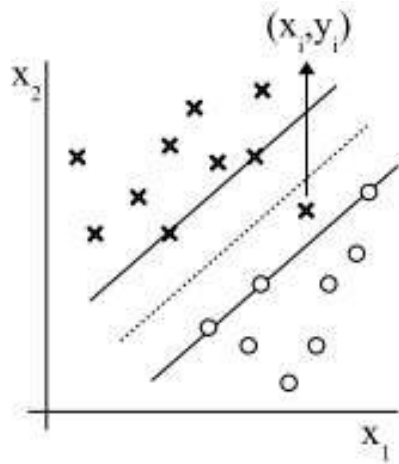


Figura 3.4: Segundo caso de violação da condição 3.15 (adaptado de (LORENA; CARVAHO, 2003; TAKAHASHI, 2012))

3. O terceiro caso de violação da condição 3.15 ocorre quando a dupla $(\mathbf{x}_i, y_i)$ encontra-se fora da região de separação e no lado incorreto da superfície de decisão (Figura 3.5).

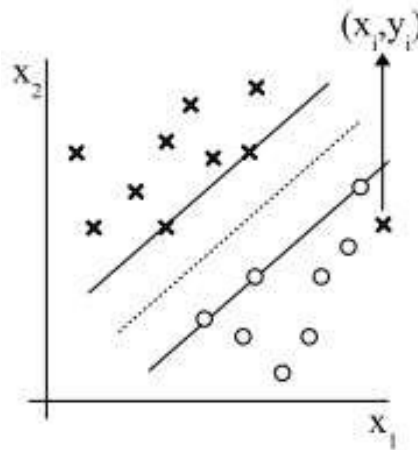


Figura 3.5: Terceiro caso de violação da condição 3.15 (adaptado de (LORENA; CARVAHO, 2003; TAKAHASHI, 2012))

Com o objetivo de se estabelecer uma solução que atenda de maneira ótima as três violações mencionadas, intruduz-se no problema uma variável $\{\xi_i\}_{1 \leq i \leq n}$ de tal forma que a condição 3.15 se torne

$$y_i[\mathbf{w}^T \mathbf{x}_i + b] \geq 1 - \xi_i, \quad (3.16)$$

em que as variáveis ξ_i são denominada variáveis soltas (LORENA; CARVAHO, 2003; TAKAHASHI, 2012).

É importante notar que, se a dupla $(\mathbf{x}_i, y_i)$ se encaixar no caso 1 de violação então $0 \leq \xi_i \leq 1$. Porém, se $(\mathbf{x}_i, y_i)$ violar a condição 3.15 segundo os casos 2 e 3 então $\xi_i > 1$ (Figura 3.6).

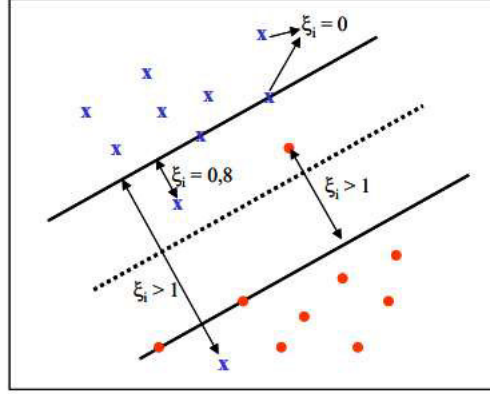


Figura 3.6: Exemplo de variáveis soltas (LORENA; CARVAHO, 2003; TAKAHASHI, 2012)

A utilização de variáveis soltas ξ_i permite estabelecer duas sentenças verdadeiras, a primeira é que sempre existirá a possibilidade de ocorrência de erros no processo de classificação e a outra é que os erros ocorrerão sempre que $\xi_i > 1$. Com isso, pode se estabelecer então que a soma das variáveis soltas ξ_i é um limite para o número de erros na fase de treinamento e esse limite deve ser minimizado. Portanto, a função desse problema passa a ser

$$\min_{w,b,\xi} \quad \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^n \xi_i, \quad (3.17)$$

C é um termo que representa um peso a minimização dos erros no conjunto de treinamento em relação a complexidade do modelo.

Diante do exposto, defini-se o problema de otimização primal para encontrar o hiperplano de separação ótimo para classes não linearmente separáveis

$$\min_{w,b,\xi} \quad \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^n \xi_i, \quad (3.18)$$

$$s.a \quad \xi_i \geq 0, \quad e \quad y_i[\mathbf{w}^T \mathbf{x}_i + b] \geq 1 - \xi_i \quad i = 1, 2, 3, \dots, n.$$

Resolvendo o problema de maneira análoga ao que foi feito no caso de separação linear, chega-se ao problema dual

$$\max \quad \sum_{i=1}^n \lambda_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j y_i y_j \mathbf{x}_i^T \mathbf{x}_j \quad (3.19)$$

$$s.a \quad \sum_{i=1}^n y_i \lambda_i \quad e \quad 0 \leq \lambda_i \leq C \quad \forall i = 1, 2, 3, \dots, n,$$

em que C é um parâmetro especificado pelo usuário.

A principal diferença entre os problemas duais para conjuntos linearmente e não linearmente separáveis é que no caso de não linearidade na separação a condição $\lambda_i \leq 0$ é substituída por $0 \leq \lambda_i \leq C$.

3.3.1 Função Kernel

Mesmo utilizando variáveis soltas ξ_i , em muitos casos, não é possível separar satisfatoriamente os padrões do conjunto de treinamento através de um hiperplano. Para resolver esse problema é realizado um mapeamento $\Psi : S \rightarrow \Omega$, em que S é o conjunto de treinamento e Ω é chamado de espaço das características. O espaço das características, corresponde a uma representação de S mapeado por funções não lineares φ (TAKAHASHI, 2012) (Figura 3.7).

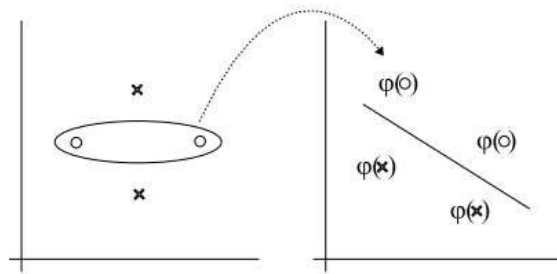


Figura 3.7: Mapeamento através de uma função kernel (adaptado de (TAKAHASHI, 2012))

Um kernel é uma função $k : S \rightarrow \Omega$, que é representado da seguinte forma:

$$k(x_i, x_j) = \varphi^T(\mathbf{x}_i)\varphi(\mathbf{x}_j) \quad (3.20)$$

Sem considerar de forma explícita o espaço de característica, a técnica de SVM consiste em utilizar o kernel para construir um hiperplano ótimo no espaço das características de tal forma que o $\mathbf{w}$ ótimo possa ser escrito da forma

$$\mathbf{w} = \sum_i^n \lambda_i y_i \varphi^T(\mathbf{x}_i) \varphi(\mathbf{x}_j). \quad (3.21)$$

Os kernels mais utilizados são os polinomiais, os gaussianos e os sigmoidais. Os parâmetros desses kernels são especificados pelo usuário (ver Tabela 3.1).

Tabela 3.1: Principais Kernels

Kernel	Função	Parâmetros
Polinomial	$[a(x_i x_j + k)]^d$	a, k e d
Gaussiano ou RBF	$e^{-\sigma x_i - x_j ^2}$	σ
Sigmoidal	$\tanh[a(x_i x_j + k)]$	a e k

Os Kernels apresentados na Tabela 3.1 são capazes de realizar mapeamentos de dados não linearmente separáveis, por se tratarem de funções não lineares. Essas funções transformam

o espaço original em um espaço de característica no qual a separabilidade dos dados é mais evidente.

3.4 Considerações

O método de classificação conhecido como Máquina de Vetores de Suporte SVM, foi apresentado nesse capítulo. Inicialmente foram apresentadas as principais características das SVM's, que tornam essa técnica atrativa.

Três situações importantes foram apresentadas. A primeira referente a problemas cujas as amostras são linearmente separáveis. Para esse caso discorreu-se sobre a formulação matemática de como se obter um hiperplano de separação ótimo. A segunda situação refere-se ao caso em que os dados não são linearmente separáveis mas ainda sim é possível estabelecer matematicamente, considerando uma margem de erro, um hiperplano de separação.

Por fim, foi apresentado o caso em que é necessário uma mudança de espaço através de um mapeamento através de funções conhecidas como kernel. Esse mapeamento é realizado a fim de projetar os dados em um novo espaço no qual se estabelece um hiperplano de separação.

Essa técnica (SVM) será utilizada para classificação de Eletroencefalogramas com e sem hirsarritmia e ainda para classificar trechos do EEG que apresentem o padrão hirsarrítmico e trechos com descargas multifocais. As motivações para escolha das SVM's foram: a sua capacidade de generalização, as características de separação dos valores de entrada e o fato dessa tecnica já ter sido utilizada nesse tipo de problema, como pode ser observado no trabalho de Smith et al. (2017).

Decomposição de Sinais em Pequenas Ondas

4.1 Introdução

As descobertas sobre como aproximar uma função por uma combinação linear de pequenas ondas tem seu ápice nos trabalhos de Jean Baptiste Joseph Fourier por volta de 1807. Fourier observou que qualquer função periódica, de período 2π e quadrado integrável, poderia ser representada por uma série de senoidais harmônicas (MORETTIN, 2014).

A análise de Fourier ainda é muito atual no processamento de sinais, porém alguns questionamentos relevantes foram levantados durante o tempo. O principal questionamento é: Os resultados de Fourier podem ser expandidos para funções não periódicas?

Na perspectiva de encontrar bases que gerem todo o espaço de funções reais de quadrado integrável a partir de uma única função ψ é que surge a análise em pequenas ondas (Wavelets). Além de não estarem limitadas apenas a construção de funções periódicas, as wavelets são ondas bem localizadas no tempo e na frequência enquanto que as bases de Fourier são localizadas apenas em frequência (MORETTIN, 2014).

Além de serem bem localizadas no tempo e na frequência a decomposição em wavelets necessita de poucos coeficientes para representar determinado sinal. Esses coeficientes assim com na análise de Fourier, revelam a regularidade do sinal através de suas amplitudes (STEPHANE, 1999).

Alguns marcos na história do desenvolvimento de pequenas ondas merecem destaque. O primeiro grande feito foi quando Haar em 1910 construiu uma função constante por partes cujas dilatações e translações geram uma base ortonormal para o espaço das funções de quadrado integrável (HAAR, 1910). Posteriormente, Sjölin e Strömberg (1983) encontraram um ψ que também gera uma base ortonormal e oferece melhores aproximações de suavidade.

Meyer (1992) construiu uma família de wavelets ortonormais, com funções infinitamente contínuas e diferenciáveis. Esse fato impulsionou buscas por novas formas ortonormais. A sistematização matemática para bases de wavelets ortonormais foi estabelecida por Meyer (1992) através da elaboração de aproximações de sinais.

Em estudos mais recentes as ondas conhecidas como funções de Gabor também tem sido usadas em processamento de sinais, principalmente para realização de filtragem, decomposição de sinais e codificação (LUCENA et al., 2011; CORREIA, 2018; CORREIA et al., 2017).

Neste capítulo são apresentados alguns conceitos relacionados a técnicas de processamento de sinais envolvendo Wavelets e funções de Gabor. Merece destaque nesse capítulo a metodologia de decomposição de sinais em funções de Gabor sintonizadas.

4.2 Wavelets

Tendo em vista que a decomposição de sinais em Wavelets surge do questionamento sobre a possibilidade de se escrever qualquer função, de quadrado integrável, como combinação linear de ondas localizadas no tempo e na frequência, verifica-se a necessidade de introduzir o conceito de espaço L_2 (espaço das funções de quadrado integráveis), pois é sobre esse espaço vetorial que as Wavelets estão inseridas. Após a definição de L_2 uma breve apresentação das Wavelets é realizada.

Definição 1. Segundo Bachman e Narici (1966), L_2 é o espaço de todas as funções $f(t)$ quadraticamente integráveis, ou seja,

$$\int_{-\infty}^{\infty} |f(t)|^2 dt < \infty. \quad (4.1)$$

Definição 2. O espaço L_2 pode ser gerado por uma única função $\psi_{a,b}(t)$ através de dilatação (ou compressão) e translação, tal que:

$$\psi_{a,b}(t) = |a|^{-1/2} \psi\left(\frac{t-b}{a}\right) \quad (4.2)$$

em que a e b números reais, $a \neq 0$ Morettin (2014), Stéphane (2009b).

A função $\psi_{a,b}(t)$ é conhecida como wavelet mãe. Algumas wavelets formam uma base ortonormal para L_2 e podem ser simétricas ou não. Para ser considerada Wavelet uma função deve satisfazer as propriedades P_1 a P_4 (MORETTIN, 2014; STÉPHANE, 2009b).

- $P_1) \int_{-\infty}^{\infty} \psi_{a,b}(t) dt = 0.$
- $P_2) \int_{-\infty}^{\infty} |\psi_{a,b}(t)| dt < \infty.$
- $P_3) \int_{-\infty}^{\infty} \frac{|\psi_{a,b}(\omega)|^2}{|\omega|} d\omega < \infty$, em que $\psi_{a,b}(\omega)$ é a transformada de Fourier de $\psi_{a,b}(t)$
- $P_4) \int_{-\infty}^{\infty} t^j \psi_{a,b}(t) dt = 0, j = 0, 1, 2, \dots, r-1$, para algum $r \geq 1$ e $\int_{-\infty}^{\infty} t^r \psi_{a,b}(t) dt < \infty.$

Segundo Morettin (2014) e Stéphane (2009b) o valor de r está ligado ao grau de suavidade de ψ , ou seja, quanto maior r mais suave será a onda.

Um exemplo de pequena onda bastante utilizado em processamento de sinais é a wavelet de Haar. Essa onda tem como função de escala a expressão:

$$\psi_{j,k}(t) = \begin{cases} 1, & 0 \leq t \leq 1 \\ 2^{j/2}, & 2^{-j}k \leq t < 2^{-j}(k + 1/2) \\ -2^{j/2}, & 2^{-j}(k + 1/2) \leq t < 2^{-j}(k + 1) \\ 0, & c.c. \end{cases} \quad (4.3)$$

A Figura 4.1 é composta por dois exemplos gráficos de Wavelets de Haar.

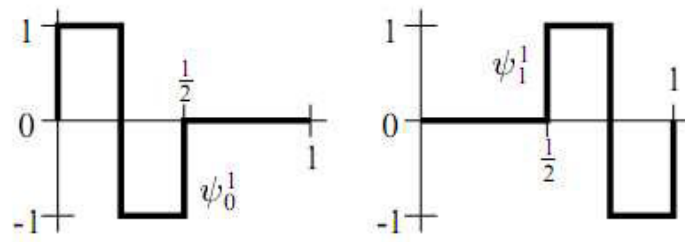


Figura 4.1: Exemplos de Wavelet de Haar (MONTEIRO, 2019)

Diversas propostas e ou descobertas de novas ondas com propriedades de wavelets têm surgido no decorrer do tempo. Porém, nem todas essas ondas apresentam expressão analítica e nem geram sistemas ortogonormais nem ortogonais.

Devido a grande quantidade de Wavelets existente na literatura e o fato de nem todas apresentarem expressões analíticas, esse trabalho limita-se a apresentar de maneira gráfica e ilustrativa apenas alguns exemplos de wavelets (Figura 4.2).

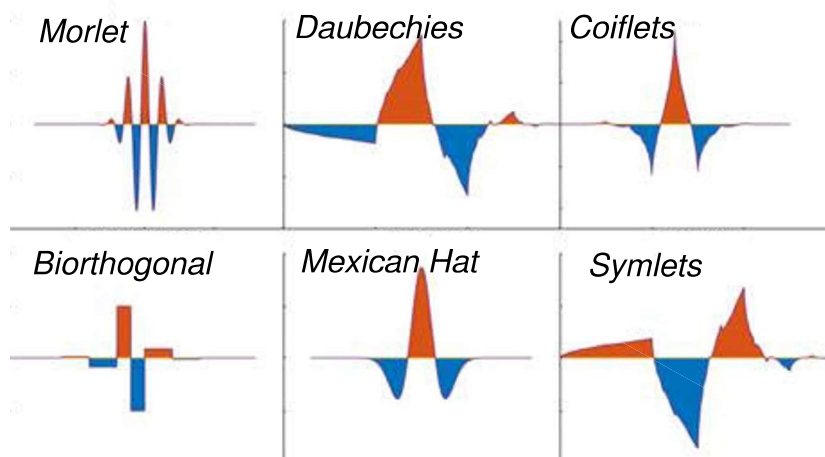


Figura 4.2: Tipos de wavelets (adaptado de Stéphane (2009a))

As pequenas ondas ilustradas na Figura 4.2, juntamente com a wavelet de Haar são, segundo Morettin (2014), as mais utilizadas para descrever funções de energia finita. Uma das vantagens

de se decompor um sinal em wavelets é a possibilidade de analisá-lo no tempo e na frequência concomitantemente. Além de poder visualizar as pequenas ondas no domínio do tempo e da frequência ao mesmo instante ainda é possível verificar graficamente sua energia.

A propagação de energia de um átomo de wavelet, corresponde a uma caixa, conhecida como caixa de Heisenberg (Figura 4.3). Nesse gráfico é possível verificar a localização em tempo x frequência de duas wavelets.

Na Figura 4.3 a área do retângulo (caixa) é constante. Isso significa que diminuir o tempo de propagação implica em aumentar o suporte de frequência e deslocá-lo para frequências mais altas (STÉPHANE, 2009a).

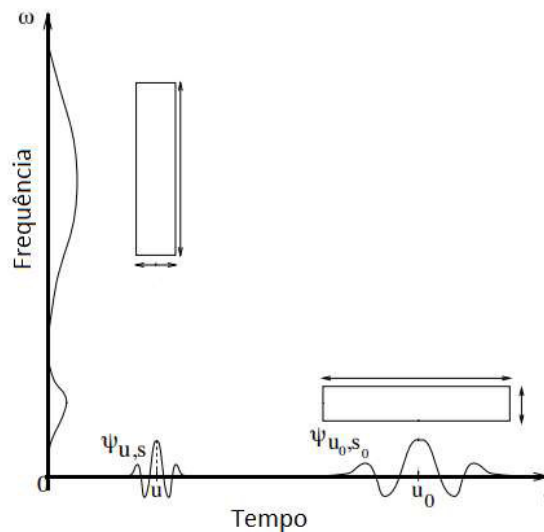


Figura 4.3: Caixas de Heisenberg de duas wavelets (adaptado de Stéphane (2009a))

A característica de ajustar a frequência de acordo com o tempo de propagação que as wavelets possuem faz das mesmas uma ferramenta importante quando utilizada como transformada. No tópico a seguir é realizada uma abordagem conceitual das Transformadas de Wavelet.

4.3 Transformadas de Wavelet

Existem na literatura diversas técnicas (transformadas) para mudança de domínio de uma função. Entre essas técnicas a Transformada de Fourier se destaca, pois permite observar uma função, que originalmente depende do tempo, no domínio da frequência. É conhecido que a Transformada de Fourier de uma função $f(t)$ pertencente ao L_2 é dada por:

$$\hat{f}(\omega) = \int_{-\infty}^{\infty} f(t)e^{-i\omega t} dt, \quad \omega \in R. \quad (4.4)$$

A princípio, observando a Equação 4.4, conclui-se que para se obter informações sobre dada frequência é necessário calcular a integral sobre todo domínio de tempo (BACHMAN; NARICI, 1966). Essa característica em algumas situações é considerada uma “deficiência” da Transformada de Fourier. A fim de amenizar esse problema Gabor (1946) propôs uma transformada

que utiliza janelas localizadas no tempo. A transformada proposta por Gabor (1946) tem a seguinte forma:

$$Gf(\omega) = \int_{-\infty}^{\infty} f(t)g_{\alpha}(t-b)e^{-i\omega t}dt, \quad (4.5)$$

em que

$$g_{\alpha}(t) = \frac{1}{\alpha\sqrt{2}}e^{-t^2/4\alpha}. \quad (4.6)$$

Apesar da Transformada de Gabor ser bem localizada no tempo e na frequência, dado um valor de α , o tamanho da janela não se modifica. Nessas condições verifica-se a necessidade de uma transformada que seja capaz de aumentar o tamanho da janela em baixas frequências e diminuir nas altas (BACHMAN; NARICI, 1966). Essa flexibilidade porém, como ilustrado na caixa de Heisenberg, é uma característica das wavelets, ou seja, uma transformada cujo núcleo seja uma wavelet mãe resolveria o problema de flexibilidade. É nesse contexto então, que surge a Transformada de Wavelet.

Além da flexibilidade mencionada, as Wavelets em sua maioria geram todo o espaço L_2 e uma operação importante em L_2 é o produto interno entre duas funções. É com base no produto interno que a Transformada de Wavelet está definida. A definição de produto interno no L_2 e, por conseguinte, a definição de Transformada de Wavelet estão apresentadas a seguir:

Definição 3. *Seja $f(t)$ e $g(t)$ pertencentes a L_2 , defini-se o produto interno entre f e g da seguinte maneira (BACHMAN; NARICI, 1966)*

$$\langle f, g \rangle \triangleq \int_{-\infty}^{\infty} f(t)g(t)dt. \quad (4.7)$$

Definição 4. *A Transformada de Wavelet contínua é o produto interno entre uma wavelet $\psi_{a,b}(t)$ e a função $f(t)$ (STÉPHANE, 2009a)*

$$CWT \triangleq \int_{-\infty}^{\infty} f(t)\psi_{a,b}(t)dt. \quad (4.8)$$

A Transformada de Wavelet pode ser reescrita como uma convolução entre $f(t)$ e $\psi_{a,b}(t)$

$$CWT \triangleq \int_{-\infty}^{\infty} f(t)|a|^{-1/2}\psi\left(\frac{t-b}{a}\right)dt = f * \psi_{a,b}(t). \quad (4.9)$$

Para o caso discreto, define-se a Transformada Discreta de Wavelet (do inglês *Discrete Wavelet Transform* - DWT) de uma série X_n da seguinte forma (MORETTIN, 2014):

$$DWT_{j,k} \triangleq \sum_{n=0}^N X_n\psi_{j,k}(n). \quad (4.10)$$

Essa transformada retorna um conjunto de coeficientes de tamanho igual a quantidade de observações. É possível verificar que a Equação 4.10 pode ser escrita na forma matricial. Outra observação interessante é que as Transformadas de Wavelet (contínua e discreta) atuam sobre o sinal como um banco de filtros.

4.4 Funções de Gabor

Diversos trabalhos na área de processamento de sinais tem utilizado funções de Gabor como centro de suas metodologias. Lucena et al. (2011) codificaram sinais de frequência cardíaca e encontraram como código neural do coração um conjunto de funções de Gabor. Por outro lado, Correia (2018) aplicou a mesma técnica e observaram que sinais de eletroencefalograma também podem ser codificados através de funções de Gabor.

Aplicando uma técnica de codificação em imagens de cenas naturais Bell e Sejnowski (1997) verificaram que tais imagens tem como base as funções de Gabor. A definição a seguir refere-se a representação genérica das funções de Gabor.

Definição 5. Uma função de Gabor consiste em uma senoide (ou cossenoide) modulada por uma função gaussiana (MARČELJA, 1980; YANG, 1992):

$$\varphi(t) = Ae^{-\frac{(t-\bar{t})^2}{2\sigma^2}} \cos[(t - \bar{t})\omega + \phi], \quad (4.11)$$

$$\bar{t}, \sigma^2, \omega \text{ e } \phi \in \mathbb{R}.$$

Os parâmetros $\bar{t}$, σ , ω e ϕ são respectivamente, a média e o desvio padrão do envelope gaussiano, a frequência e a fase do cosseno. Variando esses parâmetros é possível se obter uma infinidade de ondas. A Figura 4.4 ilustra algumas variações das funções de Gabor quando os parâmetros $\bar{t}$ e σ são modificados.

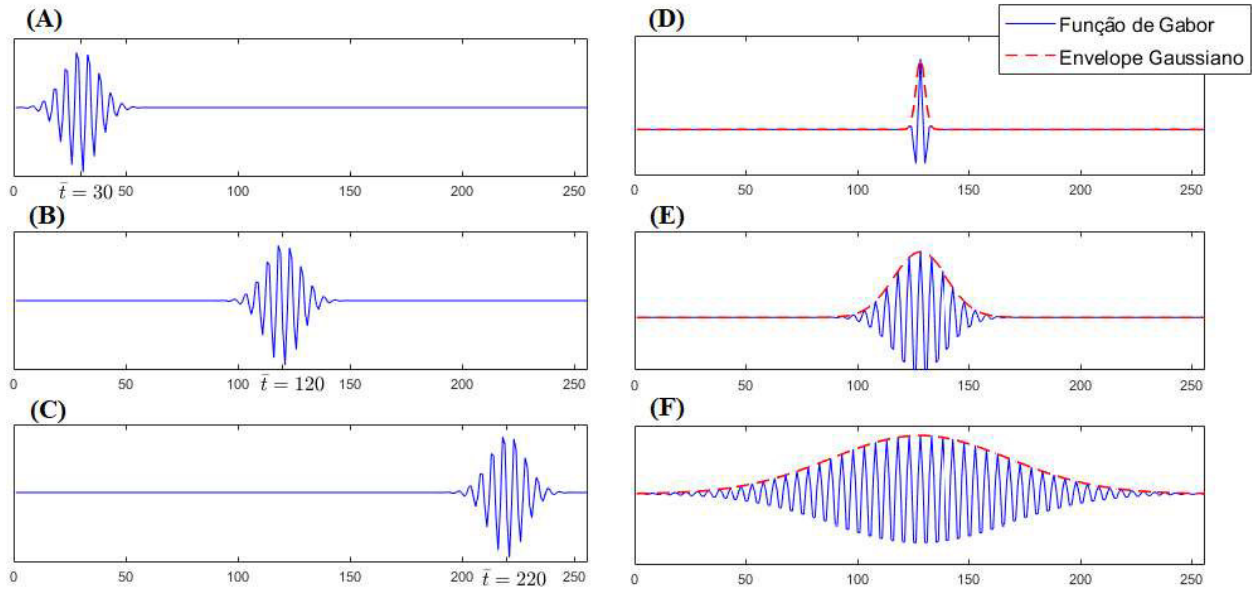


Figura 4.4: Exemplos de funções de Gabor. (A),(B) e (C) deslocamento do parâmetro $\bar{t}$ no tempo; (D),(E) e (F) impacto da mudança do parâmetro σ .

Variar os parâmetros das funções de Gabor implica em mudanças interessantes na localização e formato das ondas, uma breve descrição dessas mudanças estão listadas a seguir:

- O valor de $\bar{t}$ define a localização no tempo, desse modo, variando-o é possível percorrer todo o sinal;
- O valor de ω determina a frequência central da função de Gabor. A modificação de tal frequência pode mensurar e localizar frequências do sinal;
- O desvio padrão do envelope gaussiano σ , estabelece o tamanho da janela a ser analisada. Sua alteração determina a duração entre o início e o final da variação da função.
- A fase ϕ , pertencente aos argumentos da função cosseno, estipula a forma de onda interna ao envelope gaussiano.

As funções de Gabor quando analisadas a partir das caixas de Heisenberg possuem a característica de otimalidade da área, ou seja, o produto entre a variação em frequência e a variação no tempo é mínimo.

4.5 Funções de Gabor Sintonizadas

As funções de Gabor definidas anteriormente e que tem sua forma analítica descrita através da Equação (4.11), gozam de propriedades interessantes, o que faz das mesmas ferramentas importantes às tarefas relativas a processamento de sinais. Porém, a melhor escolha dos parâmetros $\bar{t}$, σ^2 , ω e ϕ é um desafio a ser superado.

Uma opção de escolha dos parâmetros das funções de Gabor é “sintonizá-las” ao sinal analisado. Essa técnica consiste em considerar as seguintes igualdades (TORREÃO, 2017):

$$\sigma^2 = |\tilde{I}(\omega)| \quad (4.12)$$

$$\phi = \gamma(\omega) \quad (4.13)$$

em que $|\tilde{I}(\omega)|$ é o módulo da transformada de Fourier do sinal $I(t)$ e $\gamma(\omega)$ a respectiva fase. Nessas condições define-se as funções de Gabor sintonizadas da seguinte forma:

Definição 6. *Uma função de Gabor sintonizada é toda função do tipo descrito pela Equação (4.11) com $\sigma^2 = |\tilde{I}(\omega)|$ e $\phi = \gamma(\omega)$, tal que:*

$$\varphi(t) = \frac{1}{(2\pi)^{\frac{3}{2}}} e^{-\frac{(t-\bar{t})^2}{2|\tilde{I}(\omega)|^2}} \cos[\omega t + \gamma(\omega)], \quad \bar{t} \in \mathbb{R}. \quad (4.14)$$

Segundo Oppenheim (2010) um sinal $I(t)$ pode ser escrito, utilizando a Transformada Inversa de Fourier, da seguinte forma

$$I(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{I}(\omega) e^{i\omega t} d\omega. \quad (4.15)$$

Como $\tilde{I}(\omega) = |\tilde{I}(\omega)| e^{i\gamma(\omega)}$, a Equação 4.15 torna-se

$$I(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\tilde{I}(\omega)| e^{i[\omega t + \gamma(\omega)]} d\omega. \quad (4.16)$$

Combinando a Equação (4.12) com a função de densidade gaussiana, a seguinte igualdade torna-se verdadeira

$$|\tilde{I}(\omega)| = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{\frac{(t-\bar{t})^2}{2|\tilde{I}(\omega)|^2}} dt, \quad (4.17)$$

e consequentemente,

$$I(t) = \frac{1}{(2\pi)^{\frac{3}{2}}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{\frac{(t-\bar{t})^2}{2|\tilde{I}(\omega)|^2}} e^{i[\omega t + \gamma(\omega)]} d\omega dt, \quad (4.18)$$

ou seja, as funções de Gabor sintonizadas permitem, no domínio real, a representação exata de $I(t)$.

$$I(t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{(2\pi)^{\frac{3}{2}}} e^{-\frac{(t-\bar{t})^2}{2|\tilde{I}(\omega)|^2}} \cos[\omega t + \gamma(\omega)] d\omega dt, \quad (4.19)$$

No entanto, do domínio discreto, a representação dada por integral pode ser substituída por uma aproximação somatória e a Equação 4.19 torna-se:

$$I(t) \approx \sum_{i=1}^n \varphi_i(t, \omega) \Delta\omega_i \Delta t_i. \quad (4.20)$$

A Equação (4.20) indica que todo sinal pode ser escrito como combinação linear de funções de Gabor sintonizadas, cujo os coeficientes dessa combinação refletem as variações em tempo e frequência de cada função de Gabor. Essas variações em tempo e frequência são equivalentes a energia de um átomo representado na caixa de Heisenberg.

É importante notar que, sintonizar os parâmetros das funções de Gabor ao sinal através da Transformada de Fourier, resolve o problema de três dos quatro parâmetros necessários para construir uma função de Gabor. Porém, o valor que determina a posição temporal $\bar{t}$ não é determinado nessa técnica.

Como $\bar{t}$ é um parâmetro que está relacionado ao posicionamento temporal da onda, conjecturou-se nesse trabalho o seguinte: Se $\bar{t}$ for modificado de tal forma que seja encontrado um ponto de similaridade máxima (segundo uma métrica dada) entre a função de Gabor sintonizada e o sinal, então esse ponto poderia ser definido como um valor razoável para $\bar{t}$.

A conjectura considerada para determinação de $\bar{t}$ obteve embasamento matemático e computacional quando observado o trabalho de Marçelja (1980) e o algoritmo Matching Pursuit - M P. No trabalho desenvolvido Marçelja (1980) é provado que todo sinal pode ser escrito como combinação linear de funções de Gabor, ou seja, dado um sinal $I(t)$ e um conjunto de n funções de Gabor $\varphi(t)$ então:

$$I(t) \approx \sum_{i=1}^n a_i \varphi_i(t, \bar{t}), \quad a_i = \langle I(t), \varphi(t, \bar{t}) \rangle \quad (4.21)$$

Considerando a Equação (4.21) pode-se inferir sobre a métrica de similaridade que determina $\bar{t}$, afirmando ser esta o produto interno $\langle I(t), \varphi(t, \bar{t}) \rangle$, ou seja, localiza-se $\bar{t}$ deslocando $\varphi(t)$ sobre $I(t)$ até o produto interno ser máximo. Encontrando o momento em que o produto interno é máximo localiza-se o valor de $\bar{t}$.

A metodologia de busca por similaridade entre funções através do produto interno já é bastante conhecida na literatura. Um dos algoritmos que realiza essa tarefa é conhecido como Matching Pursuit-MP. Na subseção a seguir é apresentada uma descrição do funcionamento desse algoritmo.

4.5.1 O Algoritmo Matching Pursuit-MP

A busca por similaridade através do produto interno é uma tarefa realizada pelo algoritmo Matching Pursuit-MP. Esse algoritmo permite, em alguns casos, decompor um sinal na soma ponderada de outros sinais, assim como descrito na Equação (4.21).

O algoritmo MP faz uso de um dicionário de funções Φ , no qual faz uma busca e “elege” a função com maior similaridade com o sinal, em seguida calcula-se um resíduo que consiste na diferença entre o sinal e a função eleita. Esse processo é repetido até todas as funções do dicionário serem utilizadas e ou o resíduo alcançar valor definido pelo usuário. Em resumo, a sequência de passos do MP pode ser observada no algoritmo descrito abaixo.

Algoritmo Matching Pursuit

- Step 1: Inicialize $n = 1$ e defina $R^0 I(t) = I(t)$ como resíduo.
- Step 2: Defina o dicionário Φ composto por $\varphi_j(t, \bar{t})$, definindo $\bar{t} = 0$.
- Step 3: Compute $|\langle R^{n-1} I(t), \varphi_j(t, \bar{t}_k) \rangle|$ para todas $\varphi_j(t, \bar{t}_k) \in \Phi$ e $j = 0, 1, 2, \dots, m$.
- Step 4: encontre $\varphi_n^* = \arg \max_{\varphi_j(t, \bar{t}), k} |\langle R^{n-1} I(t, \bar{t}_k), \varphi_j(t, \bar{t}_k) \rangle|$.
- Step 5: Compute o novo resíduo,

$$R^n I(t) = I(t) - \sum_{i=1}^n \langle R^{i-1} I(t), \varphi_i^*(t) \rangle \varphi_i^*(t).$$
- Step 6: Remova φ_n^* de Φ .
- Step 7: Se $n = m$ ou algum critério de parada é atingido, **PARE!**;
 caso contrário defina $n = n + 1$ e vá para Step 3.

Após remover todas as φ_n^* do dicionário Φ , ou então o critério de parada for atingido, pode-se estabelecer a seguinte aproximação:

$$I(t) \cong \sum_{i=1}^n a_i \varphi_i(t, \bar{t}), \quad (4.22)$$

A combinação entre as funções de Gabor sintonizadas e o algoritmo MP permite então, escrever uma função (sinal) a partir de uma base de funções que trazem consigo informações da função original, como por exemplo, frequência e módulo da Transformada de Fourier.

4.5.2 Decomposição em Funções de Gabor Sintonizadas

Considerando todos os conceitos abordados nesse tópico nota-se a possibilidade de se estabelecer uma metodologia capaz de transformar uma função em uma combinação linear de funções de Gabor sintonizadas. Essa metodologia permite que se obtenha uma $n - \text{upla}$ de coeficientes a_i que representem similaridades no tempo e na frequência entre a n -ésima Gabor e um conjunto de pontos centrados em $\bar{t}$.

Para realizar a tarefa de decomposição de um sinal em Gabor sintonizadas, primeiramente realiza-se a Transformada Rápida de Fourier (FFT) deste sinal a fim de extrair um conjunto de parâmetros que serão utilizados para construir as funções. Em seguida define-se um dicionário de funções de Gabor cujos parâmetros foram obtidos através da FFT (fixando $\bar{t} = 0$).

O dicionário obtido através da sintonização do sinal às funções de Gabor é então utilizado como entrada para algoritmo Matching Pursuit e finalmente o sinal pode ser escrito como combinação linear de funções de Gabor sintonizadas. A sequência de passos para realização de tal decomposição está descrito na Figura 4.5.

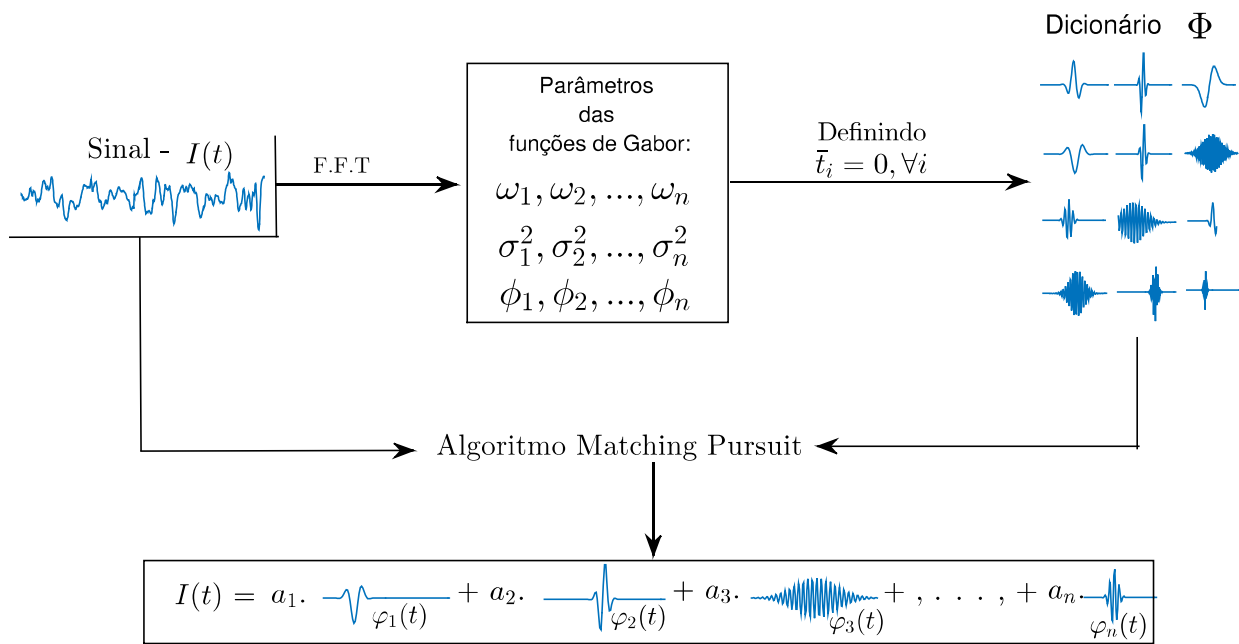


Figura 4.5: Decomposição em funções de Gabor sintonizadas

A decomposição proposta nesse subtópico é uma metodologia que permite realizar diversas análises. Essa técnica permite, por exemplo, realizar uma representação esparsa, pois o número de coeficientes a_i pode ser menor do que o número de pontos do sinal original.

Métricas clássicas de análise de sinais como medidas estatísticas, entropia e etc, podem ser aplicadas a representação obtida a partir da decomposição proposta.

4.6 Considerações

Neste capítulo foram apresentados conceitos importantes para compreensão do desenvolvimento desse trabalho. Algumas informações matemáticas sobre Wavelets e as transformadas (contínua e discreta) de Wavelete também foram discutidas. Foi apresentada a maneira que se pode observar um átomo de Wavelet no gráfico de tempo-frequência através das caixas de Heisenberg. Conceitos e exemplos de funções de Gabor foram percorridos nesse capítulo, destacando a maneira como os parâmetros influenciam na morfologia desse tipo de pequena onda.

No último tópico desse capítulo foram desenvolvidos temas que compõem a proposta de decomposição de sinais em funções de Gabor sintonizadas. Esses temas passaram pela definição de

Gabor sintonizadas, pela construção do dicionário utilizado pelo algoritmo Matching Pursuit.

É importante salientar que posteriormente esses procedimentos serão utilizados para definição matemática dos índices utilizados para identificação de hipsarritmia em sinais de eletroencefalograma de crianças com microcefalia infectadas pelo vírus Zika.

Materiais e Métodos para Identificação de Hipsarritmia

5.1 Introdução

Nesse capítulo é apresentada uma visão geral de todo o processo de realização dessa pesquisa. Uma descrição das fases e etapas é apresentada no decorrer do texto. Propõe-se também uma metodologia matemática e computacional capaz de auxiliar na identificação do padrão de hipsarritmia com base na definição de três índices matemáticos. Esses índices são baseados em técnicas de processamento de sinais que utilizam pequenas ondas com Wavelets, Funções de Gabor e Gabor sintonizadas.

O primeiro índice é uma aplicação da *Discrete Wavelet Transform* - DWT. A aplicação da DWT resulta em um conjunto de coeficientes capazes de representar o sinal em espaço de pequenas ondas. A definição do segundo índice será feita também com base na utilização da DWT. Porém nesse caso, a wavelet mãe é substituída por uma função de Gabor e os coeficientes resultantes da transformação definem o índice.

O terceiro índice é definido com base na decomposição de sinais em funções de Gabor sintonizadas. Primeiramente decompõe-se o sinal em uma combinação linear de pequenas ondas e os coeficientes da transformação linear são usados para cálculo do índice.

Dois grupos distintos de experimentos são analisadas através do cálculo dos índices. Nos primeiros experimentos o cálculo de cada índice é aplicado a EEG's com traçado normal e sinais com padrão hipsarrítmico. A média e a log energia das séries de índices servem de entrada para uma SVM para classificação dos EEG's. Em seguida serão realizados alguns testes estatísticos com a finalidade de verificar se os índices são capazes de realçar diferenças significativas entre os sinais avaliados.

No segundo grupo foram escolhidas amostras de sinais e um processo de identificação dos trechos dos sinais com hipsarritmia é realizado. Uma especialista em eletroencefalografia identificou três características nos EEG's com padrão de hipsarritmia, a saber, descargas multifocais, hipsarritmia e sem alterações.

Nos dois grupos de experimentos o processo de classificação é avaliado utilizando a técnica de validação cruzada e as métricas de sensibilidade, especificidade e acurácias serão verificadas.

5.2 Delineamento da Pesquisa

Para Zanella (2006) uma pesquisa científica pode ser classificada de acordo com seus objetivos, quanto à sua abordagem e quanto aos procedimentos adotados na coleta de dados. Os objetivos de uma pesquisa podem determinar se a mesma é do tipo exploratória, explicativa ou descritiva. Com relação à abordagem, uma pesquisa pode ser classificada como sendo qualitativa ou quantitativa. No que diz respeito aos procedimentos adotados na coleta de dados pode ser considerada do tipo bibliográfica, documental, experimental ou ex-post-facto (“a partir do fato passado”).

Considerando a classificação definida por Zanella (2006) pode-se afirmar que esta pesquisa é do tipo explicativa, quantitativa e ex-post-facto. Considera-se este trabalho como ex-post-facto, pois o autor não participou da coleta de dados, ou seja, toda pesquisa se deu após a coleta já ter sido realizada (após o fato).

Essa pesquisa foi realizada em quatro fases. Primeiramente uma revisão bibliográfica foi realizada a fim de se tomar conhecimento sobre o estado da arte. Em seguida foi executada uma pesquisa exploratória sobre os principais temas abordados em todo o desenvolvimento desse trabalho. As duas fases finais consistiram nas formulações matemáticas e computacionais e nas análises dos resultados. A Figura 5.1 resume todo processo metodológico utilizado na execução dessa pesquisa.

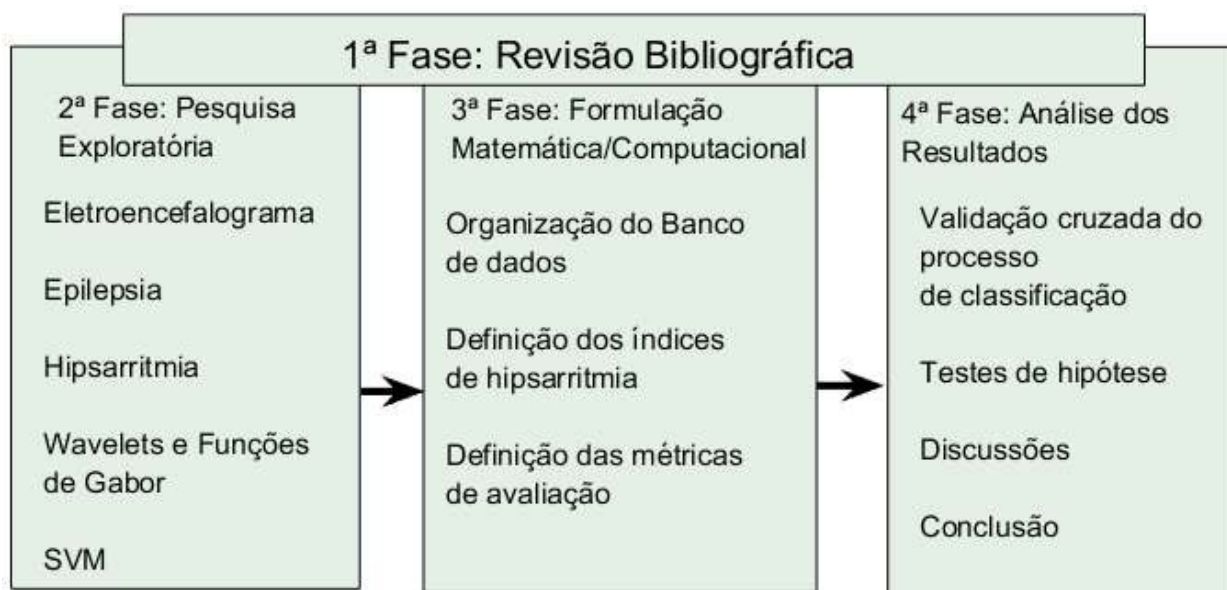


Figura 5.1: Delineamento da pesquisa.

A primeira fase dessa pesquisa constituiu-se de um levantamento bibliográfico sobre o tema proposto. Essa fase porém, foi além do “ponta pé” inicial, pois em todo o desenvolvimento desse trabalho realizou-se pesquisas bibliográficas a fim de manter-se uma sintonia entre as atividades de pesquisa e as atualizações sobre tema. Já na segunda fase foi realizada uma pesquisa exploratória na qual os principais assuntos que compõe o referencial teórico foram estudados da maneira mais aprofundada possível.

A terceira fase de realização dessa tese foi composta por quatro etapas: organização do banco de dados, definição dos índices de hipsarritmia, descrição das métricas de avaliação e os testes computacionais.

5.2.1 Organização da Base de dados

A base de dados utilizada é composta por EEGs de 30 crianças com Síndrome Congênita do Vírus Zika (SCVZ). Estes sinais foram obtidos junto a instituição governamental "Casa Ninar". Esta instituição realiza atendimentos relacionados a saúde de crianças com microcefalia (NINAR, 2018). Uma breve descrição de algumas características das crianças, cujo EEG foi utilizado nesse trabalho, é apresentada na Tabela 5.1.

Tabela 5.1: Características das crianças que tiveram os EEG's utilizados nas simulações: PC- Perímetro Cefálico; M- Masculino; F- Feminino; SD- Sem Dados; SL- Sem Acesso ao Laudo.

Código	Sexo	PC-Nascimento	PC - 1 ^a consulta	Idade	Presença de Hipsarritmia
A1	F	31 cm	36,5 cm	8 meses	Não
A2	F	29 cm	30 cm	13 dias	Não
A3	M	SD	40 cm	1 ano	Não
A4	F	29,5 cm	32,5 cm	3 meses	Sim
A5	F	28 cm	34 cm	5 meses	Sim
A6	M	SD	32 cm	5 meses	Sim
A7	M	SD	27,5 cm	5 meses	Sim
A8	F	26 cm	30 cm	3 meses	Sim
A9	M	32,5 cm	30,5 cm	5 meses	Sim
A10	F	26 cm	30,5 cm	2 meses	Sim
A11	M	28 cm	28 cm	13 dias	Não
A12	F	26,5 cm	28 cm	14 dias	Não
A13	M	31,5 cm	35 cm	2 meses	Sim
A14	M	SD	35 cm	3 meses	Sim
A15	M	SD	31 cm	4 meses	Sim
A16	F	31 cm	35 cm	3 meses	Sim
A17	M	28 cm	28,5 cm	17 dias	Não
A18	M	32 cm	35 cm	3 meses	Sim
A19	F	28 cm	33,5 cm	8 meses	Não
A20	F	29,5 cm	35,5 cm	8 meses	Não
A21	F	29,5 cm	32 cm	2 meses	Não
A22	M	SD	38,5 cm	9 meses	Não
A23	F	28,5 cm	36,5 cm	1 ano	Não
A24	M	33 cm	37 cm	3 meses	Não
A25	M	29 cm	34 cm	1 ano	Não
A26	M	29 cm	33 cm	4 meses	Sim
A27	F	SD	39,5 cm	9 meses	Não
A28	F	SL	SL	SL	Não
A29	M	SL	SL	SL	Não
A30	F	SL	SL	SL	Não

As crianças atendidas na Casa Ninar apresentam alterações eletroencefalográficas variadas e, em alguns casos, crises epiléticas. Clinicamente, as crises se caracterizam por espasmos infantis associados a presença de hipsarritmia.

Os EEG's utilizados formam em um total de 203, com duração de 5 minutos cada (98 com características de hipsarritmia e 105 sem alterações). Esses registros eletroencefalográficos foram coletados com eletrodos posicionados segundo o padrão 10-20 e adquiridos em montagem bipolar durante períodos de sono espontâneo. O software utilizado para coleta foi o Neuromap e os sinais foram amostrados a 128 Hz. O consentimento para uso dos dados foi obtido junto ao comitê de ética da Universidade Federal do Maranhão CAAE 65897317.1.0000.5086.

A presença ou ausência de hipsarritmia nos EEGs foi identificada por uma especialista. Nos EEG's identificados com hipsarritmia a especialista realizou marcações identificando os trechos nos quais constavam hipsarritmia e onde tinha presença de descargas multifocais Figura 5.2.

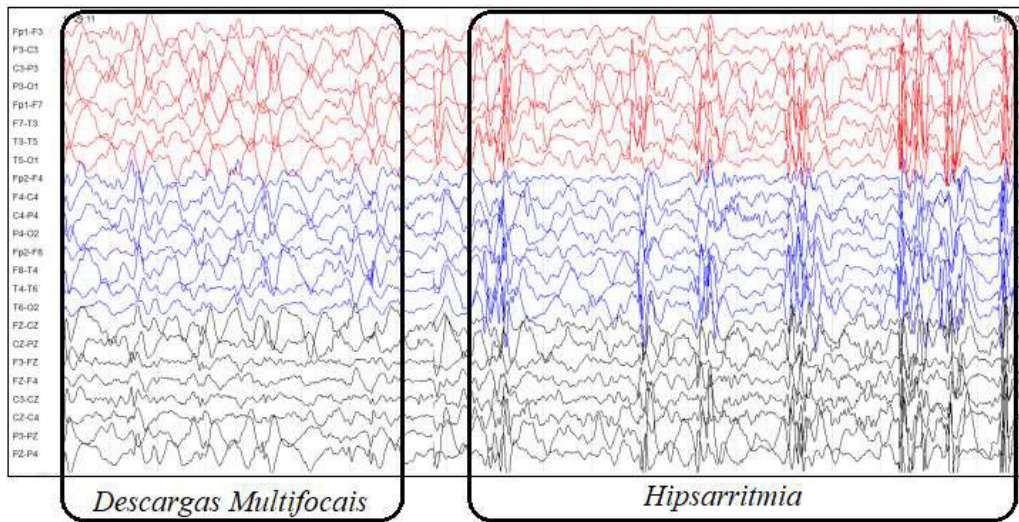


Figura 5.2: EEG com identificação de hipsarritmia e descargas multifocais.

Observando a Figura 5.2, percebe-se então que um exame de EEG com padrão de hipsarritmia pode ser composto por três características: trechos com hipsarritmia, trechos com descargas multifocais e alguns trechos que não foram classificados, a esses trechos denominou-se como “sem alteração”.

5.2.2 Índices de Hipsarritmia

A fim de se estabelecer parâmetros capazes de identificar o padrão de hipsarritmia, são propostos nesse trabalho três índices. Na apresentação de cada índice procurou-se estabelecer definições matemáticas para os mesmos. Nesse sentido, se fez necessário observar cada sinal de EEG como ente matemático sobre o qual é possível se estabelecer definições e propriedades.

Em linguagem matemática portanto, cada janela do EEG pode ser representada por uma matriz de números reais $A_k = a_{ij}$. Dessa maneira é possível, para fins didáticos, realizar o “janelamento” do EEG em um conjunto de k-matrizes de tamanho $m \times n$ (Figura 5.3).

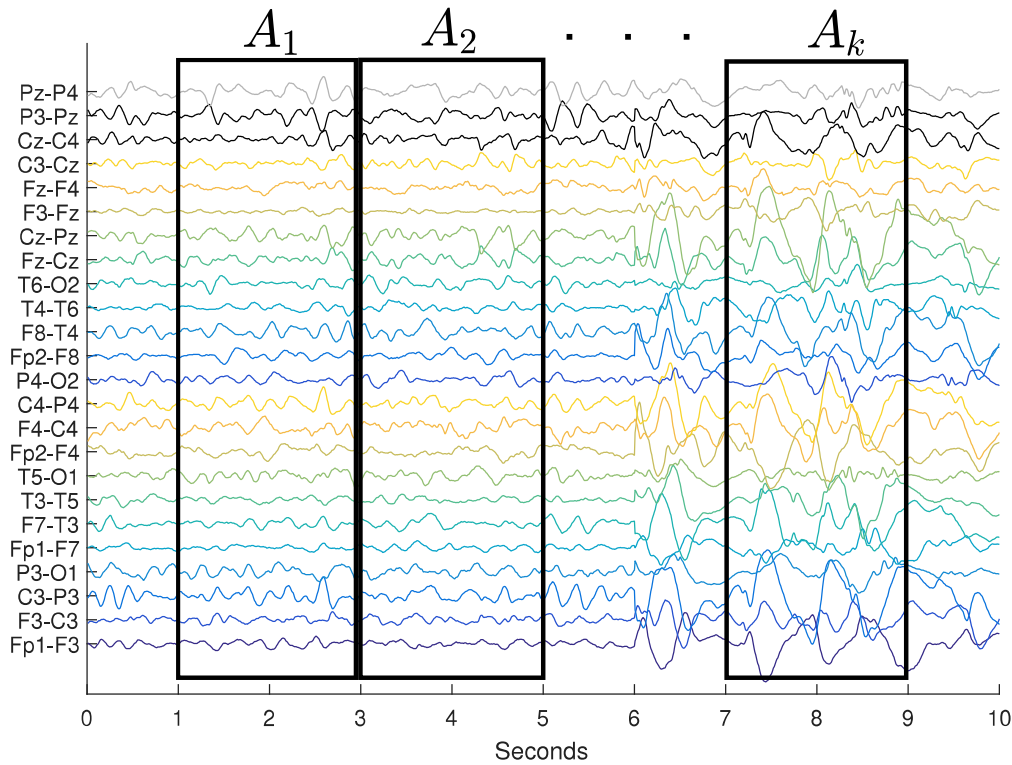


Figura 5.3: Janelamento do sinal de EEG

Com base no janelamento ilustrado na Figura 5.3, foram definidos três índices utilizados na identificação de hipsarritmia, a saber, I^w , I_k e $\bar{I}_k$. Concomitantemente às definições dos índices são apresentados o método de extração de característica em cada situação.

Em resumo, a extração de característica é realizada da seguinte maneira: primeiramente o sinal de EEG é janelado, em seguida calcula-se o respectivo índice de hipsarritmia em cada janela obtendo assim uma sequência de índices. Por fim, a log energia ou a média da sequência de índices é calculada. A log energia é utilizada para avaliar um conjunto de índices do sinal e é definida da seguinte forma (HAN; DONG; XU, 2009; SHARMA; KANKAR, 2014; GUPTA et al., 2017):

$$E_{LgEn} = \sum_{i=1}^N \log(x_i^2), \quad (5.1)$$

em que, x_i é a i -ésima amostra do sinal de comprimento N . Nesse trabalho, a sequência de índices de um sinal de EEG é utilizada para computar a log energia.

- Índice de Hipsarritmia e Extração de Características Utilizando a Transformada Discreta de Wavelet

Dada uma wavelet mãe (discreta) e um sinal de EEG segmentado em janelas como mostrado na Figura 5.3 define-se o índice de hipsarritmia baseado na DWT da seguinte forma:

Definição 7. Seja $\beta_k = \{\beta_{i1}, \beta_{i2}, \beta_{i3}, \dots, \beta_{in}\}$ o conjunto de coeficientes da DWT da i -ésima linha de matriz $A_k(m \times n)$. Definimos o índice I_k^w de tal forma que

$$I_k^w = \frac{\sum_{i=1}^m \sum_{j=1}^n \beta_{ij}}{m \cdot n} \quad (5.2)$$

Assim, para cada janela de w segundos do sinal EEG, obtém-se um índice I_k^w e, consequentemente, para um sinal de k -janelas, uma sequência $I^w = \{I_1^w, I_2^w, I_3^w, \dots, I_k^w\}$.

De posse da sequência de índices, calcula-se a log energia da mesma com a finalidade de extrair um valor peculiar ao EEG e assim concluir o processo de extração de característica. Um resumo de como realizar tal procedimento é ilustrado da Figura 5.4.

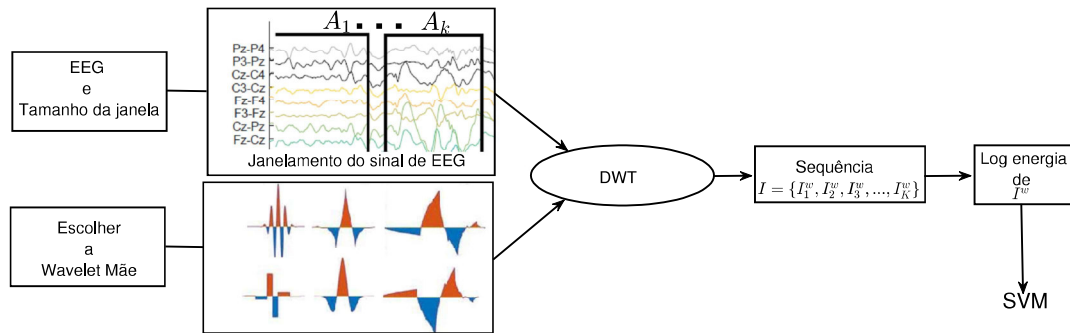


Figura 5.4: Extração de características e classificação do EEG.

- Índice de Hipsarritmia e Extração de Características Utilizando a Transformada Discreta de Wavelet com Núcleo de Gabor

A proposta de utilização de funções de Gabor como núcleo da DWT é fundamentada na percepção de que existe semelhança morfológica entre uma espícula e uma função de Gabor (Figura 5.5).

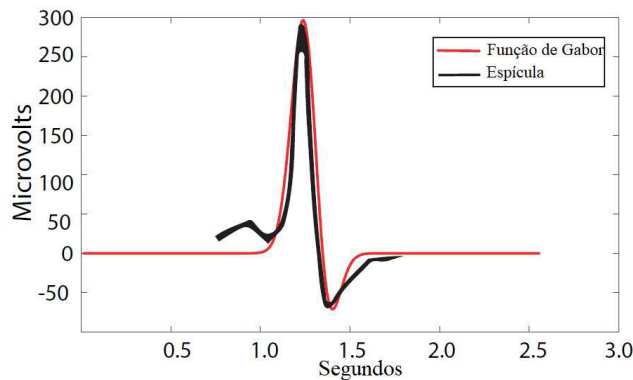


Figura 5.5: Comparação entre uma espícula e uma função de Gabor.

Espera-se que o produto interno entre uma função de Gabor e trechos do sinal que contém espículas, apresente sempre valores superiores aos produtos entre a mesma Gabor e o restante do sinal.

A proposta então, é fazer $\psi_{j,k}(n)$ igual a função de Gabor e, em seguida analisar os coeficientes dessa transformação. A hipótese principal é que como a quantidade de espículas (ondas semelhantes a funções de Gabor) é sempre maior em EEGs com características hipsarrítmicas será possível obter diferenças significativas entre um sinal de EEG com hipsarritmia e um normal.

Nessas condições define-se o índice baseado na DWT com núcleo de Gabor da seguinte forma:

Definição 8. *Seja $\alpha_k = \{\alpha_{i1}, \alpha_{i2}, \alpha_{i3}, \dots, \alpha_{in}\}$ o conjunto de coeficientes da DWT com Núcleo de Gabor da i -ésima linha de matriz $A_k(m \times n)$. Definimos I_k da seguinte forma:*

$$I_k = \frac{\sum_{i=1}^m \sum_{j=1}^n \alpha_{ij}}{m \cdot n} \quad (5.3)$$

De maneira análoga ao observado na Definição 7, para cada janela de w segundos do sinal EEG, obtemos um índice I_k e uma sequência $I = \{I_1, I_2, I_3, \dots, I_K\}$.

Desta forma, pode-se então concluir o processo de extração de característica calculando a log energia da sequência I_k . A Figura 5.6 ilustra como funciona o processo desde a entrada com o EEG e as informações sobre a função de Gabor até o cômputo da log energia.

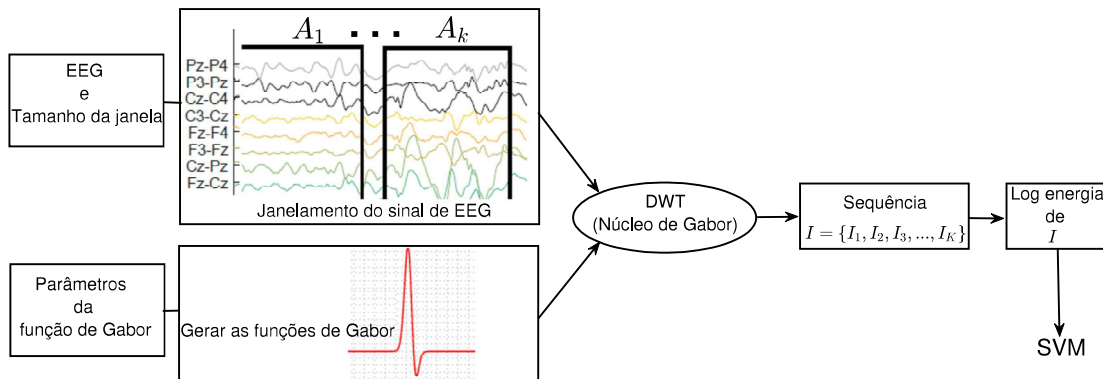


Figura 5.6: Extração de características e classificação do EEG.

- Índice de Hipsarritmia e Extração de Características Utilizando Decomposição em Funções de Gabor Sintonizadas

O cálculo dos índices definidos pelas expressões 5.2 e 5.3 necessitam, além das características de janelamento do EEG, de informações referentes aos tipos de ondas a serem utilizadas como núcleo da DWT. Por outro lado, a decomposição em funções de Gabor sintonizadas é realizada utilizando apenas informações do próprio sinal.

Considerando as matrizes $A_k(m \times n)$ como elementos de um EEG e que cada linha da mesma pode ser decomposta em uma combinação linear de funções de Gabor sintonizadas e ponderadas pelos coeficientes λ_k , define-se o índice $\bar{I}_k$ da seguinte forma:

Definição 9. *Seja $\lambda_k = \{\lambda_{i1}, \lambda_{i2}, \lambda_{i3}, \dots, \lambda_{in}\}$ o conjunto de coeficientes da Decomposição em Funções de Gabor Sintonizadas da i -ésima linha de matriz $A_k(m \times n)$. Definimos $\bar{I}_k$ da seguinte forma:*

$$\bar{I}_k = \frac{\sum_{i=1}^m \sum_{j=1}^n \lambda_{ij}}{m \cdot n} \quad (5.4)$$

Também de modo análogo a I_k^w e a I_k , para cada janela de w segundos, obtemos um índice $\bar{I}_k$ e uma sequência $\bar{I} = \{\bar{I}_1, \bar{I}_2, \bar{I}_3, \dots, \bar{I}_K\}$.

Após calcular os índices em todas as janelas a log energia (ou a média) é calculada. Para fins de se concluir o processo de extração de características do EEG. A Figura 5.7 resume o processo de extração de características utilizando a decomposição em funções de Gabor sintonizadas.

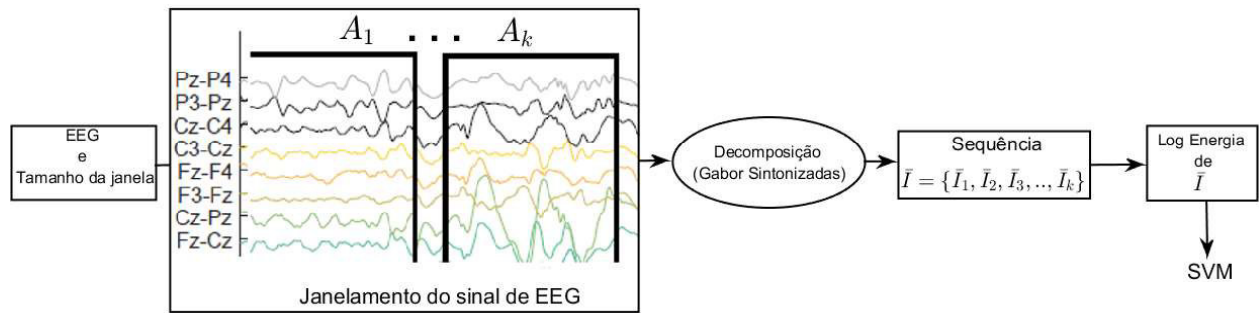


Figura 5.7: Extração de características e classificação do EEG.

É interessante ressaltar que este processo difere dos demais descritos anteriormente principalmente por não necessitar de informações que não sejam oriundas do EEG. Sendo assim esse método de extração de características esquiva-se de questionamentos como por exemplo qual wavelet usar, ou quais parâmetros da função de Gabor deve-se impor.

5.2.3 Métricas de Avaliação

Na etapa de definição das métricas de avaliação (Figura 5.1) ficou estabelecido que os processos de classificação propostos nesse trabalho seriam realizados conforme a técnica de validação cruzada. Além disso a avaliação dos resultados é realizada de duas maneiras, através de testes de hipótese estatísticos e das métricas de especificidade, sensibilidade e acurácia. Desta maneira, são apresentados a seguir os conceitos relativos a tais métricas e técnicas.

- Testes de hipótese

Na realização de uma pesquisa é comum que o pesquisador tenha interesse em tomar uma decisão a respeito de suas descobertas a respeito dos dados coletados. Normalmente essas decisões são baseadas em um conjunto de evidências e em resultados estatísticos. Para uma tomada de decisão bem fundamentada estatisticamente é necessário que o pesquisador defina, a princípio, quais são suas hipóteses. Define-se como hipótese estatística qualquer afirmação acerca de parâmetros da distribuição de probabilidade de uma ou mais variáveis aleatórias.

Para fins matemáticos define-se por H_0 a hipótese que se pretende testar, e a esta denomina-se de hipótese nula. Em contraponto a H_0 é importante que o pesquisador defina uma outra hipótese, a saber, H_1 que consiste na hipótese alternativa. Sobre essas hipóteses então são realizados testes estatísticos que tem como finalidade a minimização de erros. Nos testes estatísticos é possível cometer dois tipos de erros: erro tipo I e tipo II. Na Tabela 5.2 verifica-se a descrição dos possíveis tipos de erro.

Tabela 5.2: Tipos de erros em testes de hipóteses

Decisão	H_0 é verdadeira	H_0 é falsa
Aceitar H_0	Decisão correta	Erro tipo II
Rejeitar H_0	Erro tipo I	Decisão correta

Para que os erros descritos na Tabela 5.2 sejam o menor possível são realizados os chamados testes de hipótese. A seguir a definição matemática de um teste estatístico.

Definição 10. *Um teste de hipótese estatístico é uma função de decisão $d: \chi \rightarrow a_0, a_1$, em que a_0 corresponde à ação de considerar a hipótese H_0 como verdadeira e a_1 corresponde à ação de considerar a hipótese H_1 como verdadeira.*

A realização de um teste de hipótese segue uma sequência de passos. Essa sequência pode ser constatada a seguir:

1º Passo - Estabelecer qual será a hipótese H_0 que contradiz a hipótese alternativa H_1 ;

2º Passo - Formular a hipótese alternativa H_1 , que será adotado se H_0 for rejeitada;

3º Passo- Decidir sobre o valor do nível de significância;

4º Passo - Escolher o teste estatístico adequado em relação ao objeto da investigação e ao tipo de dados coletados;

5º Passo - Comparar o resultado obtido com os valores tabelados de acordo com o nível de significância previamente definidos;

6º Passo - Estabelecer a conclusão em relação ao trabalho efetuado.

- Técnica de Validação cruzada

Em modelos baseados em aprendizado de máquina é possível que ocorra a super adequação do modelo aos dados, o que impossibilita a generalização do modelo para dados novos. Essa super adequação recebe o nome de *overfitting*. Para evitar que ocorra *overfitting* é comum realizar algumas técnicas em modelos baseados em aprendizado de máquina. Uma dessas técnicas é o método de validação cruzada “*k-fold cross-validation*”.

No método de validação cruzada, o conjunto de exemplos é dividido em k subconjuntos de tamanhos aproximadamente iguais. Os objetos de $k - 1$ partições são utilizados no treinamento do classificador e este é então testado na partição restante. Esse processo é repetido k vezes. O desempenho então, é avaliado a partir da média dos desempenhos observados. Essa técnica pode ser ilustrada através da Figura 5.8.

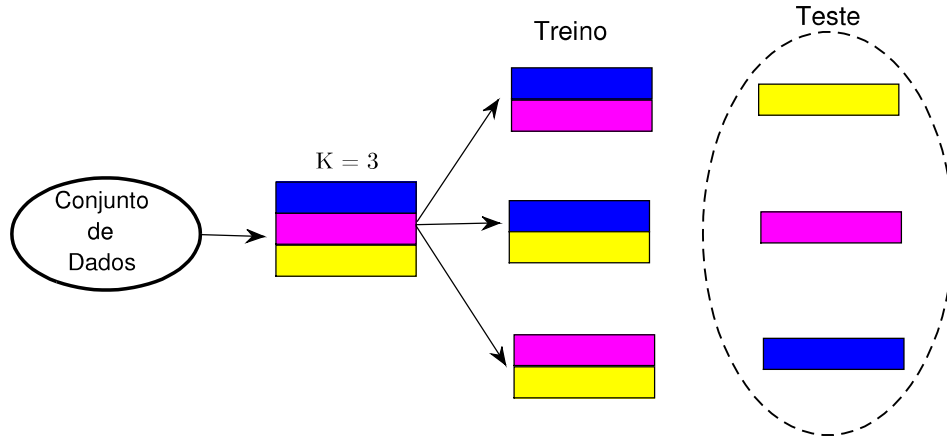


Figura 5.8: Método de validação cruzada

- Métricas de avaliação de desempenho do classificador

Como métrica para avaliar o desempenho do método proposto, são utilizadas também a especificidade, sensibilidade e acurácia. A especificidade mede o quão eficiente é o método para o diagnóstico de pacientes saudáveis. Já sensibilidade mede a eficiência no que tange a tarefa de identificar os sinais com hipsarritmia e acurácia que avalia o quão eficiente é o método em relação ao diagnóstico. A especificidade (SPEC) e sensibilidade (SENS) são definidas, respectivamente, como:

$$SPEC = \frac{VN}{VN + FP} \times 100, \quad (5.5)$$

e

$$SENS = \frac{VP}{VP + FN} \times 100. \quad (5.6)$$

A acurácia (ACC) é definida como

$$ACC = \frac{VP + VN}{VP + VN + FN + FP} \times 100, \quad (5.7)$$

sendo verdadeiro positivo (VP) o sinal de EEG hipsarrítmico é classificado como tal, verdadeiro negativo (VN) ocorre quando o sinal de EEG sem padrão hipsarrítmico é classificado como sem alteração, no falso negativo (FN) o sinal de EEG com hipsarritmia é classificado como normal, por fim, no falso positivo (FP) o sinal de EEG normal é classificado com hipsarrítmico.

5.3 Considerações

Foi apresentado nesse capítulo o delineamento geral dessa pesquisa, como ocorreram as fases do desenvolvimento da mesma e alguns conceitos metodológicos. O banco de dados utilizado foi apresentado de maneira descritiva e os índices matemáticos para identificação de hipsarritmia foram definidos. Além disso foram expostas nesse capítulo as ferramentas estatísticas de avaliação do índice (testes de hipótese) e as de avaliação do processo de classificação (especificidade, sensibilidade e acurácia). Uma breve ilustração da estratégia de validação também esteve presente no capítulo.

No delineamento da pesquisa ficou claro que a mesma foi realizada em quatro fases, e nesse capítulo foi percorrido sobre as três primeiras. Os elementos da quarta e última fase de desenvolvimento desse trabalho porém, serão detalhados nos capítulos seguintes pois tratam dos resultados obtidos.

Resultados

A organização deste capítulo é composta de duas seções. Na primeira seção consta a classificação de EEG's com ou sem hipsarritmia. Já na segunda seção foram utilizados os EEG's com hipsarritmia a fim de se identificar os trechos que possuem o padrão hipsarrítmico e os que são oriundos de descargas multifocais.

6.1 Identificação de EEG's com Hipsarritmia

Os resultados apresentados nessa seção foram obtidos a partir da utilização de 203 sinais de eletroencefalograma de crianças com microcefalia. Esses EEGs tem duração de 5 minutos cada e foram segmentados em 30 janelas de 10 segundos. Para cada janela, foi calculado um índice I_k^w , um I_k e um $\bar{I}_k$, compondo três sequências de 30 índices para cada EEG. A média - μ e a log energia - E_{LgEn} de cada sequência foram calculadas, considerando diferentes tipos de Wavelets, diferentes combinações de parâmetros das funções Gabor e a decomposição em funções de Gabor Sintonizadas.

A apresentação dos resultados está organizada em três subtópicos. No primeiro considera-se o índice I_k^w baseado na DWT, variando as wavelets mãe. No segundo subtópico são apresentados resultados da aplicação do índice I_k definido a partir de funções de Gabor. No terceiro e último subtópico os resultados apresentados são oriundos da aplicação de $\bar{I}_k$, que por sua vez, é obtido a partir da decomposição em funções de Gabor sintonizadas.

Em cada subtópico são apresentados boxplots, testes de hipótese e os resultados de classificação. O primeiro teste de hipótese a ser aplicado foi o de Kolmogorov - Smirnov. O objetivo do teste de Kolmogorov - Smirnov é verificar a hipótese de que os dados provêm de uma distribuição normal.

Após a observação dos resultados obtidos pelo teste de Kolmogorov - Smirnov, foi realizado o teste de Mann - Whitney, a fim de testar a hipótese nula de que os dados são amostras de distribuições contínuas com medianas iguais (ao nível de significância de 5 %).

Para classificação foi utilizada uma Máquina de Vetores de Suporte-SVM. Em todas as simulações, os parâmetros utilizados foram: kernel linear e slack igual a 1. Os resultados de classificação para cada índice são apresentados em tabelas no final de cada subtópico. Nas tabelas estão presentes informações sobre as médias e os desvios da acurácia, sensibilidade e

especificidade, quando a técnica de validação cruzada foi aplicada.

6.1.1 Resultados do cálculo e aplicação do índice de Hipsarritmia baseado na DWT

Para aplicar a Transformada Discreta de Wavelet é necessário primeiramente escolher uma "Wavelet mãe". Atualmente na literatura existem diversas wavelets mãe, e esse fato em alguns casos pode implicar na dúvida sobre qual wavelet deve ser utilizada.

Para essas simulações foram utilizadas como wavelets mãe as seguintes ondas: Daubechies3-db3, Coiflet-coif, Haar e Symlets-sym4. Essas Wavelets foram escolhidas depois de se observar duas coisas: a primeira observação é que Daubechies3-db3, Coiflet-coif, Haar e Symlets-sym4 são ondas bastante utilizadas no processamento de sinais, a outra observação é que durante os testes computacionais não se percebeu diferenças significativas quando se utilizou outros tipos de Wavelets. Um resumo da metodologia utilizada nessas simulações pode ser encontrado na Figura 5.4. Os boxplots da Figura 6.1 apresentam as distribuições das médias dos índices I_k^w dos EEG's com e sem hipsarritmia para cada wavelet mãe utilizada.

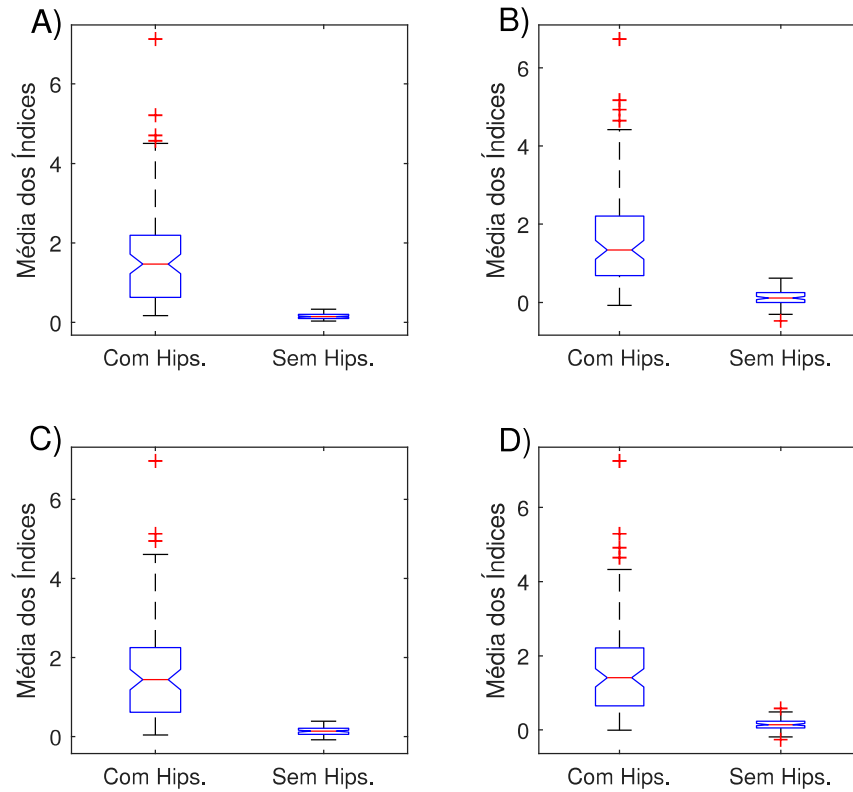


Figura 6.1: Boxplot com as médias dos índices de hipsarritmia: A) coif B) db3 C) Haar D) sym4

Aparentemente existem diferenças importantes entre as distribuições apresentadas na Figura 6.1 para todos os casos simulados. Percebe-se que, em média, os coeficientes da Transformada de Wavelet dos sinais com hirsarritmia são maiores do que os de sinais sem alteração. Uma outra observação importante é que as distribuições apresentadas nos boxplots são bastante semelhantes quando se modifica a wavelet mãe, ou seja, modificar a wavelet aparentemente não traz mudanças nos resultados.

As observações destacadas quanto a média dos índices I_k^w também são evidenciadas quando se analisa as log energia dos mesmos (Figura 6.2).

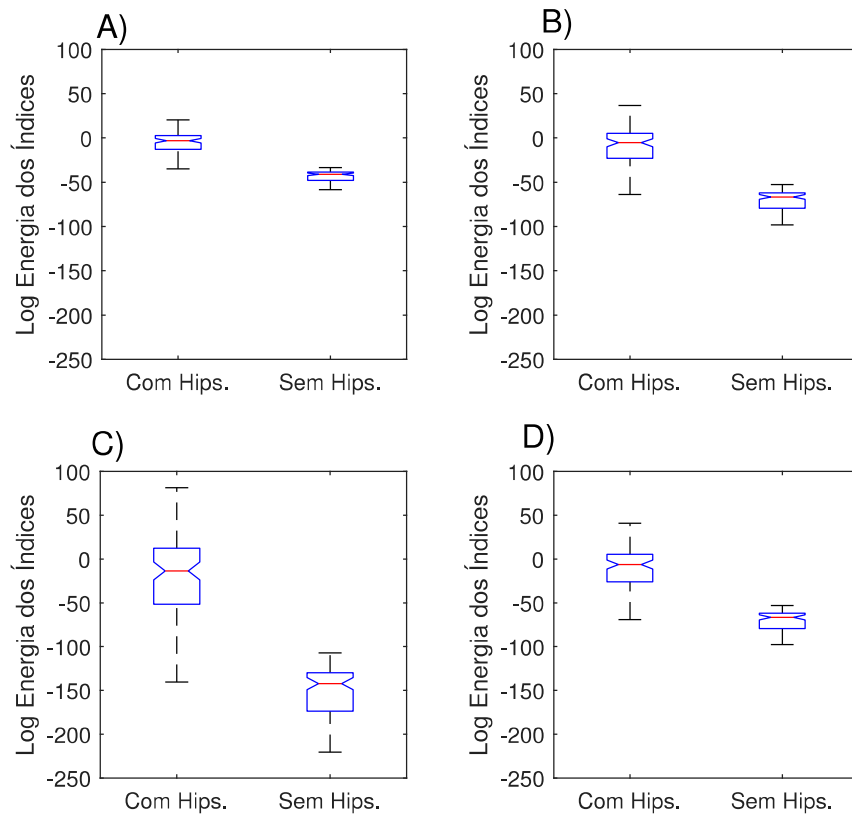


Figura 6.2: Boxplot com as log energia dos índices de hirsarritmia: A) coif B) db3 C) Haar D) sym4

A partir dos dados apresentados nos boxplots da Figura 6.1 e da Figura 6.2, foram realizados dois testes de hipóteses. No primeiro teste (Kolmogorov - Smirnov), foi averiguada a hipótese nula de que os dados provêm de uma distribuição normal. Neste teste, o resultado é 1 se o teste rejeitar a hipótese nula em um nível de significância de 5 %. Os resultados deste teste indicaram a rejeição da hipótese nula em todas distribuições.

Dado que o teste de Kolmogorov - Smirnov apontou para rejeição da hipótese de que os dados são oriundos de uma distribuição normal então optou-se por realizar o teste de Mann - Whitney, a fim de testar a hipótese nula de que os dados são amostras de distribuições contínuas

com medianas iguais ao nível de significância de 5 %.

Os resultados obtidos com o teste de Mann-Whitney indicaram que a hipótese nula deve ser rejeitada em todas as situações de comparação apresentadas. As informações referentes aos de p-valor de cada situação estão descritas na Tabela 6.1.

Tabela 6.1: Teste de Hipótese.

Wavelet Mãe	Característica	p-valor
Coiff	Média	$3,34 \cdot 10^{-33}$
Coiff	Log Energia	$1,10 \cdot 10^{-33}$
Db3	Média	$3,04 \cdot 10^{-33}$
Db3	Log Energia	$2,16 \cdot 10^{-33}$
Haar	Média	$3,00 \cdot 10^{-33}$
Haar	Log Energia	$1,04 \cdot 10^{-33}$
Sym4	Média	$3,28 \cdot 10^{-33}$
Sym4	Log Energia	$1,01 \cdot 10^{-33}$

Os testes de hipótese mostraram portanto, que existem diferenças significativas entre a média dos índices I_k^w dos EEG's com e sem hirsarritmia. Essas médias foram utilizadas como entrada em uma SVM. Os resultados de classificação estão apresentados na Tabela 6.2.

Tabela 6.2: Classificação quando as médias de I_k^w são utilizadas como entrada da SVM.

Wavelet Mãe	Acurácia (%)	Sensibilidade (%)	Especificidade (%)
Coif	$90,00 \pm 5,19$	$89,78 \pm 7,56$	$91,67 \pm 7,56$
Db3	$93,84 \pm 7,06$	$91,23 \pm 9,32$	$90,55 \pm 10,07$
Haar	$81,34 \pm 6,53$	$80,46 \pm 7,37$	$81,46 \pm 6,07$
Sym4	$91,53 \pm 7,64$	$89,69 \pm 1,45$	$88,32 \pm 2,22$

Como observado na Tabela 6.1, existem diferenças estatisticamente significativas entre as log energias dos índices dos EEG's com e sem hirsarritmia. Esses valores de log energia foram testados como valores de entrada do classificador (SVM). Os resultados da classificação estão expostos na Tabela 6.3.

Tabela 6.3: Classificação quando as log energias de I_k^w são utilizadas como entrada da SVM.

Wavelet Mãe	Acurácia (%)	Sensibilidade (%)	Especificidade (%)
Coiff	$93,84 \pm 4,86$	$92,16 \pm 6,83$	$91,76 \pm 7,33$
Db3	$95,38 \pm 5,37$	$96,34 \pm 5,94$	$94,23 \pm 6,31$
Haar	$95,38 \pm 5,37$	$96,13 \pm 10,09$	$94,22 \pm 10,09$
Sym4	$96,15 \pm 5,43$	$95,98 \pm 5,32$	$94,09 \pm 6,27$

Em uma avaliação inicial percebe-se que quando o método baseado na DWT é utilizado a log energia dos índices, no geral, apresenta maior eficácia no processo de classificação.

6.1.2 Resultados do cálculo e aplicação do índice baseado na DWT com núcleo de Gabor

Os resultados apresentados nessa subseção são implicações do uso de funções de Gabor como núcleo da DWT. Essa metodologia foi definida anteriormente e resumida na Figura 5.7 e consiste no cálculo de uma sequência de índices I_k .

Vários experimentos computacionais foram realizados para identificar os EEGs com padrões hipsarrítmicos. O boxplot apresentado na Figura 6.3 apresenta as distribuições de médias de μ dos índices para os grupos EEGs com hipsarritmia e os EEGs com traçado sem anormalidades para diferentes parâmetros da função de Gabor.

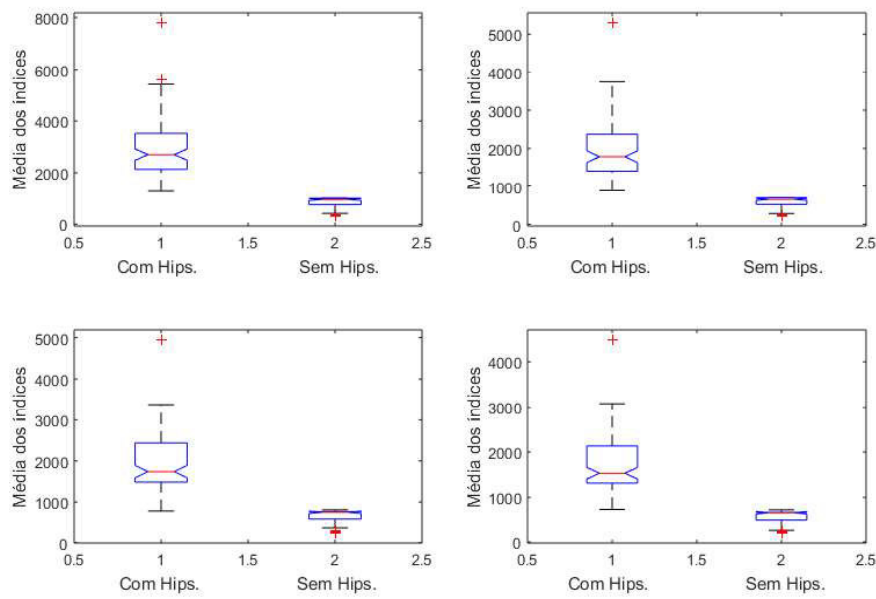


Figura 6.3: Boxplot com as médias dos índices de hipsarritmia. Parâmetros da função de Gabor: A) $\sigma^2 = 20$, $\omega = 1.02$ e $\phi = 5$, B) $\sigma^2 = 5$, $\omega = 1,40$ e $\phi = 5$, C) $\sigma^2 = 15$, $\omega = 2,68$ e $\phi = 5$ e D) $\sigma^2 = 10$, $\omega = 3,20$ e $\phi = 5$.

Nestes gráficos observou-se que, em média, os índices dos EEGs com hipsarritmia são maiores que os índices com traçados normais. Esse fato também pode ser verificado nos boxplots da Figura 6.4. Essa figura mostra a distribuição da E_{LgEn} dos índices para os grupos de EEGs com hipsarritmia e os com traçados sem anormalidades.

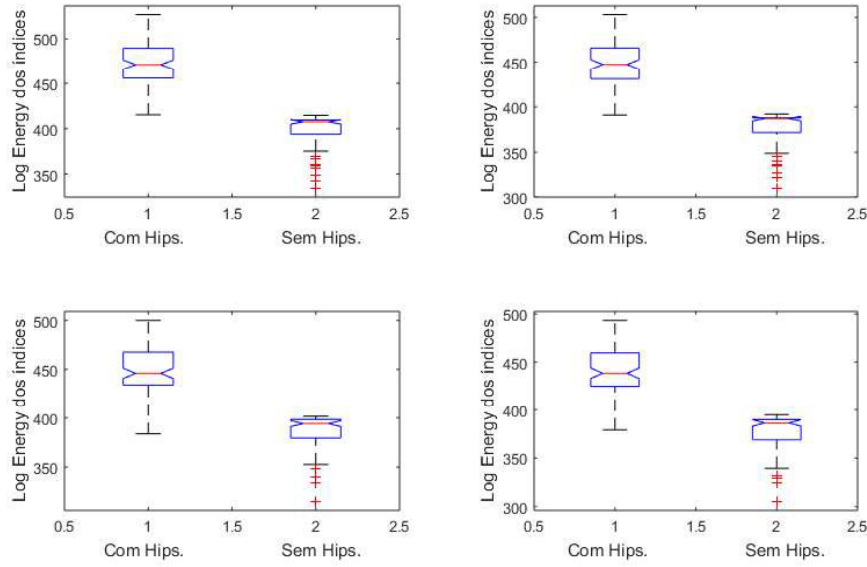


Figura 6.4: Boxplot com as log energia dos índices de hirsarritmia. Parâmetros da função de Gabor: A) $\sigma^2 = 20$, $\omega = 1.02$ e $\phi = 5$, B) $\sigma^2 = 5$, $\omega = 1,40$ e $\phi = 5$, C) $\sigma^2 = 15$, $\omega = 2,68$ e $\phi = 5$ e D) $\sigma^2 = 10$, $\omega = 3,20$ e $\phi = 5$.

Para os valores de log energia dos índices também foram aplicados a sequência de aplicação dos testes de hipóteses Kolmogorov - Smirnov e Mann-Whitney. Os resultados mostraram diferenças significativas entre as log energia dos índices de EEG's com e sem hirsarritmia, com os seguintes valores de p-valor: $p = 9,77 \times 10^{-34}$ (A), $p = 3,71 \times 10^{-33}$ (B), $p = 4,84 \times 10^{-32}$ (C), $p = 3,18 \times 10^{-32}$ (D).

Os valores de E_{LgEn} , calculados através da Equação(5.1), e as médias dos índices de hirsarritmia de cada EEG foram utilizados como características de entrada em uma Máquina de Vetores de Suporte, para realizar a classificação.

É exibido a seguir uma visão geral de algumas simulações, considerando a mudança nos parâmetros das Funções do Gabor. Foi utilizada a metodologia de validação cruzada de dez vezes para cada função Gabor. Os valores médios de sensibilidade, especificidade e acurácia são exibidos nas tabelas a seguir.

Na Tabela 6.4 foi utilizado com característica de entrada da SVM apenas a média μ dos valores dos índices de cada EEG.

Tabela 6.4: Avaliação dos resultados para diferentes funções de Gabor usando somente as médias dos índices como entrada da SVM.

Parâmetros (%)	Acurácia (%)	Sensibilidade (%)	Especificidade (%)
$\sigma^2 = 5, \omega = 1,02$ e $\phi = 5$	98,52 %	100,00%	96,94 %
$\sigma^2 = 20, \omega = 1,02$ e $\phi = 5$	98,52 %	100,00 %	96,94%
$\sigma^2 = 5, \omega = 1,40$ e $\phi = 5$	98,52 %	100,00 %	96,94%
$\sigma^2 = 10, \omega = 1,53$ e $\phi = 4$	98,52 %	100,00 %	96,94%
$\sigma^2 = 15, \omega = 1,53$ e $\phi = 5$	97,54 %	100,00 %	94,90%
$\sigma^2 = 15, \omega = 1,92$ e $\phi = 4$	98,52 %	100,00 %	96,94%
$\sigma^2 = 5, \omega = 2,56$ e $\phi = 4$	98,52 %	100,00 %	96,94%
$\sigma^2 = 15, \omega = 2,68$ e $\phi = 4$	96,71 %	100,00 %	96,94%
$\sigma^2 = 15, \omega = 2,68$ e $\phi = 5$	95,57 %	100,00 %	90,82%
$\sigma^2 = 10, \omega = 3,20$ e $\phi = 5$	95,57 %	100,00 %	90,82%

Na Tabela 6.5, estão expressos os resultados de classificação quando os valores de E_{LgEn} foram utilizados como entrada na SVM.

Tabela 6.5: Avaliação dos resultados para diferentes funções de Gabor usando somente as E_{LgEn} dos índices como entrada da SVM.

Parâmetros (%)	Acurácia (%)	Sensibilidade (%)	Especificidade (%)
$\sigma^2 = 5, \omega = 1,02$ e $\phi = 5$	96,55 %	100,00%	92,85 %
$\sigma^2 = 20, \omega = 1,02$ e $\phi = 5$	96,88 %	100,00 %	93,53%
$\sigma^2 = 5, \omega = 1,40$ e $\phi = 5$	96,71 %	100,00 %	93,19%
$\sigma^2 = 10, \omega = 1,53$ e $\phi = 4$	96,22 %	100,00 %	92,17%
$\sigma^2 = 15, \omega = 1,53$ e $\phi = 5$	96,55 %	100,00 %	92,85%
$\sigma^2 = 15, \omega = 1,92$ e $\phi = 4$	96,22 %	100,00 %	92,17%
$\sigma^2 = 5, \omega = 2,56$ e $\phi = 4$	96,55 %	100,00 %	92,85%
$\sigma^2 = 15, \omega = 2,68$ e $\phi = 4$	96,71 %	100,00 %	93,19%
$\sigma^2 = 15, \omega = 2,68$ e $\phi = 5$	95,40 %	99,36 %	91,15%
$\sigma^2 = 10, \omega = 3,20$ e $\phi = 5$	95,73 %	99,68 %	91,49%

6.1.3 Resultados do cálculo e aplicação do índice Baseado na Decomposição em Funções de Gabor Sintonizadas

A utilização da DWT para processamento de sinais traz consigo um questionamento ainda em aberto, sobre qual é a melhor onda a ser utilizada. Por outro lado, inserir funções de Gabor no núcleo da DWT também implica no fato de não ter com clareza quais os parâmetros que se deve utilizar.

Nesse contexto é que nesse subtópico são apresentados os resultados obtidos quando se decompõe um sinal em funções de Gabor sintonizadas. Nessa metodologia todas as informações utilizadas como parâmetros de entrada do processo são oriundas do sinal analisado.

O processo de decomposição em Gabor sintonizadas consiste em aproximar um sinal através de uma combinação linear de pequenas ondas. A Figura 6.5 reflete uma decomposição utilizando 100 funções de Gabor sintonizadas. O Erro Quadrático Médio - EQM entre o sinal original - $X(t_i)$ e o sinal aproximado pela decomposição - $\tilde{X}(t_i)$ foi de $28,91 \mu V^2$ ($EQM = \frac{1}{n} \sum_{i=1}^n (X(t_i) - \tilde{X}(t_i))^2$). Já o Erro Absoluto Médio - EAM foi igual a $3,74 \mu V$ ($EAM = \frac{1}{n} \sum_{i=1}^n |X(t_i) - \tilde{X}(t_i)|$).

Os valores de erro indicam que a metodologia proposta resulta em uma representação esparsa do sinal original de tal forma que a decomposição/aproximação seja apropriada. Além da boa aproximação refletida nas métricas de erro, percebe-se uma redução na quantidade de informação necessária para representar o sinal original, pois partindo de um conjunto de 1280 amostras (Fig. 6.5A) reduziu-se para 100 coeficientes (Fig. 6.5B) e 400 parâmetros de formação das funções de base (100 parâmetros de cada).

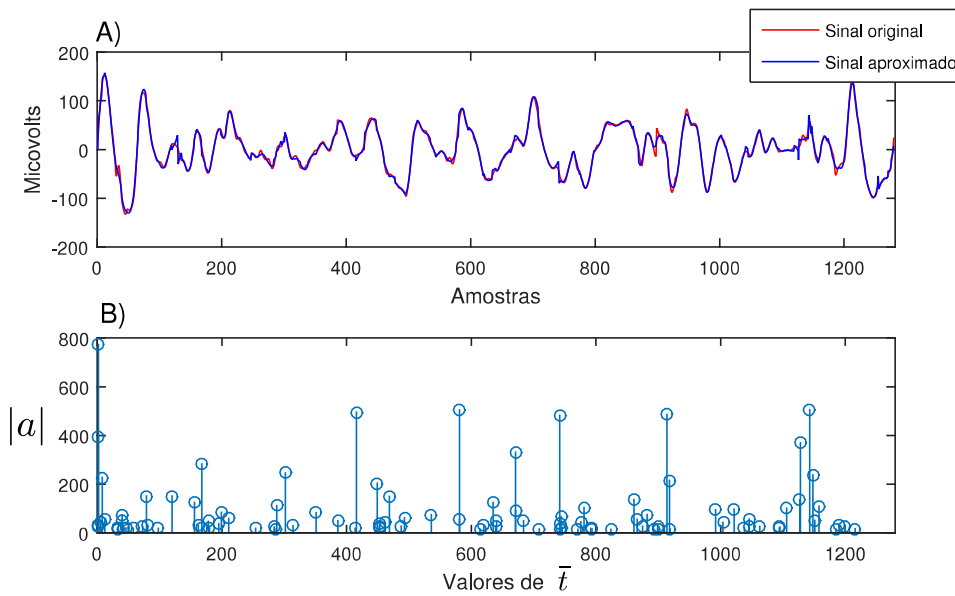


Figura 6.5: Exemplo de decomposição em funções de Gabor: A) comparação entre o sinal original e o aproximado para 100 funções de Gabor sintonizadas B) representação esparsa do sinal original

A partir do processo de decomposição exemplificado na Figura 6.5, utilizou-se os valores de $|a|$ para o cálculo dos índices $\bar{I}_k$. As distribuições das médias e das log energia das sequências de $\bar{I}_k$ para os EEG's com e sem hirsarritmia estão evidenciadas nos boxplots da Figura 6.6.

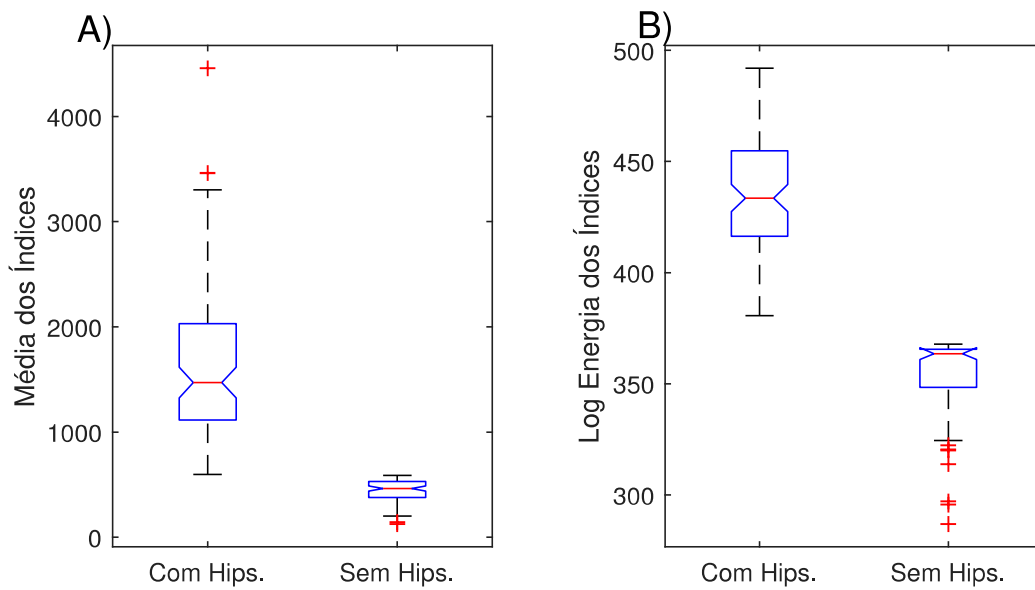


Figura 6.6: Boxplot com as médias e log energia dos índices baseados na decomposição de funções de Gabor sintonizadas: A) Média dos índices B) Log energia dos índices

Os boxplots apresentados Figura 6.6B mostram que as características log energia dos índices são linearmente separáveis. Esse fato é importante pois implica na possibilidade de 100% de classificação correta.

Para o processo de classificação de EEG's tendo como entrada informações como média e log energia dos índices $\bar{I}_k$ foi utilizada uma SVM e o método de validação cruzada. Os resultados, em termos de acurácia, sensibilidade e especificidade, estão dispostos na Tabela 6.6.

Tabela 6.6: Avaliação dos resultados de classificação

Característica	Acurácia (%)	Sensibilidade (%)	Especificidade (%)
Média	100,00± 0,00	100,00± 0,00	100,00± 0,00
Log energia	100,00± 0,00	100,00± 0,00	100,00± 0,00

Os resultados de classificação expostos na Tabela 6.6, ratificam a observação sobre a linearidade de separação evidenciada nos box plots.

6.2 Identificação de Hipsarritmia e Descargas Focais em Eletroencefalogramas

A fim de verificar se os índices propostos nesse trabalho seriam capazes de evidenciar diferenças entre trechos dos EEG's com características hipsarrítmicas alguns testes foram realizados. Primeiramente uma amostra aleatória de 19 EEG's com hipsarritmia foi escolhida.

A duração de cada EEG selecionado foi de 300 segundos, totalizando 5700s de sinal. Uma especialista em leitura eletroencefalográfica fez marcações nos mesmos. Essas marcações cul-

minaram na definição de três grupos de EEG's com as seguintes características: hipsarritmia, descargas multifocais e sem alteração.

Após a marcação dos EEG's ficaram dispostos 3 grupos de trechos, a saber, os marcados como hipsarritmia, os com descargas focais e os trechos sem marcação. O objetivo então, foi aplicar a metodologia propostas a fim de classificar os trechos mencionados. Os trechos de cada grupo foram concatenados e segmentados em janelas de 20s. A duração total de cada grupo após a concatenação e a quantidade de janelas estão dispostas na Tabela 6.7.

Tabela 6.7: Duração total dos segmentos de cada grupo identificado

Grupo	Tempo total	Número de janelas
Hipsarritmia	2980 seg.	149
Descargas focais	1120 seg.	56
Sem marcação	1200 seg.	60

Foram descartados 400 segundos de EEG, pois correspondiam a presença de artefatos. Para cada grupo identificado foram calculadas as sequências de índices I^w , I e $\bar{I}$. Essas sequências foram utilizadas como entrada da SVM. O método de validação cruzada de 10 vezes foi aplicado para avaliação da classificação.

Foram avaliadas combinações de dois a dois grupos no processo de classificação. As combinações se deram da seguinte maneira Hipsarritmia x Sem marcação, Hipsarritmia x Descargas focais e Descargas focais x Sem marcação. Na Tabela 6.8 constam os resultados do processo de classificação quanto aos grupos de segmentos de hipsarritmia e de descargas focais.

Tabela 6.8: Resultados da classificação de Hipsarritmia e Descargas focais

Kernel	Método	Acurácia (%)	Sensibilidade (%)	Especificidade (%)
RBF	Gabor Sintonizada	88,46 ± 9,7	90,3 ± 7,1	89,3 ± 8,2
	DWT (Gabor)	78,46 ± 12,9	80,29 ± 10,09	79,02 ± 11,09
	DWT	72,31 ± 5,3	71,43 ± 5,3	70,21 ± 7,3
LINEAR	Gabor Sintonizada	87,69 ± 5,38	88,16 ± 2,7	86,71 ± 4,7
	DWT (Gabor)	73,85 ± 16,7	74,01 ± 17,16	72,01 ± 16,02
	DWT	75,38 ± 10,7	76,49 ± 9,06	74,08 ± 11,7

Observando a Tabela 6.8 percebe-se uma maior eficiência em termos de acurácia, sensibilidade e especificidade quando o índice baseado em decomposição de funções de Gabor sintonizadas é utilizado.

Outra tarefa realizada foi classificar os segmentos das descargas focais e os que não obtiveram marcação. Na Tabela 6.9 constam os resultados dos testes de classificação.

Tabela 6.9: Resultados da classificação de Descargas focais e Sem marcação

Kernel	Método	Acurácia (%)	Sensibilidade (%)	Especificidade (%)
RBF	Gabor Sintonizadas	$96,36 \pm 4,6$	$96,90 \pm 6,5$	$96,90 \pm 6,5$
	DWT (Gabor)	$91,22 \pm 9,6$	$92,25 \pm 8,4$	$90,25 \pm 10,2$
	DWT	$70,9 \pm 11,18$	$68,12 \pm 13,9$	$68,12 \pm 13,9$
LINEAR	Gabor Sintonizadas	$96,18 \pm 4,75$	$97,14 \pm 3,04$	$95,23 \pm 5,04$
	DWT (Gabor)	$96,36 \pm 4,69$	$97,33 \pm 3,09$	$95,22 \pm 5,09$
	DWT	$50,00 \pm 15,00$	$50,00 \pm 15,00$	$50,00 \pm 15,00$

Por fim a última combinação de grupos foi avaliada, ou seja, o processo de classificação foi aplicado para diferenciar os EEG's com Hipsarritmia e os que não obtiveram marcação. A Tabela 6.10 expressa tais resultados.

Tabela 6.10: Resultados da classificação de Hipsarritmia e Sem marcação

Kernel	Método	Acurácia (%)	Sensibilidade (%)	Especificidade (%)
RBF	Gabor Sintonizadas	$100 \pm 0,0$	$100 \pm 0,0$	$100 \pm 0,0$
	DWT (Gabor)	$100 \pm 0,0$	$100 \pm 0,0$	$100 \pm 0,0$
	DWT	$76,00 \pm 9,60$	$77,00 \pm 10,86$	$75 \pm 11,62$
LINEAR	Gabor Sintonizadas	$100 \pm 0,0$	$100 \pm 0,0$	$100 \pm 0,0$
	DWT(Gabor)	$98,5 \pm 2,4$	$99,2 \pm 1,3$	$99,1 \pm 1,5$
	DWT	$71,50 \pm 8,18$	$73,02 \pm 8,18$	$70,98 \pm 11,31$

Uma breve análise dos resultados apresentados nas Tabelas 6.8 - 6.10, permite supor que o desafio maior consiste em diferenciar hipsarritmia das descargas focais. Porém, o índice baseado na decomposição em funções de Gabor captura características que, quando aplicado na SVM, resulta em taxas de classificação importantes.

6.3 Considerações

Nesse capítulo foram apresentados os principais resultados obtidos nas simulações realizadas na execução desse trabalho. Para fins didáticos, os experimentos foram divididos em dois grupos, um com a finalidade de classificar EEG's e outro com a finalidade de classificar trechos de eletroencefalograma. De maneira resumida pode se enfatizar que os resultados oriundos do índice baseado na decomposição em funções de Gabor sintonizadas foi, em geral, melhores do que os resultantes dos demais índices apresentados.

Discussões

Neste trabalho, foi proposto um método computacional capaz de identificar o padrão de hipsarritmia em exames de eletroencefalogramas de crianças com microcefalia infectadas pelo vírus Zika. Este método foi criado combinando índices matemáticos e uma SVM. No desenvolvimento desse trabalho três índices foram propostos e os resultados obtidos com a aplicação de cada índice foram apresentados. O primeiro índice foi definido tendo como base a decomposição de sinais em wavelets. O segundo teve como fundamentação teórica a combinação entre Transformada de Wavelet e funções de Gabor. O terceiro índice é resultado de uma proposta de decomposição de sinais em funções de Gabor sintonizadas.

Uma das contribuições desse trabalho consiste na proposta de metodologia para decomposição de sinais em pequenas ondas. Apesar da estratégia de decompor sinais em combinação linear de outros sinais não ser em si uma novidade, vale ressaltar que a proposta apresentada aqui difere das demais pelo fato do processo de decomposição em funções de Gabor sintonizadas utilizar apenas informações oriundas do sinal analisado. A Figura 6.5 combinada com os valores de erro ($EQM = 28,91$ e $EAM = 3,74$) demonstram a eficácia do método.

Além das contribuições teóricas e ou metodológicas fundamentada na definição dos índices e no método de decomposição de sinais, neste trabalho também está contida uma aplicação prática. Essa aplicação consiste na identificação de hipsarritmia em eletroencefalograma. A aplicação da metodologia proposta nesse trabalho pode auxiliar o especialista em eletroencefalografia na identificação de hipsarritmia e de descargas epilépticas multifocais.

Os índices apresentados nesse trabalho foram utilizados em dois tipos de experimentos. No primeiro o objetivo foi identificar EEG's com padrão de hipsarritmia. O segundo grupo de experimento consistiu em identificar trechos com hipsarritmia e trechos com descargas focais.

O primeiro índice testado foi o que tem como base teórica a $DWT-I^w$. Avaliando as Figuras 6.1 e 6.2 percebe-se que tanto a log energia quanto a média dos índices, aparentemente, são métricas capazes de evidenciar diferenças entre os EEG's com e sem hipsarritmia. Essa sensação é confirmada quando se analisa os resultados apresentados pelos testes de hipóteses 6.1 e os resultados obtidos na classificação, pois os valores de sensibilidade, especificidade e acurácia variam em torno de 80% a 90%.

Tendo em vista que a utilização da DWT é uma metodologia clássica de análise de sinais, os resultados apresentados evidenciam que essa técnica captura informações importantes nos

EEG's com hipsarritmia, o que possibilita estabelecer diferenças estatisticamente significativas entre os grupos em análise. Essas diferenças ficam evidentes quando se compara, por exemplo, os resultados apresentados nas Tabelas 6.2 e 6.3. Ao compara essas tabelas fica evidente que a energia (ou log energia) dos índices baseados na DWT é uma característica que possibilita uma melhor discriminação dos grupos.

Os resultados apresentados nas Figura 6.3 e 6.4 confirmam as hipóteses sugeridas na motivação para definição do índice baseado na DWT com núcleo de Gabor- I , de que os valores do índice dos eletroencefalogramas com hipsarritmia são, em geral, maiores que os dos encefalogramas sem anormalidades. Essa observação é confirmada ao se aplicar os testes de hipótese, pois verificou-se com os testes estatísticos que as diferenças entre os EEG's com e sem hipsarritmia são estatisticamente significativas.

No que tange ao processo de classificação quando I é utilizado, vale ressaltar que os valores de sensibilidade obtidos nas simulações merecem destaque, pois a taxa 100 % foi alcançada em quase todas as simulações, sendo as exceções os casos em que os parâmetros ($\sigma^2 = 15$, $\omega = 2,68$ e $\phi = 5$) e ($\sigma^2 = 10$, $\omega = 3,20$ e $\phi = 5$) foram utilizados (Tabela 6.4 e Tabela 6.5). As altas taxas de sensibilidade alcançadas em todas as simulações realizadas sugerem que o índice baseado na DWT com núcleo de Gabor é capaz de identificar, com alta precisão, os eletroencefalogramas com padrão de hipsarritmia. Apesar do destaque dado aos resultados de sensibilidade encontrados, as métricas acurácia e especificidade também obtiveram percentuais elevados.

Um outro ponto positivo da aplicação do índice I é que o mesmo pode ser facilmente implementado, uma vez que se baseia na combinação de operações matemáticas já bem estabelecidas na literatura, como, por exemplo, a média, a energia e a DWT. Esse fato torna o cálculo computacional do índice bastante rápido e passível de implementação em sistemas de auxílio médico.

É importante destacar ainda o fato de que os parâmetros das funções de Gabor escolhidos (com base nas características morfológicas das espículas), destacam de forma matemática, uma das características fisiológicas das crianças com epilepsia. Esta característica consiste em um grande número de descargas elétricas neuronais na forma de ondas muito peculiares (Espículas).

Ao fixar os parâmetros da função de Gabor e realizar a DWT pode potencializar a energia de trechos dos sinais com ondas semelhantes a Gabor utilizada. Porém, vale lembrar que as características morfológicas das ondas que compõe os EEG's com padrão hipsarritmia são bastante diversificadas. Nesse contexto então, parece ser mais salutar não utilizar apenas uma onda e sim um conjunto de funções de Gabor. Assim, portanto, surge a proposta de decompor o sinal de EEG em um conjunto de funções de Gabor e a partir de então definir um índice.

Os resultados obtidos com a aplicação do índice baseado na decomposição de funções de Gabor- $\bar{I}$ se mostraram, em geral, melhores do que os obtidos com os demais índices. A log energia dos $\bar{I}$, por exemplo, apresentou uma separabilidade linear entre os grupos de EEG's com e sem hipsarritmia (Figura 6.6B). No processo de classificação através da SVM a taxa de 100% foi alcançada em todas as simulações (Tabela 6.6). O bom desempenho da utilização do índice $\bar{I}$ deve-se principalmente ao fato do mesmo computar a energia de diversas ondas presentes no EEG.

O segundo grupo de experimentos computacionais realizados consistiu na tarefa de classi-

ficar trechos dos sinais com hirsarritmia. Três conjuntos com características diferentes foram classificados. A tarefa de classificação dos trechos com hirsarritmia e sem alterações foi bem executada quando os índices I e $\bar{I}$ foram utilizados, pois as taxas de acerto chegaram a 100% (Tabela 6.10).

Por outro lado, a classificação de trechos com descargas multifocais e trechos sem alterações já não alcançou desempenho que chegasse aos 100% (Tabela 6.9). Nessa tarefa destaca-se o bom desempenho dos índices I e $\bar{I}$ em contraponto ao índice I^w . As taxas de classificação quando I^w foi utilizado demonstra uma dificuldade da DWT em diferenciar as descargas multifocais dos trechos normais.

Considerando todas as simulações, incluindo as do primeiro grupo de experimentos, o desafio maior se apresentou quando o objetivo foi classificar trechos com hirsarritmia e trechos com descargas multifocais, pois esses sinais apresentam características morfológicas bastante semelhantes. Para essa tarefa no entanto, ficou evidente a superioridade da metodologia que combina a decomposição em funções de Gabor sintonizadas com a definição do índice $\bar{I}$.

Observando os valores apresentados na Tabela 6.8 percebe-se que enquanto a utilização de I e I^w culminam em taxas de sensibilidade, especificidade e acurácia em torno de 75%, a aplicação de $\bar{I}$ se aproxima dos 90%. Esses resultados torna a decomposição proposta não apenas uma metodologia de análise de EEG's com hirsarritmia, mas uma possibilidade de um novo método de análise de sinais em geral.

Conclusões e Perspectivas

Algumas definições importantes, como epilepsia, hipsarritmia, transformada de Gabor e SVM foram abordadas no decorrer desse trabalho. Sobre epilepsia foram apresentadas, além da definição, as várias maneiras que a Liga Internacional Contra a Epilepsia classifica essa patologia. Em meio a classificação dos diversos tipos de epilepsia destacou-se as síndromes que normalmente ocorrem na infância como por exemplo a síndrome de West.

A principal característica da síndrome de West abordada nesse trabalho foi os espasmos infantis, caracterizados por mudanças no EEG conhecidas como hipsarritmia ou padrão hipsarrítmico. Apesar das características de hipsarritmia estarem bem definidas na literatura, verificou-se que o consenso entre especialistas sobre a presença ou ausência desse padrão é baixo, e esse fato foi o principal motivador da realização desse trabalho.

Apesar da característica de hipsarritmia ser intrínseca a patologias como a síndrome de West verificou-se recentemente a existência de uma relação tríplice Zika vírus, microcefalia e hipsarritmia. Foi considerando as novas descobertas sobre essa relação que se estabeleceu todo o desenvolvimento desse trabalho. Para a identificação do padrão de hipsarritmia nos sinais de EEG de lactentes com microcefalia foi realizada extração de características com base na proposta de definição e utilização de três índices matemáticos: um oriundo da DWT, outro definido a partir de funções de Gabor e um terceiro índice resultante de uma proposta de decomposição de sinais em funções de Gabor sintonizadas.

Os resultados obtidos pela aplicação dos três índices foram comparados e observou-se que, em geral, o terceiro índice tem melhor capacidade de capturar características de sinais e trechos de EEG's com padrão de hipsarritmia. Nesse sentido verifica-se que a utilização desse índice é um fator promissor na análise de EEG's. Além da possibilidade de utilização do índice $\bar{I}$ vale ressaltar que a metodologia de decomposição proposta também pode ser promissora no que tange a análise de qualquer tipo de sinais.

Diante das tarefas já realizadas nesse trabalho surgem alguns desafios para trabalhos futuros. Um desses desafios consiste em estabelecer o formalismo matemático do método de decomposição de sinais em funções de Gabor sintonizadas. Outro desafio é aplicar o método de decomposição em outros tipos de sinais e analisar a eficácia da metodologia em realizar representação esparsa.

Propriedades matemáticas das funções de Gabor sintonizadas também carecem de análises e esse fato é sem dúvida uma tarefa a ser investigada em trabalhos futuros. Entre essas

propriedades podem ser citadas:

- As funções de Gabor satisfazem as propriedades das wavelets?
- Se o produto interno entre uma wavelet e um sinal evidencia características de multifractalidade dos sinais, então as funções de Gabor sintonizadas também são capazes de evidenciar as mesmas características?

No estágio atual desse trabalho foi possível realizar a publicação do artigo intitulado “Identification of Hypsarrhythmia in Children with Microcephaly Infected by Zika Virus”. *Entropy* 2019, 21(3), 232; doi:10.3390/e21030232. Esse artigo encontra-se no apêndice para verificação.

Referências Bibliográficas

- ALTUNEL, A.; SEVER, A.; ALTUNEL, E. O. Hypsarrhythmia paroxysm index: A tool for early prediction of infantile spasms. *Epilepsy Research*, v. 111, p. 54 – 60, 2015. ISSN 0920-1211. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0920121115000157>>.
- BACHMAN, G.; NARICI, L. *Functional Analysis*. Academic Press, 1966. ISBN 9780486402512. Disponível em: <<http://store.doverpublications.com/0486402517.html>>.
- BELL, A. J.; SEJNOWSKI, T. J. The independent components of natural scenes are edge filters. *Vision Research*, v. 37, n. 23, p. 3327 – 3338, 1997. ISSN 0042-6989. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0042698997001211>>.
- BERG, A. T. et al. Revised terminology and concepts for organization of seizures and epilepsies: Report of the ilae commission on classification and terminology, 2005–2009. *Epilepsia*, v. 51, n. 4, p. 676–685, 2010. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1528-1167.2010.02522.x>>.
- BOYD, S.; VANDENBERGHE, L. *Convex optimization*. [S.l.]: Cambridge university press, 2004.
- BRINCIOTTI, M.; MATRICARDI, M. Epileptic encephalopathies of infancy and childhood. In: *Clinical Electroencephalography*. [S.l.]: Springer, 2019. p. 413–430.
- CARVALHO, M. D. C. G. et al. Sleep eeg patterns in infants with congenital zika virus syndrome. *Clinical Neurophysiology*, v. 128, n. 1, p. 204 – 214, 2017. ISSN 1388-2457. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1388245716306484>>.
- CASPERS, H. Electroencephalography: Basic principles, clinical applications and related fields. ernst niedermeyer , fernando lopes da silva. *The Quarterly Review of Biology*, v. 58, n. 2, p. 301–302, 1983. Disponível em: <<https://doi.org/10.1086/413356>>.
- CORREIA, L. C. *Codificação Eficiente para Caracterização e Eletroencefalograma de Pacientes Epilépticos*. 1–55 p. Dissertação (Dissertação de Mestrado) — Universidade Federal do Maranhão, 2018.
- CORREIA, L. C. et al. Learning efficient codes from epileptic eeg. 2017.
- CUKIERT, A. et al. Secondary bilateral synchrony due to fronto-mesial lesions - an invasive recording study. v. 57, p. 636–42, 10 1999.

FILHO, J. G. *APLICAÇÃO DE TÉCNICAS DE RECONFIGURAÇÃO DINÂMICA A PROJETO DE SVM*. 1–107 p. Dissertação (Dissertação de mestrado) — Escola Politécnica, 2010.

FISHER, R. S. et al. Epileptic seizures and epilepsy: Definitions proposed by the international league against epilepsy (ilae) and the international bureau for epilepsy (ibe). v. 46, p. 470–2, 05 2005.

GABOR, D. Theory of communication. part 1: The analysis of information. *Journal of the Institution of Electrical Engineers - Part III: Radio and Communication Engineering*, v. 93, p. 429–441(12), November 1946. ISSN 0367-7540. Disponível em: <<http://digital-library.theiet.org/content/journals/10.1049/ji-3-2.1946.0074>>.

GASTAUT, H. et al. Encéphalopathie myoclonique infantile avec hypsarythmie (syndrome de west) et sclérose tubéreuse de bourneville. *Journal of the Neurological Sciences*, v. 2, n. 2, p. 140 – 160, 1965. ISSN 0022-510X. Disponível em: <<http://www.sciencedirect.com/science/article/pii/0022510X65900778>>.

GATHERER, D.; KOHL, A. Zika virus: a previously slow pandemic spreads rapidly through the americas. *Journal of General Virology*, Microbiology Society, v. 97, n. 2, p. 269–273, 2016. ISSN 0022-1317. Disponível em: <<https://www.microbiologyresearch.org/content/journal/jgv/10.1099/jgv.0.000381>>.

GEERTS, A. *DETECTION OF INTERICTAL EPILEPTIFORM DISCHARGES IN EEG*. 1–68 p. Dissertação (Master Thesis in Applied Mathematics) — Faculty of Electrical Engineering, Mathematics and Computer Science, 2012.

GIBBS, F. A.; GIBBS, E. L. Atlas of electroencefalography. 1952.

GUPTA, V. et al. Automated detection of focal eeg signals using features extracted from flexible analytic wavelet transform. *Pattern Recognition Letters*, v. 94, p. 180 – 188, 2017. ISSN 0167-8655. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0167865517300934>>.

HAAR, A. Zur theorie der orthogonalen funktionensysteme. *Mathematische Annalen*, Springer, v. 69, n. 3, p. 331–371, 1910.

HAGA, Y. et al. Asymmetric spasms in west syndrome. *Journal of Epilepsy*, v. 8, n. 1, p. 61 – 67, 1995. ISSN 0896-6974. Disponível em: <<http://www.sciencedirect.com/science/article/pii/089669749400011N>>.

HAN, J.; DONG, F.; XU, Y. Y. Entropy feature extraction on flow pattern of gas/liquid two-phase flow based on cross-section measurement. v. 147, p. 012041, 03 2009.

HAUSER, W. A.; KURLAND, L. T. The epidemiology of epilepsy in rochester, minnesota, 1935 through 1967. *Epilepsia*, v. 16, n. 1, p. 1–66, 1975. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1528-1157.1975.tb04721.x>>.

HRACHOVY, R. A.; FROST, J. D.; KELLAWAY, P. Hypsarrhythmia: Variations on the theme. v. 25, p. 317–25, 07 1984.

HUMPHREY, A. et al. The early childhood epilepsy severity scale (e-chess). v. 79, p. 139–45, 05 2008.

HUSSAIN, S. A. et al. Hypsarrhythmia assessment exhibits poor interrater reliability: A threat to clinical trial validity. v. 56, 11 2014.

IVES, J.; GLOOR, P. Automatic nocturnal sleep sampling: A useful method in clinical electroencephalography. *Electroencephalography and Clinical Neurophysiology*, v. 43, n. 6, p. 880 – 884, 1977. ISSN 0013-4694. Disponível em: <<http://www.sciencedirect.com/science/article/pii/0013469477900116>>.

JANCA, R. et al. Detection of interictal epileptiform discharges using signal envelope distribution modelling: Application to epileptic and non-epileptic intracranial recordings. v. 28, 06 2014.

KANDA, P. et al. Sleep eeg of microcephaly in zika outbreak. *The Neurodiagnostic Journal*, v. 58, p. 11–29, 01 2018.

KASINATHAN, A. et al. Pseudo hypsarrhythmia: An early marker of angelman syndrome. *Annals of Indian Academy of Neurology*, Wolters Kluwer–Medknow Publications, v. 22, n. 3, p. 359, 2019.

KRAMER, U.; SUE, W. C.; MIKATI, M. Hypsarrhythmia: Frequency of variant patterns and correlation with etiology and outcome. v. 48, p. 197–203, 01 1997.

KWAN, P.; SANDER, J. W. The natural history of epilepsy: an epidemiological view. *Journal of Neurology, Neurosurgery & Psychiatry*, BMJ Publishing Group Ltd, v. 75, n. 10, p. 1376–1381, 2004. ISSN 0022-3050. Disponível em: <<http://jnnp.bmj.com/content/75/10/1376>>.

LIMA, J. G. C.; LONGO, R. H.; MOREIRA, M. H. Franco da R. Hipsarritmia: estudo clínico-eletrencefalográfico. v. 32, p. 303–309, 12 1974.

LINDEN, H. van der et al. Epilepsy profile in infants with congenital zika virus infection. *New England Journal of Medicine*, v. 379, n. 9, p. 891–892, 2018. PMID: 30157396. Disponível em: <<https://doi.org/10.1056/NEJMc1716070>>.

LOPES, M. V. et al. *ICA feature extraction for spike sorting of single-channel records*. 2013. 1-5 p.

LORENA, A. C.; CARVAHO, A. C. P. L. F. de. *Introdução às Máquinas de Vetores Suporte*. [S.l.]: Instituto de Ciências Matemáticas e de Computação, 2003. ISSN 0103-2569.

LUCENA, F. et al. Statistical coding and decoding of heartbeat intervals. In: *PloS one*. [S.l.: s.n.], 2011.

MARÇELJA, S. Mathematical description of the responses of simple cortical cells. *J. Opt. Soc. Am.*, OSA, v. 70, n. 11, p. 1297–1300, Nov 1980. Disponível em: <<http://www.osapublishing.org/abstract.cfm?URI=josa-70-11-1297>>.

MARCHI, L. R. D. et al. Epileptic spasms without hypsarrhythmia in infancy and childhood: Tonic spasms as a seizure type. v. 17, p. 188–193, 06 2015.

- MEYER, Y. *Wavelets and applications*. [S.l.]: Masson Paris, 1992. v. 31.
- MONTEIRO, R. Transformada de haar. *Universidade de Lisboa*, 2019.
- MORETTIN, P. A. *ONDAS E ONDALETAS: Da Análise de Fourier à Análise de Ondaletas de Séries Temporais*. edusp, 2014. ISBN 978-85-314-1478-7. Disponível em: <<https://www.ime.usp.br/~pam/ondas.html>>.
- MUIR, A. M. et al. Genetic heterogeneity in infantile spasms. *Epilepsy Research*, Elsevier, v. 156, p. 106181, 2019.
- MYTINGER, J. R. et al. Improving the inter-rater agreement of hypsarrhythmia using a simplified eeg grading scale for children with infantile spasms. *Epilepsy Research*, v. 116, p. 93 – 98, 2015. ISSN 0920-1211. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0920121115300309>>.
- MYTINGER, J. R.; WEBER, A.; VIDAURRE, J. High amplitude background slow waves in normal children aged 3 to 18 months - implications for the consideration of hypsarhythmia. v. 35, p. 1, 01 2018.
- NINAR, C. de A. *Governo do Estado do Maranhão*, 2018. Disponível em: <<https://www.institutoacqua.org.br/casa-de-apoio-ninar>>.
- OLIVEIRA, W.; CORTEZ-ESCALANTE, J.; OLIVEIRA, W. Increase in reported prevalence of microcephaly in infants born to women living in areas with confirmed zika virus transmission during the first trimester of pregnancy - brazil, 2015. *MMWR Morb Mortal Wkly Rep*, v. 65, p. 242–247, 03 2016.
- OPPENHEIM, A. V. *Sinais e sistemas*. [S.l.]: Prentice-Hall, 2010.
- PACOLA, E. R. et al. *RECONHECIMENTO DE ESPÍCULAS EM SINAIS DE EEG UTILIZANDO WAVELETS E LDA*. 2012.
- PESSOA, A. et al. Motor abnormalities and epilepsy in infants and children with evidence of congenital zika virus infection. *Pediatrics*, American Academy of Pediatrics, v. 141, n. Supplement 2, p. S167–S179, 2018. ISSN 0031-4005. Disponível em: <<https://pediatrics.aappublications.org/content/141/Supplement\2/S167>>.
- PHILIPPI, H. et al. Electroencephalographic evolution of hypsarrhythmia: Toward an early treatment option. v. 11, p. 1859–1864, 11 2008.
- RICHHARIYA, B.; TANVEER, M. Eeg signal classification using universum support vector machine. *Expert Systems with Applications*, v. 106, p. 169 – 182, 2018. ISSN 0957-4174. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0957417418302100>>.
- RUFINO, H. L. P. *Algoritmo de Aprendizado Baseado em Máquina de Vetores de Suporte-Uma contribuição Para o Reconhecimento de Dados Desbalanceados*. 1–107 p. Dissertação (Tese de Doutorado) — Universidade Federal de Uberlândia, 2011.
- RUSSO, F.; JUNGSMANN, P.; BRAGA, P. B. Zika infection and the development of neurological defects. *Cellular microbiology*, v. 19, 03 2017.

SAKAMOTO, A. C. Classificação das crises epiléticas. I, p. 13–20, 02 2007.

SANEI, S.; CHAMBERS., J. *EEG signal processing*. [S.l.]: British Library, 2007. ISBN 13 978-0-470-02581-9.

SCARAMELLI, A. et al. Prodromal symptoms in epileptic patients: Clinical characterization of the pre-ictal phase. *Seizure*, v. 18, n. 4, p. 246 – 250, 2009. ISSN 1059-1311. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S105913110800232X>>.

SEHGAL, R. et al. "prognostic utility of clinical epilepsy severity score versus pretreatment hypsarrhythmia scoring in children with west syndrome.". v. 48, 07 2016.

SHARMA, A.; KANKAR, P. Feature extraction and fault severity classification in ball bearings. v. 22, 04 2014.

SJÖLIN, P.; STRÖMBERG, J.-O. Basis properties of hardy spaces. *Arkiv för Matematik*, Springer, v. 21, n. 1, p. 111–125, 1983.

SMITH, R. J. *Computational characterization of the electroencephalogram in patients with infantile spasms syndrome*. Tese (Doutorado) — UC Irvine, 2019.

SMITH, R. J. et al. Long-range temporal correlations reflect treatment response in the electroencephalogram of patients with infantile spasms. *Brain Topography*, v. 30, n. 6, p. 810–821, Nov 2017. ISSN 1573-6792. Disponível em: <<https://doi.org/10.1007/s10548-017-0588-5>>.

SPENNER, B. et al. The prognostic value of sleep spindles in long-term outcome of west syndrome. *European Journal of Paediatric Neurology*, Elsevier, 2019.

STEPHANE, M. A wavelet tour of signal processing. *The Sparse Way*, Elsevier, 1999.

STÉPHANE, M. Chapter 1 - sparse representations. In: STÉPHANE, M. (Ed.). *A Wavelet Tour of Signal Processing (Third Edition)*. Third edition. Boston: Academic Press, 2009. p. 1 – 31. ISBN 978-0-12-374370-1. Disponível em: <<http://www.sciencedirect.com/science/article/pii/B9780123743701000057>>.

STÉPHANE, M. Chapter 7 - wavelet bases. In: STÉPHANE, M. (Ed.). *A Wavelet Tour of Signal Processing (Third Edition)*. Third edition. Boston: Academic Press, 2009. p. 263 – 376. ISBN 978-0-12-374370-1. Disponível em: <<http://www.sciencedirect.com/science/article/pii/B9780123743701000112>>.

SWEILAM, N. H.; THARWAT, A.; MONIEM, N. A. Support vector machine for diagnosis cancer disease: A comparative study. *Egyptian Informatics Journal*, v. 11, n. 2, p. 81 – 92, 2010. ISSN 1110-8665. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1110866510000241>>.

TAKAHASHI, A. *Máquina de Vetores de Suporte Intervalar*. 1–72 p. Dissertação (Tese de Doutorado) — Universidade Federal do Rio Grande do Norte, 2012.

THIJS, R. D. et al. Epilepsy in adults. *The Lancet*, v. 393, n. 10172, p. 689 – 701, 2019. ISSN 0140-6736. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0140673618325960>>.

TORREÃO, J. R. A. Generalized signal-tuned gabor approach for signal representation and analysis. *International Journal of Modern Physics C*, v. 28, n. 01, p. 1750001, 2017. Disponível em: <<https://www.worldscientific.com/doi/abs/10.1142/S0129183117500012>>.

TRAITRUENGSAKUL, S. et al. Automatic localization of epileptic spikes in eegs of children with infantile spasms. In: *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. [S.l.: s.n.], 2015. p. 6194–6197. ISSN 1094-687X.

TRAITRUENGSAKUL, S. et al. Developing a novel epileptic discharge localization algorithm for electroencephalogram infantile spasms during hypsarrhythmia. v. 55, p. 1659–1668, 09 2017.

WALTER, W. G. Slow potential waves in the human brain associated with expectancy, attention and decision. *Archiv für Psychiatrie und Nervenkrankheiten*, v. 206, n. 3, p. 309–322, May 1964. ISSN 1433-8491. Disponível em: <<https://doi.org/10.1007/BF00341700>>.

YACUBIAN, E. M. T.; KOCHEN, S. *Crises epilépticas*. Leitura Médica Ltda, 2014. ISBN 978-85-61125-46-2. Disponível em: <<http://epilepsia.org.br/wp-content/uploads/>>.

YANG, J. Do gabor functions provide appropriate descriptions of visual cortical receptive fields?: comment. *J. Opt. Soc. Am. A*, OSA, v. 9, n. 2, p. 334–336, Feb 1992. Disponível em: <<http://josaa.osa.org/abstract.cfm?URI=josaa-9-2-334>>.

ZANELLA, L. C. H. *Metodologia da pesquisa*. [S.l.]: SEAD/UFSC, 2006.

ZHANG, Y. et al. Prognostic utility of hypsarrhythmia scoring in children with west syndrome after ketogenic diet. *Clinical Neurology and Neurosurgery*, v. 184, p. 105402, 2019. ISSN 0303-8467. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0303846719301982>>.

ANEXO

The BASED* score – An EEG grading scale for children with infantile spasms

Goal: To use a simplified EEG grading scale to improve the inter-rater reliability of hypsarrhythmia.

This scale is founded on the following: The **burden of interictal epileptiform discharges** and the **background amplitudes** are important features of hypsarrhythmia and are perhaps the most objective (in that they can be measured).

***BASED:** *Burden of Amplitudes and Epileptiform Discharges*

BASED score	Description
0	Normal
1	Any definite nonepileptiform abnormality
2	<3 spike foci AND no common background slow waves $\geq 200 \mu\text{v}^{\#\Psi}$
3	MFS <50% of one second bins [^] AND no common background slow waves $\geq 200 \mu\text{v}^{\#\Psi}$, OR No MFS but common background slow waves $\geq 200 \mu\text{v}^{\#\Psi}$
4 (Hyps ^Ω)	MFS <50% of one second bins [^] AND common background slow waves $\geq 200 \mu\text{v}^{\#\Psi}$
5 (Hyps ^Ω)	MFS $\geq 50\%$ of one second bins [^] , OR Common background slow waves $\geq 300 \mu\text{v}^{\#\Phi}$ in 2 or more bilateral head regions

Table. The BASED Score.

MFS: multifocal spikes – at least three different foci; μv : microvolts

[^]The percentage of one second bins that include one or more spikes in the most severely abnormal five minute epoch (i.e. the epoch that gives the highest BASED score).

[#]Peak-to-peak amplitude on a longitudinal bipolar montage, refers to background slow waves and excludes 1) the slow wave associated with a preceding spike, 2) hypnagogic patterns, and 3) arousal rhythms

^Ψ: May be one or more head regions, must be a common finding, may regional (e.g. left posterior), and may exist in the presence of other lower amplitude background activities

^Φ: Must be two of: bilateral frontal, bilateral temporal, bilateral parietal, bilateral occipital; must be a common finding

^Ω: To determine the presence or absence of hypsarrhythmia, findings should be representative of the most severely abnormal five minute epoch of the study (i.e. the epoch that gives the highest BASED score); both a score of 4 or 5 suggest electrographic evidence of hypsarrhythmia

Patient Information: _____

Final Score (0, 1, 2, 3, 4, or 5): _____

***Answer all questions though skip question 4 if no MFS**

1) Is the background ≥ 300 microvolts in **2 or more bilateral head regions** (circle)? (yes) (no)

#Answer no if high amplitude background is just regional (e.g. posterior)

If yes, enter a Final Score of 5

2) Does the clip include $\geq 50\%$ multifocal spikes (circle)? (yes) (no)

(If needed, use the table)

If yes, enter a Final Score of 5

	15 sec with <50% MFS	15 sec with $\geq 50\%$ MFS
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		

	15 sec with <50% MFS	15 sec with $\geq 50\%$ MFS
11		
12		
13		
14		
15		
16		
17		
18		
19		
20		

3) Does clip include MFS (circle)? (yes) (no)

If yes, go to the next question

If no, skip question 4

4*) If MFS, does the clip include <50% MFS (circle)? (yes) (no)

Go to the next question

5) Does the background include common slow waves of ≥ 200 uv (circle)? (yes) (no)

(If you don't yet have a final score, use these questions to find one)

If yes, and your answer to 3 is yes, enter Final Score of 4

If yes, and your answer to 3 is no, enter Final Score of 3

If no, and your answer to 3 is yes, enter Final Score of 3

If no, and your answer to 3 is no, enter Final Score 0, 1 or 2 using the remainder of the tracing

Article

Identification of Hypsarrhythmia in Children with Microcephaly Infected by Zika Virus

Gean Carlos Sousa ^{1,*}, Claudio M. Queiroz ², Patrícia Sousa ³, Priscila Lima ¹, Antônio Augusto Moura da Silva ⁴, Nilviane Pires ⁵  and Allan Kardec Barros ¹ 

¹ Department of Electrical Engineering, Federal University of Maranhão (UFMA), São Luís-MA 65080-805, Brazil; priscilalima_rocha@hotmail.com (P.L.); akduailibe@gmail.com (A.K.B.)

² Brain Institute, Federal University of Rio Grande do Norte, Natal 59078-970, Brazil; clausqueiroz@neuro.ufrn.br

³ Department of Medicine, University Hospital of the Federal University of Maranhão, São Luís-MA 65080-805, Brazil; cdneuropatricia@gmail.com

⁴ Department of Public Health, Federal University of Maranhão, São Luís-MA 65080-805, Brazil; aamouradasilva@gmail.com

⁵ Program in Biotechnology, Center for Biological Sciences and Health, Federal University of Maranhão (UFMA), São Luís-MA 65080-805, Brazil; nilvipires@gmail.com

* Correspondence: gean.pib@gmail.com; Tel.: +55-98-98841-7717

Received: 11 January 2019; Accepted: 22 February 2019; Published: 28 February 2019



Abstract: Hypsarrhythmia is an electroencephalographic pattern specific to some epileptic syndromes that affect children under one year of age. The identification of this pattern, in some cases, causes disagreements between experts, which is worrisome since an inaccurate diagnosis can bring complications to the infant. Despite the difficulties in visually identifying hypsarrhythmia, options of computerized assistance are scarce. Aiming to collaborate with the recognition of this electropathological pattern, we propose in this paper a mathematical index that can help electroencephalography experts to identify hypsarrhythmia. We performed hypothesis tests that indicated significant differences in the groups under analysis, where the *p*-values were found to be extremely small.

Keywords: hypsarrhythmia; electroencephalographic; epileptic; identification; index

1. Introduction

The Congenital Zika Virus Infection can cause severe brain abnormalities and diverse electroencephalographic abnormalities [1]. Among these is hypsarrhythmia, an unusual activity of the electroencephalogram (EEG), found in a few types of epilepsy and known for interictal and spike-and-wave discharges in an irregular and disorganized background [2–6].

Identifying hypsarrhythmia patterns is challenging, and only experienced professionals in the reading of EEG tend to identify the presence of hypsarrhythmia distinctly [4,7,8]. This task tends to be even harder when the analysis is made in EEG exams of nursing infants infected by the Zika virus and with microcephaly, because, as is pointed out by Carvalho et al. (2017), these exams are consistently abnormal even in infants who have not yet developed epilepsy [1]. Additionally, brief and subtle seizures may not be detected, or can cause diagnostic errors [4].

Carvalho et al. (2017) found hypsarrhythmia in EEGs of newborns infected by the Zika Virus and with microcephaly [1]. The presence of hypsarrhythmia in EEGs in the sleep of children with microcephaly was also found by Kanda et al. (2018) [9]. It is worth mentioning that these works performed only the visual description of EEG signals due to the lack of mathematical and computational methods that can assist experts carrying out these types of tasks.

According to Roger (2015), the conventional systems of classification do not take into account all the ramifications of the impact of diagnostic errors [10]. For this reason, we propose a mathematical index capable of assisting the identification of hypsarrhythmia in a fast and efficient way. The mathematical features of this index are grounded on the morphological similarities that exist between a spike and a Gabor function. The main hypothesis for the definition of the index is that the inner product between a Gabor function and the parts of the signal that contain spikes always presents higher values than the products between the same Gabor and the rest of the signal.

The proposed methodology tends to provide evidence for the differences between electroencephalograms with a hypsarrhythmia pattern and normal electroencephalograms with normal tracings, because according to Mytinger et al. (2018), electroencephalographic tracings with hypsarrhythmia have a greater number of spikes than the normal tracings [11].

2. Materials and Methods

2.1. Database

The database used was composed by EEGs of 30 children with the Congenital Zika Syndrome, with a total of 203 EEGs of 5 minutes each (98 with traces of hypsarrhythmia and 105 without any abnormalities). These signals were obtained by the government institution “Casa Ninar” [12], which offers medical assistance to children with microcephaly. The children assisted in the institution have different electroencephalographic abnormalities and, in some cases, have epileptic crises. Clinically, these crises are marked by infantile spasms linked with the presence of hypsarrhythmia. The electroencephalographic records were collected with electrodes positioned in accordance with the 10–20 system during periods of spontaneous sleep. The software Neuromap was used to collect data, and the signals were sampled at 128 HZ. An expert identified the presence or absence of hypsarrhythmia on the EEGs. Consent for the use of the collected data was obtained by means of the Ethics Committee of the Federal University of Maranhão, under the registration code CAAE 65897317.1.0000.5086.

2.2. Hypsarrhythmia Index

A large number of spikes in an electroencephalographic record can be characterized as a hypsarrhythmic record [7]. Spikes are electrographic markers where its quantification is used in the surgical planning of patients with epilepsy. A spike is clinically defined as a transitory acute that lasts between 20 and 70 ms and that is clearly distinguishable from background activity [13–16].

The index proposed in this paper is based on an inner product between an EEG and a Gabor function, by means of the discrete wavelet transform. A Gabor function is a sinusoidal function modulated by a Gaussian function [17–20] (1).

$$\varphi(t) = Ae^{-\frac{(t-\xi)^2}{2\sigma^2}} \cos[(t-\xi)\omega + \phi] \quad (1)$$

The parameters ξ , σ , ω , and ϕ are, respectively, the average and the standard deviation of the Gaussian envelope, the frequency, and the cosine phase. By varying the parameters, it is possible to achieve an infinity of waves.

The Discrete Wavelet Transform (DWT) is the inner product between a signal X_n and a Wavelet $\psi_{a,b}(n-b)$, moving $\psi_{a,b}(n-b)$ through the parameter b [21,22]. The DWT is given as follows (2):

$$DWT = \sum_{n=1}^N X_n \psi_{a,b}(n-b). \quad (2)$$

The function $\psi_{a,b}(n-b)$ used in our experiments will be the Gabor function. The Gabor functions were chosen as the core of the DWT, since they are morphologically similar to the waves known as

spikes. The DWT will be applied in windows of w seconds of EEG and after that, the energy of the coefficients of this transformation will be analyzed. Each window assessed corresponds mathematically to a Matrix A_k of dimensions $m \times n$, in which m is the number of channels and n is the number of samples (Figure 1).

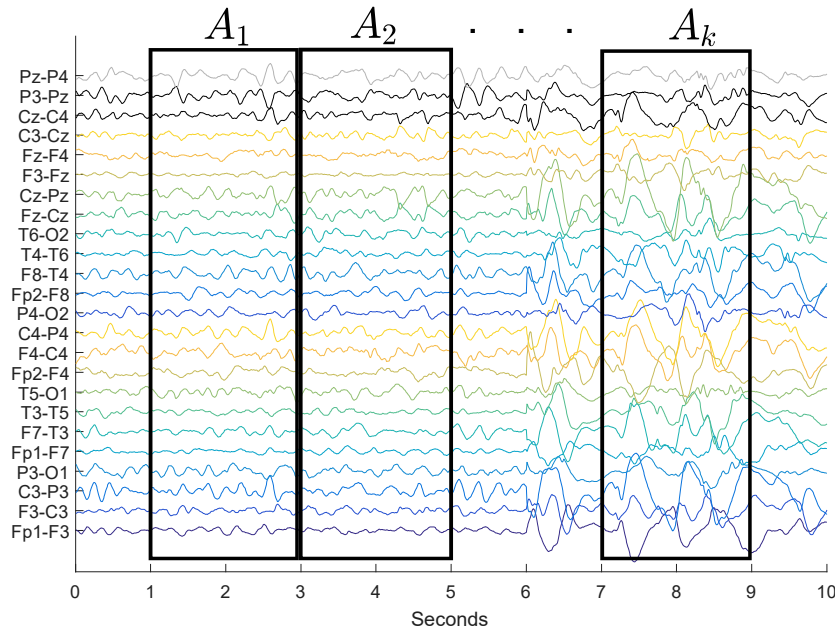


Figure 1. Windowing of the EEG's signal in A_k matrices.

Definition 1 (Hypsarrhythmia Index). Given that $\alpha_k = \{\alpha_{i1}, \alpha_{i2}, \alpha_{i3}, \dots, \alpha_{in}\}$ the set of coefficients of the DWT of the i -th matrix line $A_k(m \times n)$. We defined the hypsarrhythmia index as follows (3):

$$I_k = \frac{\sum_{i=1}^m \sum_{j=1}^n |\alpha_{ij}|}{m \cdot n} \quad (3)$$

Thus, for each window of w seconds of the EEG signal, we get an I_k index, and consequently, for a signal of K -windows, we get a $\beta = \{I_1, I_2, I_3, \dots, I_K\}$ sequence.

2.3. Feature Extraction and Classification

The process of feature extraction proposed in this paper was accomplished in three stages. First, the EEG went through a signal windowing (Figure 1). Afterward, the hypsarrhythmia index was calculated in each window, thus obtaining a β sequence of indexes. Finally, the average and the log energy entropy of the sequence were calculated. The log energy entropy is used to assess the set of indexes of a signal, and is defined as follows [23–25]:

$$E_{LgEn} = \sum_{i=1}^N \log(x_i^2), \quad (4)$$

In which x_i is the i -th signal sample of length N . The sequence of indexes $\beta = \{I_1, I_2, I_3, \dots, I_K\}$ from an EEG signal to compute the log energy entropy (4). A summary of the feature extraction process can be seen in Figure 2.

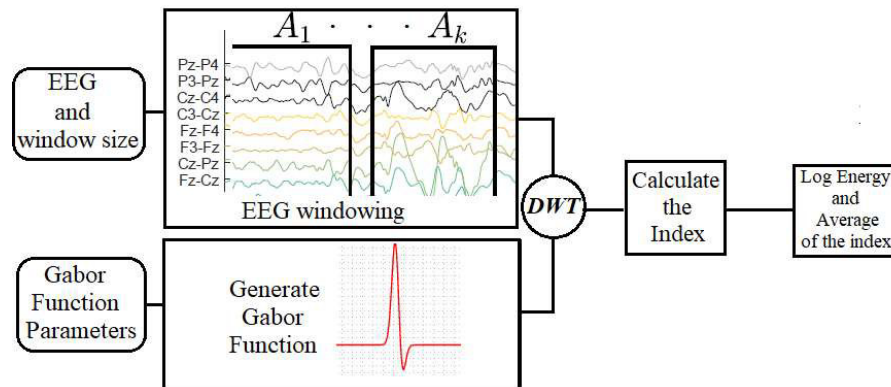


Figure 2. Algorithm for feature extraction.

3. Results

In our simulations, we used 203 electroencephalogram signals of children with microcephaly. These EEGs had a length of 5 min each, and were segmented in 300 windows of 10 s. We performed experiments with windows of 3 and 20 s, and we observed that there were not any significant differences in the results. For each window, an index I_k was calculated in Equation (3), composing a sequence of 300 indexes for each EEG. The average μ and the log energy entropy E_{LgEn} of each sequence were calculated, considering different parameter combinations of the Gabor functions.

Many computational experiments were made in order to identify the EEGs with hypsarrhythmic patterns. The boxplot presented in Figure 3 had the distributions of averages μ of the indexes for the EEGs groups with hypsarrhythmia and the EEGs with tracing without abnormalities. In these charts, it was observed that, on average, the indexes of the EEGs with hypsarrhythmia are bigger than the indexes with normal tracings.

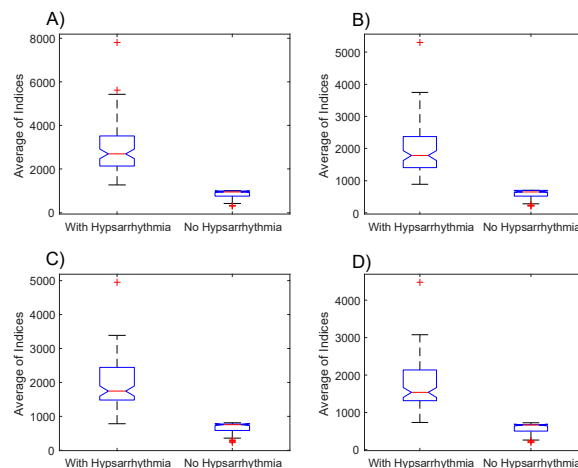


Figure 3. Boxplot with the averages of the EEGs' indexes with and without hypsarrhythmia. Indexes generated from the Gabor function with the parameters: (A) $\sigma^2 = 20$, $\omega = 1.02$ and $\phi = 5$, (B) $\sigma^2 = 5$, $\omega = 1.40$ and $\phi = 5$, (C) $\sigma^2 = 15$, $\omega = 2.68$ and $\phi = 5$, (D) $\sigma^2 = 10$, $\omega = 3.20$ and $\phi = 5$.

From the data presented in the boxplots of Figure 3, we performed two hypothesis tests. In the first test (Kolmogorov–Smirnov), we verified the null hypothesis that the data comes from a normal

pattern distribution. In this test, the result is 1 if the test rejects the null hypothesis in a significance level of 5%. The results of this test indicated the rejection of the null hypothesis in all datasets.

After we observed the results obtained through the Kolmogorov–Smirnov test, we decided to perform the Mann–Whitney test, in order to test the null hypothesis that the data were samples of continuous distributions with equal medians, against the alternative that they were not. The logical result 1 indicates a null hypothesis rejection and 0 indicates a failure to reject the null hypothesis at a significance level of 5%.

The results obtained with the Mann–Whitney test indicated that the null hypothesis must be rejected in all the comparison situations presented in Figure 4, with the following values of p -value: $p = 9.68 \times 10^{-34}$ (3A), $p = 9.67 \times 10^{-34}$ (3B), $p = 1.12 \times 10^{-33}$ (3C) and $p = 9.67 \times 10^{-34}$ (3D).

The boxplot of Figure 4 shows the distribution of the log energy entropy E_{LgEn} of the indexes for the EEGs groups with hypsarrhythmia and the EEGs with tracings without abnormalities. A large part of EEGs with a hypsarrhythmic pattern have log energy entropies of indexes bigger than those of the EEGs without abnormalities. This fact is similar to the results obtained with the μ averages.

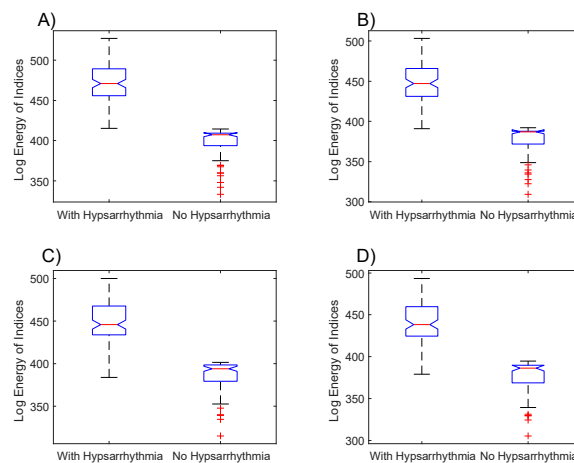


Figure 4. Boxplot with the log energy entropies of the EEGs' indexes with and without hypsarrhythmia. Indexes generated from the Gabor function with the parameters: (A) $\sigma^2 = 20$, $\omega = 1.02$ and $\phi = 5$, (B) $\sigma^2 = 5$, $\omega = 1.40$ and $\phi = 5$, (C) $\sigma^2 = 15$, $\omega = 2.68$ and $\phi = 5$, (D) $\sigma^2 = 10$, $\omega = 3.20$ and $\phi = 5$.

The sequence of application of the hypothesis tests applied to the data presented in Figure 4 was repeated in the data of the log energy entropy. The results obtained with the Kolmogorov–Smirnov test were the same as those previously obtained. The results obtained with the Mann–Whitney test denote that the null hypothesis must be rejected in all comparison situations presented in Figure 5, with the following values of p -value: $p = 9.77 \times 10^{-34}$ (4A), $p = 3.71 \times 10^{-33}$ (4B), $p = 4.84 \times 10^{-32}$ (4C), $p = 3.18 \times 10^{-32}$ (4D).

Figure 5 shows the dispersion of the values of the average μ and the log energy entropy E_{LgEn} of the indexes of each EEG, both for the group with hypsarrhythmia and for the group without abnormalities.

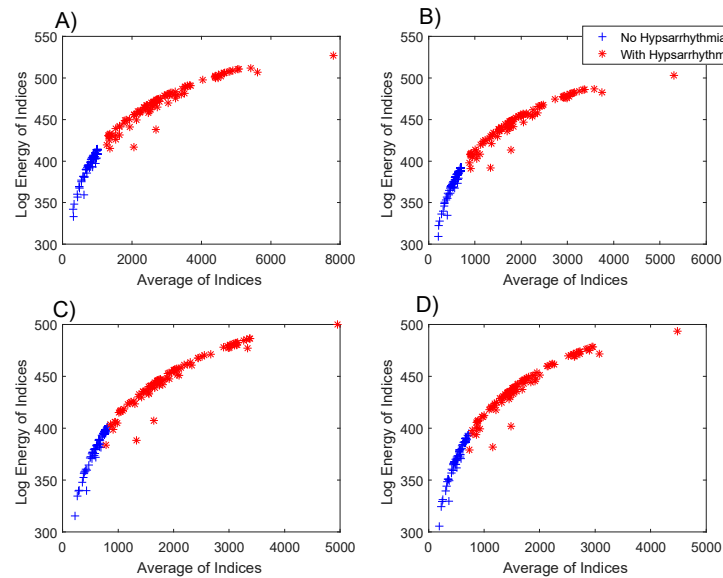


Figure 5. Dispersion of the average features and the log energy entropy of the EEGs with and without hypsarrhythmia. Features obtained from the Gabor function with the parameters: (A) $\sigma^2 = 20$, $\omega = 1.02$ and $\phi = 5$, (B) $\sigma^2 = 5$, $\omega = 1.40$ and $\phi = 5$, (C) $\sigma^2 = 15$, $\omega = 2.68$ and $\phi = 5$, (D) $\sigma^2 = 10$, $\omega = 3.20$ and $\phi = 5$.

4. Discussion

In this paper, we proposed a methodology that aims to identify hypsarrhythmia patterns in EEGs. This methodology consisted of the mathematical elaboration of an index capable of highlighting significant differences between EEG signals with normal tracings and those with characteristics of hypsarrhythmia. In an initial analysis, it was possible to verify that there are important differences between the medians and the indexes' entropies of the two groups under analysis, even when the parameters of the Gabor function are modified in Figure 3.

The results from the box plots of the indexes' averages (Figure 3) and of the log energy entropy (Figure 4) were very similar, but the combination of these two metrics (Figure 5) resulted in a possible linear separation between the groups being analyzed.

The results presented in Figure 3 confirm the hypotheses suggested at the beginning of this paper that the index values of the electroencephalograms with hypsarrhythmia are, in general, bigger than the ones of encephalograms without any abnormalities. It is possible to observe, by comparing the boxplots of Figure 4A,B with the ones of Figure 4C,D, that the changes in the Gabor function's parameters caused important implications for the results. The hypothesis tests applied showed that the proposed index is able to highlight the statistically significant differences, which enables this methodology to be used as a form of triage in the daily routine of the neuroscientists investigating epileptic syndromes in infants.

The proposed index can be easily implemented since it is based on the combination of mathematical operations already well-established in the literature, such as, for example, the median, the energy, and the DWT. This fact makes the computational calculation of the index rather quick.

It is relevant to emphasize that the parameters of the Gabor function chosen (based on the morphological characteristics of the spikes), highlight, in a mathematical way, one of the physiological features of the children with epilepsy. This feature consists of a large number of electric discharges in the shape of very peculiar waves (spikes).

5. Conclusions

Although it is important to discuss and elaborate on computational methods capable of performing the identification of hypsarrhythmia, it is also valid to emphasize that the existence of a concentration of kids (such as the ones assisted by the Casa Ninar) with this pathology is not common, and this can justify the lack of studies under this scope. However, the correct identification of hypsarrhythmia is an important step for the treatment of infant patients with refractory epilepsy, and the use of an index capable of assisting the electroencephalography expert is an important fact.

The findings confirm the hypotheses that the Gabor functions can enhance important differences between the groups in question. These results also suggest that the methodology of hypsarrhythmia's identification based on an index is able to distinguish signals without abnormalities from signals with hypsarrhythmic characteristics.

The proposed index can be easily implemented and used as a screening factor, since it consists of a metric capable of summarizing in a single number (or in a sequence of indexes) the characteristics that indicate the presence or absence of hypsarrhythmia. It is important to stress that the proposed method aims only to assist the experts involved.

Author Contributions: Conceptualization, G.C.S. and N.P.; Methodology, G.C.S. and P.L.; Software, G.C.S.; Validation, G.C.S., C.M.Q. and P.S.; Formal Analysis, A.K.B.; Investigation, P.S.; Resources, G.C.S.; Data Curation, C.M.Q., P.S. and A.A.M.d.S.; Writing—Original Draft Preparation, G.C.S.; Writing—Review & Editing, G.C.S.; Visualization, A.K.B.; Supervision, C.M.Q.

Funding: This research received no external funding.

Acknowledgments: The Foundation for Research and Scientific and Technological Development of Maranhão (FAPEMA).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Carvalho, M.D.C.G.; de Barros Miranda-Filho, D.; van der Linden, V.; Sobral, P.F.; Ramos, R.C.F.; Rocha, M.A.W.; Cordeiro, M.T.; de Alencar, S.P.; Nunes, M.L. Sleep EEG patterns in infants with congenital Zika virus syndrome. *Clin. Neurophysiol.* **2017**, *128*, 204–214. [[CrossRef](#)] [[PubMed](#)]
2. Gibbs, F.A.; Gibbs, E.L. *Atlas of Electroencefalography*; Addison-Wesley: Reading, MA, USA, 1952.
3. Millichap, J.; Millichap, J. Hypsarrhythmia or Hypsarrhythmia? *Pediatr. Neurol. Briefs* **2015**, *29*, 64. [[CrossRef](#)] [[PubMed](#)]
4. Hussain, S.A.; Kwong, G.; Millichap, J.; Mytinger, R.J.; Ryan, N.; Matsumoto, J.; Wu, Y.J.; Lerner, J.T.; Sankar, R. Hypsarrhythmia assessment exhibits poor interrater reliability: A threat to clinical trial validity. *Epilepsia* **2015**, *56*, 77–81. [[CrossRef](#)] [[PubMed](#)]
5. Altunel, A.; Sever, A.; Altunel, E.Ö. Hypsarrhythmia paroxysm index: A tool for early prediction of infantile spasms. *Epilepsy Res.* **2015**, *111*, 54–60. [[CrossRef](#)] [[PubMed](#)]
6. Marchi, L.R.; Seraphim, E.A.; Corso, J.T.; Naves, P.V.; Carvalho, K.C.; Ramirez, M.D.; Ferrari-Marinho, T.; Guaranha, M.S.; Yacubian, E.M. Epileptic spasms without hypsarrhythmia in infancy and childhood: Tonic spasms as a seizure type. *Epileptic Disord.* **2015**, *17*, 188–193. [[PubMed](#)]
7. Mytinger, J.R.; Hussain, S.A.; Islam, M.P.; Millichap, J.J.; Patel, A.D.; Ryan, N.R.; Twanow, J.D.E.; Heyer, G.L. Improving the inter-rater agreement of hypsarrhythmia using a simplified EEG grading scale for children with infantile spasms. *Epilepsy Res.* **2015**, *116*, 93–98. [[CrossRef](#)] [[PubMed](#)]
8. Smith, R.J.; Sugijoto, A.; Rismanchi, N.; Hussain, S.A.; Shrey, D.W.; Lopour, B.A. Long-Range Temporal Correlations Reflect Treatment Response in the Electroencephalogram of Patients with Infantile Spasms. *Brain Topogr.* **2017**, *30*, 810–821. [[CrossRef](#)] [[PubMed](#)]
9. Kanda, P.A.M.; de Almeida Xavier Aguiar, A.; Miranda, J.L.; Falcao, A.L.; Andrade, C.S.; dos Santos Reis, L.N.; Almeida, E.W.R.B.; Bello, Y.B.; Monfredinho, A.; Kanda, R.G. Sleep EEG of Microcephaly in Zika Outbreak. *Neurodiagn. J.* **2018**, *58*, 11–29. [[CrossRef](#)] [[PubMed](#)]

10. Rodger, J.A. Discovery of medical Big Data analytics: Improving the prediction of traumatic brain injury survival rates by data mining Patient Informatics Processing Software Hybrid Hadoop Hive. *Inform. Med. Unlocked* **2015**, *1*, 17–26. [CrossRef]
11. Mytinger, J.R.; Weber, A.; Vidaurre, J. High Amplitude Background Slow Waves in Normal Children Aged 3 to 18 Months: Implications for the Consideration of Hypsarhythmia. *J. Clin. Neurophysiol.* **2018**, *35*, 151–154. [CrossRef] [PubMed]
12. Casa de Apoio Ninar. Government of the State of Maranhão. Available online: <http://www.institutoacqua.org.br/casa-de-apoio-ninar> (accessed on 27 February 2019).
13. Traitruangsakul, S.; Seltzer, L.E.; Paciorkowski, A.R.; Ghoraani, B. Automatic localization of epileptic spikes in eegs of children with infantile spasms. In Proceedings of the 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Milan, Italy, 25–29 August 2015; pp. 6194–6197.
14. Traitruangsakul, S.; Seltzer, L.E.; Paciorkowski, A.R.; Ghoraani, B. Developing a novel epileptic discharge localization algorithm for electroencephalogram infantile spasms during hypsarrhythmia. *Med. Biol. Eng. Comput.* **2017**, *55*, 1659–1668. [CrossRef]
15. Ives, J.; Gloor, P. Automatic nocturnal sleep sampling: A useful method in clinical electroencephalography. *Electroencephalogr. Clin. Neurophysiol.* **1977**, *43*, 880–884. [CrossRef]
16. Caspers, H. Electroencephalography: Basic Principles, Clinical Applications and Related Fields. Ernst Niedermeyer, Fernando Lopes da Silva. *Q. Rev. Biol.* **1983**, *58*, 301–302. [CrossRef]
17. Gabor, D. Theory of communication. Part 1: The analysis of information. *J. Inst. Electr. Eng. Part III Radio Commun. Eng.* **1946**, *93*, 429–441. [CrossRef]
18. Marčelja, S. Mathematical description of the responses of simple cortical cells*. *J. Opt. Soc. Am.* **1980**, *70*, 1297–1300. [CrossRef] [PubMed]
19. Yang, J. Do Gabor functions provide appropriate descriptions of visual cortical receptive fields? Comment. *J. Opt. Soc. Am. A* **1992**, *9*, 334–336. [CrossRef]
20. Torreão, J.R.A. Generalized signal-tuned Gabor approach for signal representation and analysis. *Int. J. Mod. Phys. C* **2017**, *28*, 1750001. [CrossRef]
21. Stéphane, M. CHAPTER 7—Wavelet Bases. In *A Wavelet Tour of Signal Processing*, 3rd ed.; Stéphane, M., Ed.; Academic Press: Boston, MA, USA, 2009; pp. 263–376.
22. Morettin, P.A. *ONDAS E ONDALETAS: Da Análise de Fourier à Análise de Ondas de Séries Temporais*; Edups: Sao Paulo, Brazil, 2014.
23. Han, J.; Dong, F.; Xu, Y.Y. Entropy feature extraction on flow pattern of gas/liquid two-phase flow based on cross-section measurement. *J. Phys.* **2009**, *147*, 012041. [CrossRef]
24. Sharma, A.; Kankar, P. Feature extraction and fault severity classification in ball bearings. *J. Vib. Control* **2014**, *22*, 176–192. [CrossRef]
25. Gupta, V.; Priya, T.; Yadav, A.K.; Pachori, R.B.; Acharya, U.R. Automated detection of focal EEG signals using features extracted from flexible analytic wavelet transform. *Pattern Recognit. Lett.* **2017**, *94*, 180–188. [CrossRef]

